

UNIVERSIDAD DE CANTABRIA

PROGRAMA DE DOCTORADO EN
ciencia y tecnología



TESIS DOCTORAL

BÚSQUEDA DE LA PRODUCCIÓN DE PAREJAS DE CHARGINOS Y
SQUARKS TOP EN ESTADOS FINALES CON DOS LEPTONES DE
CARGA OPUESTA EN COLISIONES PROTÓN-PROTÓN A $\sqrt{s} = 13$ TeV

PH.D. THESIS

SEARCHES FOR PAIR PRODUCTION OF CHARGINOS AND TOP
SQUARKS IN FINAL STATES WITH TWO OPPOSITELY CHARGED
LEPTONS IN PROTON-PROTON COLLISIONS AT $\sqrt{s} = 13$ TeV

Realizada por:

Bárbara Chazin Quero

Dirigida por:

Dr. Luca Scodellaro

Dr. Jónatan Piedra Gómez

Escuela de Doctorado de la Universidad de Cantabria

Santander 2019

Acknowledgements

En primer lugar quiero agradecer a Francisco Matorras el haberme dado la oportunidad de realizar este doctorado el cual me ha permitido adentrarme un poquito en la física de partículas. A Jesús Marco de Lucas por tutorizar mis primeros pasos de esta pequeña aventura y en la cual siempre ha estado presente. Quiero expresar en especial mi agradecimiento a Luca Scodellaro, director de esta tesis, que con su paciencia, buen juicio y su sentido del humor ha hecho posible este trabajo. También destacar la ayuda brindada por Jónatan Piedra, codirector de esta tesis, con quien he podido contar en cada paso del camino y sin el cuál todo esto hubiera sido mucho más difícil. No hubiera podido realizar este doctorado sin la colaboración del grupo de física de altas energías del IFCA. En particular quiero agradecer al grupo de análisis: Alicia, Rocío, Teresa, Celso y Parbol, su esfuerzo y dedicación, el apoyo recibido y todos los buenos consejos que me han dado. Mencionar la suerte que he tenido de rodearme todo este tiempo de varios compañeros de doctorado, Juan y Nicolò ya doctores, Cédric, Andrea, Pedro, Celia y Pablo, en camino de serlo. Todos ellos personas de talento y con un gran futuro en la física de partículas.

Quiero también dar las gracias en general a todos los miembros del IFCA, con los que he compartido mi estrés, unas risas, las charletas de café y cervezas, o las comidas diarias en la sala Isaac Newton. Nunca podré agradecer lo suficiente a Aída y a Miguel Ángel, quienes siempre han estado ahí, y me han cubierto las espaldas cuando algo se torcía en el último momento y de la manera más rara. No quisiera olvidar, por supuesto, a Lara, quien siempre me hace reír a carcajadas y me ha escuchado y animado en muchos momentos.

No tengo suficientes palabras de agradecimiento para dedicarle a mi familia a la que valoro y estimo por encima de todo y quienes han aportado a este trabajo más de lo que pueden imaginar. Por último, quiero dar las gracias a Daniel, a quien en realidad no puedo expresar la infinidad de razones por las que le estoy profundamente agradecida, ya que es la luz que ilumina mi vida.

Contents

Acknowledgements	iii
1 Introduction	1
2 The Standard Model and supersymmetry	3
2.1 The Standard Model	3
2.1.1 The Brout-Englert-Higgs mechanism	8
The Standard Model Lagrangian	13
2.1.2 Standard Model shortcomings	13
2.2 Supersymmetry	16
2.2.1 Minimal Supersymmetric Standard Model	17
2.2.2 Simplified Supersymmetric Models	22
2.2.3 The direct search for supersymmetric particles in this thesis . .	23
2.3 Physics at LHC	24
3 Experimental environment	27
3.1 The European Organization for Nuclear Research	27
3.2 The Large Hadron Collider	28
3.2.1 Proton acceleration process	29
3.2.2 LHC performance	32
3.3 The Compact Muon Solenoid	34
3.3.1 Coordinate system convention	37
3.3.2 Magnet	38
3.3.3 Inner tracker detector	40
3.3.4 Electromagnetic calorimeter	42
3.3.5 Hadronic calorimeter	44
3.3.6 Muons system	47
3.3.7 Trigger and data acquisition system	49
3.3.8 Luminosity measurement	53
4 Analysis framework	55
4.1 Event data model and data tiers	55
4.2 Event reconstruction	57
4.2.1 Particle-flow event reconstruction	58
4.2.2 Electrons and isolated photons reconstruction	58
4.2.3 Muon reconstruction	59
4.2.4 Hadrons and non-isolated photons reconstruction	61
4.2.5 Lepton isolation	61
4.2.6 Pileup interactions and primary vertex	64
4.3 Monte Carlo simulation	66
4.3.1 Monte Carlo event generator	67
4.3.2 The Monte Carlo simulation in CMS	70

5	Analysis physics objects	73
5.1	Physics objects identification	73
5.1.1	Electron identification	73
	Electron selection	75
5.1.2	Muon identification	76
	Muon selection	78
5.1.3	Jet reconstruction	79
5.1.4	Identification of b-flavour jets	83
	Algorithms for heavy-flavour jets identification	84
5.1.5	Missing transverse momentum	87
5.2	Data samples	89
5.2.1	Triggers	90
5.3	Monte Carlo simulated samples	91
6	Event selection and background estimation	95
6.1	Event selection	95
6.2	Main backgrounds in the dilepton channel	96
6.3	Analysis strategy and signal region definition	103
6.4	Background estimation	110
6.4.1	Modeling of $m_{T2}(ll)$ in $t\bar{t}$, tW , and WW events	111
6.4.2	Estimation of $t\bar{t}Z$ background	116
6.4.3	Estimation of WZ background	117
6.4.4	Estimation of ZZ background	119
6.4.5	Estimation of Z boson background	122
6.5	Systematic uncertainties	124
7	Results and interpretation	129
7.1	Signal extraction	129
7.2	Results	130
7.2.1	Pull of nuisance parameters	131
7.3	Interpretation of results	133
8	Conclusions	141
9	Resumen	143
9.1	El gran colisionador de hadrones	145
9.2	El detector CMS	146
9.3	Reconstrucción de eventos	147
9.4	Identificación y selección de objetos	148
9.5	Selección de eventos	149
9.6	Simulaciones de Monte Carlo	150
9.7	Estrategia del análisis	151
9.8	Estimación de los fondos	153
9.9	Incertidumbres sistemáticas	157
9.10	Resultados e interpretación	158
A	Data and Monte Carlo simulated samples	163
B	Impact of the jet rate modeling on the chargino search	167
C	Tables of systematic uncertainties	171

D Pull of the nuisance parameters**177**

List of Figures

2.1	Elementary particles of the Standard Model	4
2.2	Shape of the Higgs scalar potential	10
2.3	Lowest order Feynman diagram for three Higgs boson decays	15
2.4	One-loop corrections to the Higgs boson mass	15
2.5	Theoretical production cross sections for selected supersymmetric processes	22
2.6	Feynman diagrams of the chargino pair production	23
2.7	Feynman diagrams of the top squark pair production	24
2.8	Schematic illustration of a proton-proton collision in the LHC	25
2.9	Example of PDFs calculated by the MSTW group	26
3.1	LHC layout	29
3.2	CERN accelerator complex	30
3.3	Peak luminosity evolution between 2011 and 2016	33
3.4	Integrated luminosity between 2011 and 2016	34
3.5	Integrated luminosity recorded by CMS during 2016	35
3.6	CMS detector	36
3.7	Transverse section of the CMS detector	37
3.8	CMS pixel tracker detector	41
3.9	CMS inner tracker detector	41
3.10	Tracking efficiency for muons in 2016	42
3.11	CMS calorimeter and tracker systems	44
3.12	The ECAL detector performance with the first recorded data in 2016	45
3.13	CMS hadronic calorimeter	47
3.14	Missing transverse momentum resolution for 2016 data	48
3.15	Muon momentum resolution using the tracker and the muon systems	49
3.16	CMS muons detector	50
3.17	Performance reconstruction of muons in 2016 data	51
3.18	CMS Level-1 trigger system for Run2	51
4.1	The Data flow tiers	57
4.2	Schematic illustration of the lepton isolation cone	62
4.3	Tight muons isolation efficiency for 2016 data	64
4.4	Relative charged hadron isolation for electrons in the ECAL for 2016 data	65
4.5	Relative electromagnetic hadron isolation for electrons in the ECAL for 2016 data	66
4.6	Relative neutral hadron isolation for electrons in the ECAL for 2016 data	67
4.7	Mean number of interactions per bunch crossing during the 2016 data taking	68
4.8	Sherpa simulation of a high energy hadron-hadron collision	69

5.1	Illustration of a track impact parameter	74
5.2	Electron identification efficiency for 2016 data	76
5.3	Electron selection efficiency for this analysis	77
5.4	Muon identification efficiency for 2016 data	79
5.5	Muon selection efficiency for this analysis	80
5.6	Simulated jet response as a function of the pseudorapidity	81
5.7	PF jet energy composition for different periods of the 2016 data taking	82
5.8	Recorded event by CMS detector in 2016	83
5.9	Schematic illustration of a heavy-flavour jet with secondary vertex	84
5.10	Recorded event by CMS detector in 2016	85
5.11	Jet tagging efficiency for various b-tagging algorithms of Run2	86
5.12	Geometrical schema of an event with a transverse energy imbalance	88
5.13	Distribution of missing transverse energy for events with two leptons in 2016 data	89
5.14	Distribution of missing transverse energy for events with photon and jets in 2016 data	90
6.1	Feynman diagrams for the $t\bar{t}$ production in proton-proton collisions	97
6.2	Feynman diagrams for the decay chain of a $t\bar{t}$ pair	97
6.3	Feynman diagrams for the single top quark production in proton-proton collisions	98
6.4	Feynman diagrams for the $t\bar{t}Z$ and $t\bar{t}W^+$ productions in proton-proton collisions	98
6.5	Feynman diagrams for the Z boson production in proton-proton collisions	99
6.6	Feynman diagram for the Z boson fermionic decays	100
6.7	Feynman diagrams for the WW production in proton-proton collisions	100
6.8	Feynman diagrams for the ZZ production in proton-proton collisions	101
6.9	Feynman diagrams for the WZ production in proton-proton collisions	101
6.10	Feynman diagrams for the triboson production in proton-proton collisions	101
6.11	Feynman diagrams for the triboson production in proton-proton collisions	102
6.12	Feynman diagrams for the Higgs boson production in proton-proton collisions	103
6.13	Feynman diagrams for the W +jets production in proton-proton collisions	103
6.14	Feynman diagrams for the semileptonic decay chain of a $t\bar{t}$ pair	104
6.15	Distributions of p_T^{miss} in events with two OC leptons passing the Z veto selection	106
6.16	Distributions of the $m_{T2}(ll)$ in events with two OC leptons passing the Z veto selection and with $p_T^{miss} > 140$ GeV	107
6.17	Distribution of $m_{T2}(ll)$ vs p_T^{miss} for $t\bar{t}$ simulated events with two leptons	108
6.18	Distributions of some discriminatory observables	109
6.19	Distribution of $m_{T2}(ll)$ for events with a b-tagged jet in the validation region $100 < p_T^{miss} < 140$ GeV	112
6.20	Distribution of $m_{T2}(ll)$ for events without b-tagged jet and no jets in the validation region $100 < p_T^{miss} < 140$ GeV	112
6.21	Distribution of $m_{T2}(ll)$ for events with three leptons and b-tagged jet veto in the validation region $p_T^{miss} > 140$ GeV	113
6.22	Comparison of $m_{T2}(ll)$ shape for nominal WW and mimicked WW processes	114

6.23	Distribution of $m_{T2}(ll)$ in events with nonprompt lepton candidates . . .	115
6.24	Distribution of $m_{T2}(ll)$ in events with same-charged lepton pairs	115
6.25	Distribution of $m_{T2}(ll)$ in events with positive and negative same-charged lepton pairs	116
6.26	Distribution of $m_{T2}(ll)$ for events in $t\bar{t}Z$ control region	117
6.27	Distribution of $m_{T2}(ll)$ for events in the WZ control region	118
6.28	Normalized shape of the $m_{T2}(ll)$ distribution obtained with and without different K factor sets	120
6.29	Observed and expected jet multiplicity distribution for ZZ process . . .	121
6.30	Distribution of $m_{T2}(ll)$ in events with two leptons satisfying $ m_{\ell\ell} - m_Z < 15$ GeV	123
6.31	Distribution of $m_{T2}(ll)$ in events with two leptons satisfying $ m_{\ell\ell} - m_Z < 15$ GeV after applying Drell-Yan correction	124
7.1	Distributions of $m_{T2}(ll)$ after the fit to data in the chargino signal regions for DF events	135
7.2	Distributions of $m_{T2}(ll)$ after the fit to data in the chargino signal regions for SF events	136
7.3	Distributions of $m_{T2}(ll)$ after the fit to data in the top squark signal regions for DF events	137
7.4	Distributions of $m_{T2}(ll)$ after the fit to data in the top squark signal regions for SF events	138
7.5	Upper limits at 95% CL on chargino pair production cross section . . .	139
7.6	Upper limits at 95% CL on top squark production cross section	140
9.1	Distribuciones de las variables discriminadoras	154
9.2	Límites superiores encontrados a 95% CL para la sección eficaz de producción de una pareja de charginos	160
9.3	Límites superiores encontrados a 95% CL para la sección eficaz de producción de una pareja de squarks top	161
B.1	Jet multiplicity distributions in events with $100 < p_T^{miss} < 140$ GeV and no b-tagged jets	167
B.2	Jet multiplicity distributions of events with three leptons, no b-tagged jets and $p_T^{miss} > 140$ GeV	168

List of Tables

2.1	Quiral supermultiplets in the Minimal Supersymmetric Standard Model	18
2.2	Gauge supermultiplets in the Minimal Supersymmetric Standard Model	19
2.3	Masses eigenstates in the Minimal Supersymmetric Standard Model	21
3.1	Maximum kinetic energy and speed of a proton in the different steps of the CERN accelerator complex. The rest mass of the proton is $0.938 \text{ GeV}/c^2$.	31
3.2	LHC parameters for 2012 and 2016 compared to the nominal values of design.	33
3.3	Main parameters of the CMS superconducting solenoid.	39
5.1	Observable ECAL and tracker variables used for electron identification	74
5.2	Tight and very loose electron identification requirements used in the analysis	77
5.3	Muon identification requirements for the Medium working point	78
5.4	HLT triggers paths used in this analysis	91
6.1	Baseline selection	96
6.2	Expected signal yields at different stages of the baseline event selection	96
6.3	Signal region definition for the chargino search	110
6.4	Signal region definition for the top squark search	110
6.5	Control region definition for the estimate of $t\bar{t}Z$ production	117
6.6	Control region definition for the estimate of WZ production	118
6.7	Simulated samples of the $ZZ \rightarrow 4\ell$ production events used in the analysis	119
6.8	Observed and expected events of ZZ process in the 4ℓ control region	120
6.9	Normalization scale factors for $t\bar{t}Z$, WZ , and ZZ backgrounds	122
6.10	Systematic uncertainties in the predicted yields for SM processes	128
6.11	Systematic uncertainties in the predicted yields for representative signal points	128
7.1	Observed and expected yields in the signal regions for the chargino search after the background-only fit	131
7.2	Observed and expected yields in the signal regions for the top squark search after the background-only fit	132
7.3	Expected yields for a reference signal point used in the chargino search	133
7.4	Expected yields for a reference signal point used in the top squark search	134
9.1	Número observado y esperado de eventos en las regiones de señal de búsqueda del chargino	158
9.2	Número observado y esperado de eventos en las regiones de señal de búsqueda del squark top	159

A.1	Data event samples	164
A.2	Simulated processes of the Standard Model used as backgrounds . . .	165
A.3	Simulated signal processes	166
C.1	Range of systematic uncertainties on the total background prediction in the $e\mu$ channel for the chargino search	172
C.2	Range of systematic uncertainties on the total background prediction in the $ee+\mu\mu$ channel for the chargino search	172
C.3	Range of systematic uncertainties on the predicted yield for a repre- sentative chargino mass point in the $e\mu$ channel	173
C.4	Range of systematic uncertainties on the predicted yield for a repre- sentative chargino mass point in the $ee+\mu\mu$ channel	173
C.5	Range of systematic uncertainties on the total background prediction in the $e\mu$ channel for the top squark search	174
C.6	Range of systematic uncertainties on the total background prediction in the $ee+\mu\mu$ channel for the top squark search	174
C.7	Range of systematic uncertainties on the predicted yield of a repre- sentative top squark mass point in the $e\mu$ channel	175
C.8	Range of systematic uncertainties on the predicted yield for a repre- sentative top squark mass point in the $ee+\mu\mu$ channel	175
D.1	Pull values of the nuisance parameters for the top squark search	179
D.2	Pull values of the nuisance parameters for the chargino search	180
D.3	Variation of the SM rate parameters for the chargino search	181
D.4	Variation of the SM rate parameters for the top squark search	181

...a mi familia

Chapter 1

Introduction

The origin of the Universe is a great mystery. The latest measurements date its beginning at 13.7 billions years ago. However, the most ancient sign of its creation directly observed is the Cosmic Microwave Background, which dates back thousand of million of years ago. Before this time, the Universe was so hot and dense that the light was not able to shine, preventing our telescopes from any observation beyond this point.

However, the conditions of heat and energy at the early Universe can be recreated right here, on Earth, with particle accelerators. Colliding high energy particles we can learn about the matter and the physics at the energy scale similar to that in the primitive Universe. The particle detectors can catch these collisions and provide experimental data for testing the fundamental mechanisms predicted by the theory.

The most performant theory to explain the physics in the first minutes after the Big Bang is the Standard Model of particle physics. It has so far been able to describe a wide variety of phenomena with outstanding precision at scales of energy of the order of TeV. However, this theory does not address some of the most important issues of Nature, such as the hierarchy problem between the Higgs boson mass and the Planck scale [1, 2], and does not contain a dark matter candidate to explain cosmological observations [3–5]. The supersymmetry [6–14] is an extension of the Standard Model that assigns a fermion (boson) superpartner to every boson (fermion) of the Standard Model. The introduction of this new symmetry can solve the hierarchy problem since the large quantum loop corrections to the Higgs boson mass, due mainly to the top quark, can be largely canceled by the analogous corrections from the top quark superpartner [15–17]. Moreover, if R -parity [18] is conserved, the lightest supersymmetric particle is stable and, if massive, provides a good candidate for dark matter.

This thesis presents a search for pair production of supersymmetric particles in events with two oppositely charged leptons (electrons or muons) and missing transverse momentum [19]. The data sample corresponds to an integrated luminosity of 35.9 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ collected with the CMS detector during the 2016 data taking period at the LHC. The results are interpreted in terms of several simplified supersymmetric models for the lightest chargino ($\tilde{\chi}_1^\pm$) and the lightest top squark (\tilde{t}_1) pair production assuming R -parity conservation, with the lightest neutralino as the lightest supersymmetric particle of the models. The chargino search is studied along the whole mass plane while the top squark search is focused on the compress spectrum where the top squark mass is close to the top quark mass in the context of natural supersymmetry models.

The document is organized as follows: Chapter 2 review the main aspects of the Standard Model and supersymmetry; Chapter 3 introduces the Large Hadron Collider and the CMS experiment; Chapter 4 and Chapter 5 describe the reconstruction of the physics objects and the details on the data and simulated event samples used in this search, respectively. Chapter 6 presents the strategy of the analysis, discusses the estimates of the Standard Model processes contributing to the selected events and details the sources of systematic uncertainty; Chapter 7 reports the results and their interpretation in terms of the considered simplified models. Finally, Chapter 8 summarizes the results and conclude this document. In addition, the Chapter 9 provides a short statement in Spanish of the main points covered in this thesis.

Chapter 2

The Standard Model and supersymmetry

Our understanding of Nature has been always motivated by the experience of what surrounds us. The physics sciences provide an effective mathematical framework and empirical procedures to explain very different phenomena in our world. In special, particle physics seeks the fundamental constituents of the Universe and their interactions. In the heart of our knowledge about these laws is the Standard Model of particle physics. This theory provides a unified description of the elemental structure of the matter and the fundamental forces that govern on it in terms of particles and the fundamental interactions between them. However, although experimentally successful, the Standard Model is not a complete theory. There are several open questions that lead scientists to wonder about a new physics description beyond the current formalism. Of all the different approaches that are being carried out, supersymmetry is one of the most promising theories. Proposing the expansion of the current Standard Model through the introduction of a new symmetry, the supersymmetric expansion can solve several of the most important shortcomings at the moment.

In this chapter the theoretical framework behind the work of this thesis is presented. Section 2.1 introduces the Standard Model elements as well as the main shortcomings of the model. In Section 2.2, the supersymmetric extension of the Standard Model and its consequences are explained. Finally, an overview of the underlying processes in hadron-hadron collisions at LHC is given in Section 2.3.

2.1 The Standard Model

The Standard Model (SM) is a renormalizable and relativistic Quantum Field Theory (QFT) with the local gauge symmetry $SU(3) \otimes SU(2) \otimes U(1)$ as internal symmetry [20–24]. The SM gives a quantitative description of three of the four fundamental interactions observed in Nature: electromagnetism, weak interaction and nuclear strong force; and the elementary particles that experience or propagate them.

Practical outlook

The elementary particles are characterized by the spin, charge, the mass and the quantum numbers determining its interactions. Figure 2.1 shows the complete scenario. Twelve fermions, particles of spin-1/2, four gauge bosons, particles of spin-1 which mediate the interactions, and one scalar boson, the Higgs boson, an special

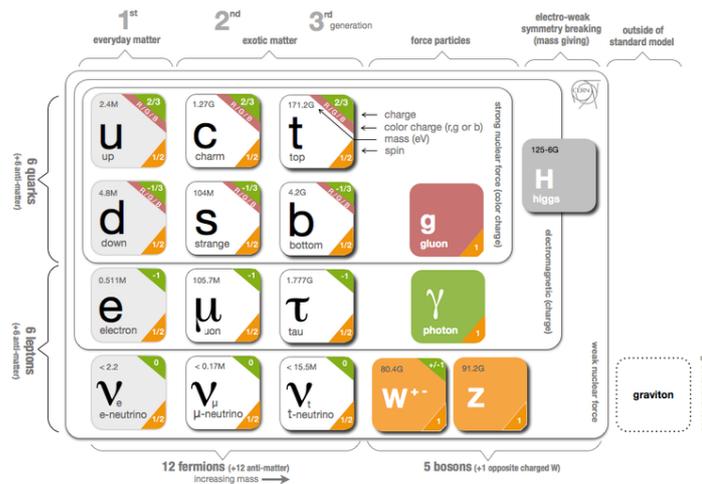


FIGURE 2.1: The Standard Model fundamental particles.

particle of spin-0. In addition a set of antiparticles defined by the same mass but opposite quantum numbers fills the schema with twelve antifermions.

The fermions are the elementary constituents of the ordinary matter. There are two types of fermions, leptons and quarks, which are in turn divided in six flavours. Both, quarks and leptons can be grouped as a function of their mass in three generations. All the stable matter observed in the Universe is made only by particles of the first generation. The particles of the second and third generations are understood as copies of the first generation particles, except for their masses. The origin of this mass hierarchy and generation structure is currently an unknown of particle physics.

All the fermions are susceptible to the weak interaction, but only those with non-zero electric charge are subject to the electromagnetic force. The strong interaction simply affects the quarks which, beside of electric charge, have color charge. Each interaction is described by a renormalizable QFT compatible with special relativity and a gauge symmetry. Essentially, each of these three interactions is described by one of the three gauge symmetries of the SM:

- Electromagnetic force ($U(1)$): mediated by the photon, is responsible for the electromagnetic behavior of Nature at low and high-energy levels, and is fully described by Quantum Electrodynamics (QED). Since the photon is massless, the range of this interaction is infinite.
- Weak interaction ($SU(2)$): mediated by the W^\pm and Z^0 bosons, it is responsible for the nuclear β -decays of certain radioactive isotopes and the nuclear fusion processes like the ones that fuel the Sun. Because of the large mass of the mediators the range of interaction is very small ($\sim 10^{-3}$ fm).
- Strong nuclear force ($SU(3)$): mediated by gluons at the nucleus length range (≈ 1 fm), it is responsible for keeping the atomic nuclei stable. It is the manifestation of a more fundamental theory of strong interaction, the Quantum Chromodynamics (QCD).

Because of the nature of QCD, quarks have been never observed as free particles but only confined in bound states, called hadrons. The hadronic states composed by a quark and an antiquark ($q\bar{q}$) are known as mesons. The proton and the neutron which are composed by three quarks (qqq) are called baryons and their antiparticles, consisting of three antiquarks ($\bar{q}\bar{q}\bar{q}$) are called antibaryons. The quark flavour is conserved in electromagnetic and strong interactions but not in the weak ones, as quark mass eigenstates do not correspond to the weak-interaction eigenstates. Their mixing is described by the Cabbibo-Kobayashi-Maskawa (CKM) matrix [25, 26].

In the 1960s, Glashow, Salam and Weinberg (GSW) developed a unified theory of the electromagnetism and the weak interactions, the GSW model [27–29]. One consequence of the model is the predictions of the Z^0 and W^\pm masses. They were awarded with the Nobel Prize in Physics in 1979 for their work.

Finally, the picture is completed by gravity, which although extremely weak, is always attractive and is therefore responsible for the large-scale structure in the Universe. Although gravity is one of the four fundamental forces, and essential part of Einstein's General Relativity theory (GR), it is not described in the SM. The purpose of establishing a unified theory implies the combination of the SM and GR resulting in a new field associated to a gravity mediator, the spin 2 particle called *graviton* for which there is not any experimental probe of existence up to date.

The general Lagrangian of the SM describes the coupling between the forces and the fermions through the gauge bosons. It can be written as the sum of three parts:

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EWK} + \mathcal{L}_H, \quad (2.1)$$

where \mathcal{L}_{QCD} is the quantum chromodynamics Lagrangian that describes the interactions of quarks and gluons, \mathcal{L}_{EWK} corresponds to the electroweak interactions between γ , W^\pm and Z^0 with fermions, and \mathcal{L}_H that is the Higgs part of the Lagrangian.

Quantum Chromodynamics

The Quantum Chromodynamics (QCD) is a non-abelian gauge quantum field theory that describes the strong interaction between quarks and gluons. The interaction is mediated by eight massless (spin-1) gluons that correspond to the eight generators of the $SU(3)$ local gauge symmetry group. In analogy with the electric charge, there are three conserved "colour" charges: red (r), blue (b) and green (g). The gluons are therefore a neutral octet of coloured states analogous to $q\bar{q}$ meson flavour states. The quarks, which carry electric and color charge, exist in three orthogonal colour states represented by the colour functions as:

$$r = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad b = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad g = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (2.2)$$

While the antiquarks carry the opposite colour charge to the quarks \bar{r} , \bar{b} , \bar{g} . Only particles with color charge can couple to gluons. Hence, the gluons can couple with themselves. The leptons are colourless and cannot feel the strong force.

The QCD Lagrangian is invariant under transformations of the local gauge symmetry group $SU(3)$ and it is given by:

$$\mathcal{L}_{QCD} = \sum_q [\bar{\psi}_{q,i} (i(\gamma^\mu D_\mu)_{ij} - m_q \delta_{ij}) \psi_{q,j} - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}], \quad (2.3)$$

where repeated indices are summed over. The γ^μ are the Dirac γ -matrices. The $\psi_{q,i(j)}$ is the quark field spinor of a flavour quark q with color index $i(j)$ ($i, j = 1, 2, 3$). D_μ is the gauge covariant derivative:

$$D_\mu = \partial_\mu - ig_c A_\mu^a T_a. \quad (2.4)$$

The factor T_a corresponds to the generators of the non-Abelian group $SU(3)$, A_μ^a are the gluon fields and g_c is the dimensionless coupling strength. Finally, $G_{\mu\nu}^a$ is the gauge invariant gluon field strength tensor, given by,

$$G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g_c f^{abc} A_\mu^b A_\nu^c \quad a = 1, \dots, 8. \quad (2.5)$$

This term includes f^{abc} , the structure constants of $SU(3)$. The $SU(3)$ color symmetry is an exact symmetry and QCD is invariant under unitary transformations in the color space. Consequently, the strength of the QCD interaction is independent of the colour charge of the quark. The parameters of QCD are the coupling strength g_c (or $\alpha_s = \frac{g_s^2}{4\pi}$) and the quark masses m_q . The QCD theory has two interesting properties:

- **Asymptotic freedom.** It has been observed that the strength of the QDC interaction decreases with the distance between the interacting coloured particles, that is, with the increases of the interaction energy. In an opposite way, the strength gets stronger at higher distances or smaller energies. When the coupling strength of QCD, α_s , is soft, the calculations can be performed using the perturbation theory. However, at low energies, when the interaction strength is large, $\alpha_s \sim \mathcal{O}(1)$, the evolution of the coupling constant cannot be analytically predicted and the other peculiar property starts to play a leading role, the color confinement. The asymptotic freedom was discovered in 1973 by David Gross and Frank Wilczek [30] and independently by David Politzer [31] who were awarded with the Nobel Prize of Physics in 2004.
- **Color confinement.** Only colorless hadrons (color-singlets) can be observed as free particles in Nature. The hypothesis of color confinement arises as a consequence of no experimental observation of free quarks or gluons and it is believed that the origin of the confinement resides in the asymptotic freedom. When two confined quarks start to separate from each other, the coupling strength of the QCD interaction increases, at certain point this separation makes the creation of a new quark-antiquark pair energetically possible. Eventually, the initial hadron is turned into two hadrons instead of two isolated color charges. This process, in the regime of non-perturbative QCD, is called hadronization, fragmentation, or string breaking, and it is the basis for the production of particle jets in high energy physics accelerators.

At the moment, there is none analytic formalism for describing the non-perturbative QCD behaviour. However many efforts are focused on the techniques of lattice QCD calculations, that tries to explain the asymptotic freedom and the color confinement properties from first principles.

The role played by QCD in high energy proton colliders is crucial. Sections 2.3 and 4.3.1 offer a brief overview of the underlying physics used to describe the high energy collisions at hadron colliders.

Electroweak interaction

The electroweak interaction is based on the gauge group $SU(2)_L \otimes U(1)_Y$. The $SU(2)_L$ refers to the local phase transformations of the weak interaction whose three generators are the 2×2 Pauli spin-matrices. The required gauge invariance is satisfied by introducing three spin-1 gauge fields, W_i ($i=1,2,3$). Related to group symmetry is the weak isospin charge (I_W), whose third component is conserved by the weak interactions. The weak charged-current interaction associated with the W^\pm bosons couples different fermions in weak isospin doublets of $I_W = 1/2$, differing in one unit of electric charge. These weak isospin doublets are constructed from weak-interaction eigenstates:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L.$$

Since only left-handed quiral (LH) fermions or the right-handed quiral (RH) antifermions have been observed coupling to the W^\pm , the unaffected chiral states of RH fermions or LH antifermions are placed in weak isospin singlets, $I_W = 0$, which are therefore unaltered by the $SU(2)_L$ local gauge transformation and hence they do not couple to the gauge bosons of the symmetry.

The $U(1)$ symmetry group represents the local gauge transformation in QED formulation. The requirement of the local $U(1)$ phase invariance is satisfied by introducing a spin-1 gauge field B , the photon (γ). This symmetry is related to the electromagnetic charge (Q). In the same way, the underlying $U(1)$ symmetry group of the electroweak unification is denoted by $U(1)_Y$ where Y refers to the weak hypercharge (Y_W). The weak hypercharge can be expressed as a linear combination of the electromagnetic charge Q and the third component of the weak isospin I_W^3 :

$$Y_W = 2(Q - I_W^3). \quad (2.6)$$

For satisfying the invariance under $U(1)_Y$ and $SU(2)_L$ local gauge transformations, the weak hypercharges in a weak isospin doublet must be the same. The unification of electromagnetic and weak interactions gives rise to a quantum field theory invariant under local $SU(2)_L \otimes U(1)_Y$ gauge symmetries, whose Lagrangian is expressed as:

$$\mathcal{L}_{EWK} = \mathcal{L}_g + \mathcal{L}_f. \quad (2.7)$$

The \mathcal{L}_g term describes the interaction between the three vector bosons W_i and the B boson,

$$\mathcal{L}_g = -\frac{1}{4}W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu}, \quad (2.8)$$

where $W^{i\mu\nu}$ ($i=1,2,3$) and $B^{\mu\nu}$ are the field strength tensors for the weak isospin and weak hypercharge gauge fields, respectively. And, the term \mathcal{L}_f is the kinetic term for the coupling of gauge bosons and fermions:

$$\mathcal{L}_f = \sum_f i\bar{\psi}_f \gamma^\mu \mathcal{D}_\mu \psi_f, \quad (2.9)$$

where the sum is extended over all fermions f and \mathcal{D}_μ is the gauge covariant derivative that guarantees the Lagrangian gauge invariance:

$$\mathcal{D}^\mu = \partial - i\frac{g}{2}\sigma_i W_\mu^i - i\frac{g'}{2}Y_W B_\mu. \quad (2.10)$$

The factors g and g' are respectively the $SU(2)_L$ and $U(1)_Y$ coupling constants. The Pauli matrix, σ_i , are the generators of the $SU(2)$ group and the weak hypercharge, Y_W , is the generator of the $U(1)$ group. The physical fields are obtained as linear combinations of these gauge fields:

$$\begin{aligned} A_\mu &= \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu, \\ Z_\mu &= \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu, \quad W_\mu^\pm = \frac{W_\mu^1 \mp iW_\mu^2}{\sqrt{2}}. \end{aligned} \quad (2.11)$$

These equations represent two neutral gauge bosons: the photon (described by the A_μ field) and the Z boson, in addition to two charged particles: W^+ and W^- bosons. The angle θ_W is the weak mixing angle or Weinberg angle.

At this point, all particles involved in the electroweak theory are massless. This fact is in apparent contradiction with the observations that give experimental evidences for massive fermions and W^\pm , Z bosons in Nature. However, the introduction of mass terms in the electroweak Lagrangian eliminates the gauge invariance and hence causes the breaking of gauge symmetries. Similarly, the fermion masses cannot be included without violating the $SU(3)$ gauge symmetry in the \mathcal{L}_{QCD} .

In 1964, Robert Brout and Francois Englert [32], Peter Higgs [33], and independently Gerald Guralnik, C. R. Hagen, and Tom Kibble [34], published almost simultaneously a theory to explain mass generation by breaking these symmetries while keeping the gauge invariant Lagrangian, one of its names is: the Brout-Englert-Higgs (BEH) mechanism. This solution proposes the Spontaneous Symmetry Breaking (SSB) of the SM $SU(2)_L \otimes U(1)_Y$ gauge group giving rise to the appearance of a physical scalar particle in the model, the so-called Higgs boson. Fermions and weak bosons would acquire their masses when coupling to this new particle. The Higgs boson is the visible manifestation of the Higgs field, and its mass is a free parameter of the model. In 2012, 48 years after this hypothesis was formulated, the Higgs boson has been observed at LHC by the ATLAS [35] and CMS [36] Collaborations.

2.1.1 The Brout-Englert-Higgs mechanism

The gauge symmetry guarantees a well defined and renormalizable Lagrangian and assures the renormalization of the associated Quantum Field Theory [37]. But, in order to generate massive weak bosons and fermions the gauge symmetry must be broken. A possible solution to this dilemma is based on the *Goldstone theorem* [38]

that allows getting non-symmetric results from a gauge invariant Lagrangian. The Goldstone theorem states that if a Lagrangian is invariant under a continuous symmetry group G , and the ground state of the related potential is only invariant under a subgroup $H \subset G$, there must exist as many massless spin-0 particles as broken G symmetry generators (generators of G that not belong to H). This degenerate set of states with minimal energy are known as Goldstone bosons and which transform, under the symmetry group G , as members of a multiplet. If one arbitrarily selects one of those states as the ground state of the system, then that symmetry G is spontaneously broken.

The idea of spontaneous symmetry breaking can be introduced through both a real or a complex scalar field. In the context of $SU(2)_L \otimes U(1)_Y$ local gauge symmetry of electroweak sector of SM, the minimal model to generate the SSB appears by defining two complex scalar fields placed in a $SU(2)_L$ weak isospin doublet as,

$$\phi(x) \equiv \begin{pmatrix} \phi^{(+)}(x) \\ \phi^{(0)}(x) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ \phi_3(x) + i\phi_4(x) \end{pmatrix}. \quad (2.12)$$

The fields ϕ_j ($j = 1,2,3,4$) are used to manifest the complexity of $\phi^{(+)}$ and $\phi^{(0)}$, and the superscripts $(+)$, (0) indicate the electric charge Q of the components. The upper and lower components of the doublet differ by one unit in charge. The gauge invariant Lagrangian under local transformations $SU(2)_L \otimes U(1)_Y$ for these fields is,

$$\mathcal{L} = (D_\mu \phi)^\dagger D^\mu \phi - V(\phi), \quad (2.13)$$

where D_μ is the gauge covariant derivative defined in Equation 2.10. Its action on the complex scalar field generates the masses of the gauge bosons. Since $\phi^{(0)}$ does not couple to the photon, the value of the scalar hypercharge Y_W (Equation 2.6), in D_μ , is fixed to $Y_W = 1/2$, giving also the right charge for $\phi^{(+)}$. The Higgs potential $V(\phi)$ is defined as,

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \quad (2.14)$$

with μ and λ the parameters of the model. In order to have non-vanishing vacuum expectation value, the potential should be bounded from below, i.e. $\lambda > 0$. For the quadratic piece there are two possibilities that will determine the shape of the potential:

- $\mu^2 > 0$: The minimum of the potential occurs when the field is $\phi(x) = 0$.
- $\mu^2 < 0$: The potential has a degenerate set of infinite states with minimum energy, satisfying

$$\phi^\dagger \phi = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}. \quad (2.15)$$

The Higgs scalar potential with $\mu^2 < 0$ is shown in Figure 2.2. The minimum is a continuous set of degenerate vacuum states in the complex plane of the Higgs field ($\phi(x)$). Once a certain state is chosen the symmetry gets spontaneously broken.

Since the interaction of the electroweak gauge field with the new complex scalar field conserve the electric charge quantity, the ground states must be invariant under gauge transformations of the QED symmetry group. Hence, only the neutral scalar

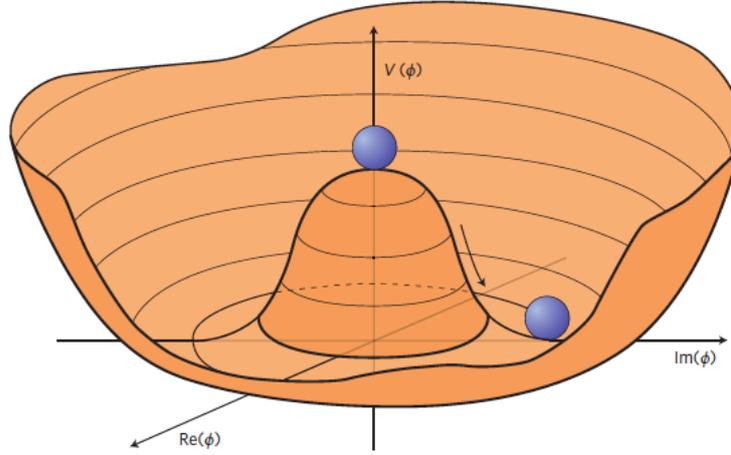


FIGURE 2.2: Shape of the Higgs scalar potential $V(\phi)$ for the case that $\mu^2 < 0$ ($\lambda > 0$), in which the minimum is a continuous set of degenerate vacuum states satisfying $|\phi|^2 = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}$ connected by the phase transformations of the rotational $U(1)$ symmetry. Choosing one of the massless states of the potential breaks spontaneously the rotational symmetry [39].

field can acquire non-zero vacuum expectation value,

$$\langle 0 | \phi^{(0)} | 0 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (2.16)$$

Once a particular ground state has been selected, the $SU(2)_L \otimes U(1)_Y$ symmetry gets spontaneously broken to the electromagnetic subgroup $U(1)_{QED}$, which by construction remains a true symmetry of the vacuum. According to the Goldstone theorem, there must exist as many massless states as generators of the broken symmetry in the $SU(2)_L \otimes U(1)_Y$ group, i.e. three massless Goldstone bosons should appear. It is possible to parametrize the excitations over the ground state in a general form as follows,

$$\phi(x) = \exp\left\{i\frac{\sigma_i}{2}\theta^i(x)\right\} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \quad (2.17)$$

The three real fields $\theta^i(x)$ ($i = 1, 2, 3$) are the Goldstone bosons associated to the SSB mechanism, and the fourth real field $H(x)$ is the Higgs field. Additional gauge symmetry constraints make the massless excitations unphysical and due to the underlying local symmetry, the dependence to $\theta^i(x)$ fields can be eliminated from the Lagrangian.

Taking the unitary (and physical) gauge value of $\theta^i(x) = 0$, the Higgs field becomes

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \quad (2.18)$$

and the corresponding part of the SM Lagrangian, Equation 2.13, can be expressed as

$$\mathcal{L}_H = (v + H)^2 \left\{ \frac{g_W^2}{4} W_\mu^\dagger W^\mu + \frac{g_W^2}{2 \cos^2 \theta_W} Z_\mu Z^\mu \right\} + \frac{1}{2} \partial_\mu H \partial^\mu H - \frac{1}{2} M_H^2 H^2 \left(1 + \frac{H}{2v} + \frac{H^2}{4v} \right).$$

The first term expresses the coupling of gauge bosons to the Higgs field. The mass of the W^\pm bosons is therefore determined by the coupling constant of the $SU(2)_L$ gauge interaction g_W and the vacuum expectation value of the neutral Higgs field v ,

$$m_W = \frac{1}{2} g_W v. \quad (2.19)$$

Here the coupling constant g has been substituted by g_W and g' has been absorbed by using the ratio of the $U(1)_Y$ and $SU(2)_L$ gauge symmetries couplings $\frac{g'}{g} = \tan \theta_W$. The acquired mass of the Z boson can be related to the W boson mass using the former relation, resulting:

$$m_Z = m_W \cos \theta_W. \quad (2.20)$$

While the photon remains massless because $U(1)_{QED}$ is an unbroken symmetry and it does not couple to the Higgs boson:

$$m_\gamma = 0. \quad (2.21)$$

The latest two terms are related with a new massive scalar particle H associated to the Higgs field, the Higgs boson. It couples with itself and mediates the Higgs interaction with the rest of the gauge fields. The last term corresponds to the potential $V(\phi)$ from which one obtains the relation between the vacuum expectation value v and the parameter μ^2 associated to the mass term of the Lagrangian:

$$v = \frac{\mu^2}{\lambda}. \quad (2.22)$$

Then, the Higgs boson mass is given by

$$m_H = v \sqrt{2\lambda}. \quad (2.23)$$

The vacuum expectation value of the Higgs field can be re-expressed as a function of the Fermi constant, thus determining the electroweak scale:

$$v = \frac{1}{\sqrt{2}G_F} = 246 \text{ GeV}. \quad (2.24)$$

The coupling λ is a free parameter of the theory, so the Higgs mass is not fixed but it has to be measured experimentally. Then, the GSW model is described by four parameters: the $SU(2)_L \otimes U(1)_Y$ gauge couplings g_W and g' , and the two free parameters of the Higgs potential, μ and λ .

The Higgs couplings are determined by m_H , m_W , m_Z and the vacuum expectation value v . These interactions are always proportional to the mass squared of the coupled bosons. Moreover, this introduction of a new additional scalar field doublet in the model can be also used to give mass to the fermions.

Fermion masses

The fermion mass terms $\mathcal{L}_m = -m\bar{\psi}\psi = -m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)$ cannot be included in the SM Lagrangian without violating the gauge symmetry $SU(2)_L \otimes U(1)_Y$. However, the BEH mechanism for the SSB generates also the mixing and masses of the fermions by their additional couplings with the scalar particle.

In the SM, the left-handed chiral fermions are placed in $SU(2)$ doublets, while the right-handed fermions are placed in $SU(2)$ singlets. Because of the two complex scalar fields of the Higgs field are placed in a $SU(2)$ doublet the combination $\bar{\psi}_L\phi$ will be invariant under $SU(2)_L$ while the combination $\bar{\psi}_L\phi\psi_R$ will be invariant under $SU(2)_L \otimes U(1)_Y$ symmetry. Hence, if one introduces in the Lagrangian a new term of the form $-g_f(\bar{\psi}_L\phi\psi_R + \bar{\psi}_R\phi^\dagger\psi_L)$ it will satisfy the electroweak gauge symmetry of the Standard Model and will obtain the fermion masses (putting aside the neutrino masses). In the case of the $SU(2)_L$ doublet containing the electron, the term of Yukawa interaction¹ is:

$$\mathcal{L}_e = -g_e \left[(\bar{\nu}_e \bar{e})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \bar{e}_R (\phi^{+*} \ \phi^{0*}) \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \right], \quad (2.25)$$

where g_e is the Yukawa coupling constant of the electron to the Higgs field. After the spontaneous symmetry breaking, in the unitary gauge (Equation 2.18) the \mathcal{L}_e becomes,

$$\mathcal{L}_e = \frac{1}{\sqrt{2}}(v + H)(g_e \bar{e}e). \quad (2.26)$$

The Yukawa coupling g_e is not predicted by the Higgs mechanism, but can be associated to the measured electron mass,

$$g_e = \sqrt{2}\frac{m_e}{v} \quad \text{and} \quad \mathcal{L}_e = -m_e \bar{e}e - \frac{m_e}{v}\bar{e}e h. \quad (2.27)$$

Therefore, the electron acquires its mass from the coupling of left-handed and right-handed massless chiral fermions through the interaction with the non-zero expectation value of the Higgs field.

In a similar way, the SSB mechanism generates the masses for the rest of fermions but the neutrinos. The mass term \mathcal{L}_m can be re-written as:

$$\mathcal{L}_m = - \left(1 + \frac{H}{v} \right) \{ m_d \bar{d}d + m_u \bar{u}u + m_e \bar{e}e \}, \quad (2.28)$$

where the Yukawa couplings for up-type quarks (u), down-type quarks (d) and leptons (e, μ, τ) are given by:

¹In particle physics, the interaction between a Dirac field ψ and a scalar field ϕ is called Yukawa interaction in honor to the Japanese physicist Hideki Yukawa. It takes the form $g_Y\bar{\psi}\phi\psi$.

$$g_f = \sqrt{2} \frac{m_f}{v}, \quad (2.29)$$

with the vacuum expectation value of the Higgs field $v = 246$ GeV. The mechanism for generating neutrino masses is more complex and can be found in the related literature of reference [40].

The Standard Model Lagrangian

The Standard Model Lagrangian can be written as the sum of the gauge field terms, dynamical terms and mass terms:

$$\mathcal{L}_{SM} = \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + i\bar{\psi} \not{D} \psi + h.c. + \psi_i y_{ij} \psi_j \phi + h.c. + |D_\mu \phi|^2 - V(\phi). \quad (2.30)$$

- The first term, $\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$, is the scalar product of $F_{\mu\nu}$, the field strength tensor. It encodes the $U(1)$, $SU(2)$, and $SU(3)$ gauge fields behaviour.
- The second term, $i\bar{\psi} \not{D} \psi$, describes the interaction between fermions and gauge bosons except for the Higgs boson. Its hermitian conjugate, denoted by $h.c.$, does not have a physical meaning.
- The third term, $\psi_i y_{ij} \psi_j \phi$, represents the coupling between fermions and the Higgs boson. Thereby, this is the fermionic mass term. Its hermitian conjugate $h.c.$, describes the same interaction, but with the antifermions.
- The fourth term, $|D_\mu \phi|^2$, is the coupling between weak gauge bosons and the Higgs field which thereby obtain their mass.
- The fifth term, $V(\phi)$, is the Higgs potential that includes the Higgs boson self-couplings.

The Standard Model has been able to describe particle physics phenomena with outstanding precision up to date. After its final formulation, in the mid-1970s, a succession of experimental results have positioned the Standard Model as one of the most successful theories in History. Some of these outstanding outcomes are the top quark (1995), the tau neutrino (2000) and the Higgs boson (2012) discoveries, as well as several accuracy predictions on the neutral weak currents properties and finally the W^\pm and Z^0 bosons (1983) discoveries. However, nowadays the SM is affected by several shortcomings such as the need of understanding the large discrepancy between the scales of the Higgs boson and Planck masses or the existence of a dark matter candidates conforming to the cosmological observations. The Section 2.2 presents an extension of the SM that deals with some of these problems by introducing a new set of particles, the supersymmetry.

2.1.2 Standard Model shortcomings

The recent precision tests and the discovery of the Higgs bosons have firmly established the validity of the SM as an effective theory at energies up to the electroweak scale (\sim TeV). But despite this success, the hints of physics beyond the SM gain importance every day and motivate a variety of experimental studies at the TeV energy domains with wide discovery potential. These discoveries could take the form

of evidences of supersymmetry, extradimensions or dark matter existence, among others [41]. A short list of the big open questions that are currently the subject of experimental and theoretical research is presented in this section.

- **Gravity.** The gravitation is an attractive force responsible for the formation of the macroscopic structures in the universe. Despite of this, at the scale of quantum mechanics, it is much less intense than the three other fundamental forces. So far, no consistent quantum description of the gravitational interaction has been achieved and the mathematical framework to describe gravity is not compatible with the current SM formalism. Even if its contribution at the electroweak scale is negligible, it is expected to play a much more relevant role at the Planck scale (2.4×10^8 GeV) where quantum gravitational effects become important, so that a complete theory able to coherently include the four fundamental interactions will be needed.
- **Matter-antimatter asymmetry.** According to the SM, antimatter is almost the mirror image of matter, and in principle there is not any reason for the dominance of one over the other. Despite of this, the Universe seem to be made almost entirely of matter. To explain this observation a new physics that provides baryon number violation, charge-parity (CP) violation, and out-of-equilibrium physics in the early Universe is required.
- **Grand Unification Theory (GUT).** In the mid 1970s Georgi and Glashow [42] proposed the unification of the three fundamental forces under a single unified theory, the GUT. They suggested that the three gauge symmetries could be accommodated in a larger $SU(5)$ symmetry group and treated as low energy representations of a single fundamental interaction. In this formulation, the three coupling constants of the SM are found to almost converge at high energy scale of about 10^{15} GeV. An interesting consequence of this convergence is found in the supersymmetric extension of the SM. The introduction of new particles at the 1 TeV mass scale makes the coupling constants converge to a single value of $\alpha_{GUT} \simeq 1/26$ at $|q| \sim 10^{16}$ GeV that reinforces the idea of a larger unified theory beyond the Standard Model. Despite further studies have demonstrated that $SU(5)$ is not the correct gauge group for a GUT and so far supersymmetry has not been discovered, the vision of the SM as low energy manifestation of something bigger and more general is getting importance in the scientific community.
- **Dark matter.** Since the mid 1930s, it has been known that a significant part of the mass in the Universe is not located in the stars as it was originally thought. The first evidences came from the velocity distribution of the stars orbiting the galactic centre. Recently, the precise measurements of the small fluctuations in the cosmic microwave background (CMB), which is related to the large scale structure in the Universe, constitute one of the basis for the existence of dark matter. The most performant cosmological standard model, or the Λ CDM Model, predicts that only 5% of the energy-matter density of the Universe corresponds to the ordinary known and SM barionic matter. The 23% is in form of cold dark matter but the vast majority of energy-matter density, of about 72%, is in form of dark energy. The experimental results and the sophisticated understanding of the cosmology challenge the Standard Model of particles physics to account for the constituents of dark matter, so that, an extension of the current theory is necessary.

- Hierarchy problem.** The quantum loop corrections from virtual effects of every particle or other phenomena that couple to the Higgs field contribute to increase the mass of the Higgs boson. If one considers that the SM is a valid theory at high energy scales such as the GUT ($\sim 10^{16}$ GeV) or the Planck mass ($\sim 10^{18}$ GeV) scales, these corrections become huge and make the Higgs boson mass incompatible with the electroweak scale ($\sim 10^2$ GeV). This is known as the hierarchy problem. The recent discovery of the Higgs boson with a mass ~ 125 GeV leaves the SM under this dilemma. To solve the situation, one possible solution is fine-tuning the new contributions to the Higgs mass such that they tend to cancel at high degree of precision. Other solutions happen to assume a new physics beyond the SM such as supersymmetry that avoids this procedure by a natural cancellation of the Higgs mass divergences.

The Standard Model Higgs boson H is a neutral scalar particle whose mass is given by $m_H = v\sqrt{2\lambda}$, and it is one of the free parameters of the model. All Higgs couplings are determined by m_H , m_W , m_Z , the vacuum expectation value v and the mass (squared) of the coupled particle. In Figure 2.3 the Feynman diagrams for three lowest-order decay modes are shown with their couplings strengths. The proportionality of the coupling constants to the mass of the coupled particle determines the dominant processes through which the Higgs boson will be produced and decay. The Higgs boson will preferentially couple to the most massive particles that are kinematically accessible.

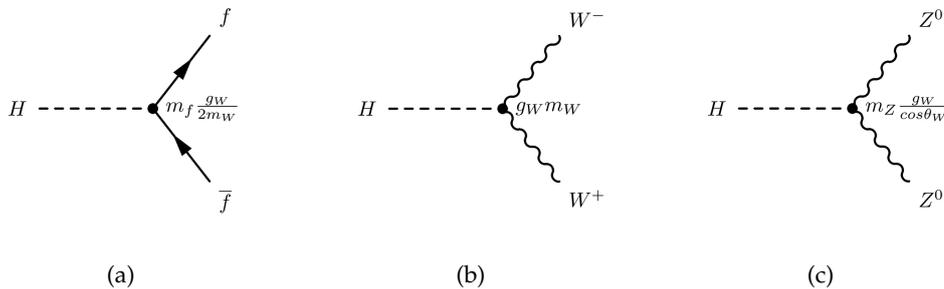


FIGURE 2.3: Lowest order Feynman diagrams for three Higgs boson decays: into fermions (a), into W^\mp boson (b), and into Z^0 bosons (c).

The contribution from virtual effects of every particle or other phenomena that couple directly or indirectly to the Higgs field generate large quantum corrections to the Higgs mass. In Figure 2.4 one-loop corrections to the Higgs squared mass parameter due to a fermion and a boson couplings are shown on the left and right, respectively.

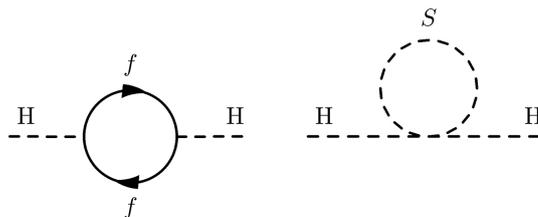


FIGURE 2.4: One-loop corrections to the Higgs boson, due to a Dirac fermion f (left) and a scalar S (right) couplings.

In the case of a Dirac fermion ψ with mass m_f coupling to the Higgs field H the interaction term $-\lambda_f \bar{\psi} H \psi$ yields a quantum correction:

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2} \Lambda_{UV}^2 + \dots, \quad (2.31)$$

where λ_f is the fermion coupling to the Higgs and Λ_{UV} is the ultraviolet momentum cutoff used to regulate the loop integral. This cutoff is interpreted as the energy scale up to which the Standard Model is valid. Just considering the value of the Planck mass as the physical ultraviolet cutoff, these corrections would be many orders of magnitude larger than the required corrections to give the recently discovered Higgs with 125 GeV of mass. Furthermore, due to the couplings of all SM particles with the Higgs boson, the entire mass spectrum of SM is directly or indirectly sensitive to the Λ_{UV} cutoff. Finally, on the other hand, if one considers smaller values of Λ_{UV} , these are not obvious in the current SM theory which will need a priori unnatural modifications. Therefore, the hierarchy problem is not only related to the quadratic divergences, but rather it is related to the quadratic sensitivity of the SM to Higgs mass scale.

The case of a heavy scalar S with mass m_S coupling to Higgs boson through a Lagrangian term $-\lambda_S |H|^2 |S|^2$ is shown in Figure 2.4 right. The Feynman diagram gives then a correction:

$$\Delta m_H^2 = +\frac{|\lambda_S|^2}{16\pi^2} \left[\Lambda_{UV}^2 + 2m_S^2 \ln\left(\frac{\Lambda_{UV}}{m_S}\right) + \dots \right]. \quad (2.32)$$

Comparing both Equations 2.31 and 2.32, one can observe that the minus sign of the dominant fermion loop contribution turns into plus for the massive boson loop contribution. Supposing that each of SM fermions coupling to Higgs boson is accompanied by an scalar partner with $\lambda_S = |\lambda_f|^2$, then the Λ_{UV}^2 contributions would be cancelled at all orders of theory of perturbation. This idea along with other considerations brought out the advantages of introducing in the SM a new symmetry relating fermions to bosons, the *supersymmetry*.

2.2 Supersymmetry

Supersymmetry is a principle that relates fermion fields to boson fields. The first ideas about this connection was proposed by Hironari Miyazawa in 1966 [43, 44]. Miyazawa worked in the field of the nuclear physics and he found a mathematical way to group baryons and mesons under the same algebraic representation. Although interesting, this supersymmetry did not involve the space-time and it could not be extended in a more general theory remaining forgotten for a long time. Some years after the first formulation of supersymmetry in the context of a quantum field theory, Julius Wess and Bruno Zumino developed a four-dimensional supersymmetric field theory [10, 11]. This will early allow the first applications in particle physics.

Finally in 1977, Pierre Fayet proposed the first supersymmetry model compatible with the SM [45, 46], the one known as Minimal Supersymmetric Standard Model (MSSM). Nowadays, the MSSM can solve some of the most important shortcomings

of the SM, such as, the hierarchy problem or the absence of a dark matter candidate. This section gives a brief introduction to supersymmetry from a phenomenological point of view regarding the experimental searches of new supersymmetric particles.

Practical outlook

A supersymmetric theory is built on a space-time symmetry that turns fermionic states into bosonic states, and vice versa. The general supersymmetric transformation is generated by the operator Q (and Q^\dagger , the hermitian conjugate of Q) that changes in $1/2$ the spin angular momentum of the SM particles:

$$Q|Boson\rangle = |Fermion\rangle, \quad Q|Fermion\rangle = |Boson\rangle. \quad (2.33)$$

The particles related under this operation are called *superpartners*. Both boson and fermion states, superpartner one of each other, are arranged into the irreducible representation of the supersymmetry algebra, the *supermultiplets*. Each supermultiplet contains the same number of boson and fermion degrees of freedom.

Because of the supersymmetry generators Q, Q^\dagger commute with the squared mass operator $(-P^2)$, the particles inhabiting the same irreducible supermultiplet will have the same eigenvalues of $-P^2$, namely the same mass value. Moreover, the operators Q, Q^\dagger also commute with the generators of the gauge transformations, therefore the particles in the same supermultiplet will have the same representation in the gauge group, that is, they will have same electric charge, weak isospin, and color degrees of freedom.

In general, the supersymmetric theories can include more than only one supersymmetry, that is, more than one distinct copy of Q, Q^\dagger generators. However, even if such models are useful tools to study the structure of the gauge theories, at this moment they do not provide any phenomenological prospect for experimental research. Under a practical view, the ordinary and phenomenologically viable type of supersymmetric model is the $\mathcal{N}=1$ model, where \mathcal{N} is the number of distinct copies of Q, Q^\dagger or supersymmetries in the theory.

The minimal supersymmetric extension of the SM is the mentioned model as Minimal Supersymmetric Standard Model, MSSM. The term "Minimal" refers to the minimum set of new stable particles and new interactions needed to complete a supersymmetric theory consistent with the known phenomenology. The next section enunciates the main features of the MSSM.

2.2.1 Minimal Supersymmetric Standard Model

The MSSM is the simplest supersymmetric extension of the Standard Model. The additional symmetry requirement associates each known fundamental particle to a superpartner with spin differing by $1/2$ unit. The number of SM fields doubles and are grouped in gauge supermultiplets.

The supersymmetric partners of the SM fermions are the spin-0 sfermions, namely *squarks* and *sleptons*, where "s" denotes scalar. They are designated using the same symbols as those used for the corresponding SM fermions but with a tilde " \sim ", making clear that they are superpartners. The associated symbol of "handedness", LH or RH will refer to that of their SM superpartners. The SM gauge interactions of each

squark and slepton field are the same as for the corresponding quarks and leptons, i.e. the bosons of the SM couple with the sfermions.

The fermionic superpartners of the SM gauge bosons are spin-1/2 particles named gauginos. In QCD, the color gauge interactions are mediated by the gluons, whose spin-1/2 color-octet supersymmetric partners are \tilde{g} , the gluinos. The electroweak gauge interaction is mediated by the spin-1/2 \tilde{W}^\pm , \tilde{W}^0 and \tilde{B}^0 gauginos, called respectively *winos* and *bino*. After the electroweak symmetry breaking, the corresponding gaugino mixtures of \tilde{W}^0 and \tilde{B}^0 generate the *photino* $\tilde{\alpha}$ and the *zino* \tilde{Z}^0 . If gravity is included, then the spin-2 *graviton* has a spin-3/2 superpartner, called the *gravitino*. Both sfermions and gauginos are members of different supermultiplets. There are two distinctions, the *chiral*, *matter* or *scalar* supermultiplets and the *gauge* or *vector* supermultiplets. Only chiral supermultiplets can contain fermions and their spin-0 boson superpartners, the sfermions (squarks and sleptons). The gauge supermultiplets, on their side, are a combination of spin-1 SM gauge bosons and their superpartners, the spin-1/2 gauginos.

The case of the scalar Higgs boson with spin 0 is special. Theoretical constraints require that the Higgs must reside in two quiral supermultiplets of weak hypercharge $Y_W = 1/2$. The SM Higgs field is then replaced by two $SU(2)_L$ -doublets of complex scalar fields with third component of the weak hypercharge $Y_W^3 = \pm 1/2$, denoted in this text by H_u and H_d respectively. The components of H_u have electric charges 1 and 0 (H_u^0 , H_u^+) with the third component of the weak isospin $I_W^3 = (1/2, -1/2)$ respectively. This doublet gives masses to the charged up-type quarks (up, charm, top). In the same way, the H_d $SU(2)_L$ -doublet components with $I_W^3 = (1/2, -1/2)$ have an electric charge of 0 and -1 respectively (H_d^0 , H_d^-), giving mass to the charged down-type quarks (down, strange, bottom) and charged leptons. The neutral scalar Higgs of the SM is a linear combination of H_d^0 and H_u^0 . Since the generic spin-1/2 superpartners are referred to by appending "-ino" to their SM partner name, the superpartner of the Higgs boson will be called *higgsino*. In the Table 2.1 the quiral supermultiplets of the MSSM are summarized as a function of their transformation properties under the $SU(3)_C \times SU(2)_L \times U(1)_Y$ Standard Model gauge symmetry group.

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

TABLE 2.1: Quiral supermultiplets in the Minimal Supersymmetric Standard Model. This table has been adapted from [14].

The higgsinos and electroweak gaugino are mixed with each other because of the effects of the electroweak symmetry breaking. The neutral gauginos (\tilde{W}^0 , \tilde{B}^0) and neutral higgsinos (\tilde{H}_u^0 and \tilde{H}_d^0) mix to form four mass eigenstates referred to as neutralinos $\tilde{\chi}_i^0$, with $i = 1, 2, 3, 4$. In a similar way, the charged gauginos (\tilde{W}^- and \tilde{W}^+) combine with the charged higgsinos (\tilde{H}_u^+ and \tilde{H}_d^-) to form four charged mass eigenstates with charge ± 1 known as charginos $\tilde{\chi}_i^\pm$ $i = 1, 2$. In both cases, these are labeled in ascending order of mass, so the index i runs from lightest to heaviest particle.

Table 2.2 summarizes the gauge multiplets of the MSSM showing their transformation properties under the $SU(3)_C \times SU(2)_L \times U(1)_Y$ Standard Model gauge symmetry group. The full picture of all the MSSM sparticles is filled with really new particles whose properties have to be confirmed experimentally.

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

TABLE 2.2: Gauge supermultiplets in the Minimal Supersymmetric Standard Model. This table has been adapted from [14].

R-Parity symmetry and dark matter candidates

In the MSSM, similarly to the SM, the conservation of either the baryon number (B_n) or total lepton number (L_n) is not postulated. But, unlike the SM, in the MSSM do exist terms in the potential that allow B_n - and L_n -violating processes. Just accepting the existence of these terms would contradict the experimental observations of B_n and L_n conservation, from which the strongest constraint is the non-observation of the proton decay that would violate both B_n and L_n in one unit.

Because of the numbers B_n and L_n are not taken as fundamental symmetries of Nature, thus, it is desirable avoiding the postulation of B_n and L_n conservation. A solution to this problem is to add a new symmetry which eliminates the terms of the renormalizable potential that possibly violate the B_n and L_n conservation. This new symmetry related to the matter parity is called *R-parity* (P_R), and it is defined for each particle of spin s as:

$$P_R = (-1)^{3(B_n - L_n) + 2s}. \quad (2.34)$$

For the SM particles and the Higgs boson, the R-parity is $P_R=+1$, while for the supersymmetric particles or sparticles (squarks, sleptons, gauginos and higgsinos) it takes the value of $P_R=-1$. Particles of the same supermultiplet do not have the same R-parity. If R-parity is an exactly conserved symmetry, there will not be mixing between sparticles ($P_R=-1$) and particles ($P_R=+1$). Furthermore, every interaction vertex in the theory must contain an even number of $P_R=-1$ sparticles. These statements have phenomenological consequences with interesting applications in the experimental searches for new physics:

- 1. The lightest sparticle with $P_R=-1$ is called the *lightest supersymmetric particle* or LSP. It is a stable particle, and if electrically neutral, it will interact only weakly with the ordinary matter becoming an interesting candidate for non-baryonic dark matter as is required by cosmology.
- 2. Each sparticle other than LSP must eventually decay into a final state that contains an odd number of LSPs.
- 3. In collider experiments, the sparticles can only be produced in even numbers.

These conclusions will be reflected in most searches for supersymmetric particles at LHC, in particular the work of this thesis.

Soft supersymmetric breaking in the MSSM

The most distasteful feature of supersymmetry theory is the fact that none of the predicted superpartners have been still observed. While the SM particles have been discovered and their properties measured, in the case of the supersymmetric particles no hints of their presence have been found at the time of this writing. If supersymmetry were an exact symmetry of Nature the superpartners of SM particles would have been detected long time ago since in the simplest models they have the same internal properties as electric and color charge, or mass. Therefore, if the supersymmetry exists, it cannot be an exact symmetry of Nature.

Because of the unbroken supersymmetry guarantees the cancellation of the quadratic divergences to the Higgs boson mass corrections, it is desirable for the broken supersymmetry that maintains the mass hierarchy between the electroweak scale and Planck mass scale (or any other large scale). Hence the breaking is considered to be "soft" in the sense that the effective Lagrangian of the MSSM can be written in two parts:

$$\mathcal{L} = \mathcal{L}_{SUSY} + \mathcal{L}_{soft}. \quad (2.35)$$

The \mathcal{L}_{SUSY} term contains all the gauge and Yukawa interactions preserving the supersymmetry invariance, and the \mathcal{L}_{soft} term that only includes terms for sparticles and not for their superpartners violates supersymmetry, but guarantees the unbroken supersymmetry properties. Therefore, the supersymmetric partners receives two contributions to their masses, from the electroweak symmetry breaking as well as from the soft supersymmetry breaking.

Mass eigenstates and naturalness

The mass eigenstates of squark and sleptons of the MSSM are predicted as combinations of right-left handed squarks and sleptons and sneutrinos with small mixing angles. They should be obtained by diagonalizing 6x6 squared-mass matrices for right-left handed squarks and sleptons, and 3x3 matrix for sneutrinos. The first-generation and second-generation of sfermions have soft Higgs couplings, and can be in a good approximation represented by the seven interaction eigenstates or gauge eigenstates. The third-generation of the sleptons and squarks (heavy flavour sfermions), on their side, can have larger couplings to the Higgs field and larger differences in mass leading to mass eigenstates different from gauge eigenstates. The gluino is a color octet fermion that cannot mix with any other particle of the MSSM. In that sense it is unique among all sparticles in MSSM and have the same mass and gauge eigenstate, as the gravitino.

In Table 2.3 it is summarized the MSSM sparticles specifying which are the gauge eigenstates and the mass eigenstates. The particular mass and mixing spectrum of the MSSM relies on the assumptions of the considered model. There are many scenarios but, in general terms, some features prevail as likely premises for many different type of models. Among them, it is the possibility for the neutralino to be the lightest supersymmetric particle, unless the gravitino is lighter or R -parity is not conserved, and that the lighter top squark \tilde{t}_1 and bottom squark \tilde{b}_1 are the lightest squarks.

The MSSM lightest mass eigenstate related to the Higgs boson is denoted as h^0 . This sparticle is likely candidate corresponding to the SM 125 GeV Higgs boson observed at LHC. However, while the soft supersymmetry breaking removes the quadratic divergences from the SM Higgs mass, it leaves corrections to the $m_{h^0}^2$ that could be

Names	Spin	P_R	Gauge Eigenstates	Mass Eigenstates
Higgs bosons	0	+1	$H_u^0 H_d^0 H_u^+ H_d^-$	$h^0 H^0 A^0 H^\pm$
squarks	0	-1	$\tilde{u}_L \tilde{u}_R \tilde{d}_L \tilde{d}_R$ $\tilde{s}_L \tilde{s}_R \tilde{c}_L \tilde{c}_R$ $\tilde{t}_L \tilde{t}_R \tilde{b}_L \tilde{b}_R$	(same) (same) $\tilde{t}_1 \tilde{t}_2 \tilde{b}_1 \tilde{b}_2$
sleptons	0	-1	$\tilde{e}_L \tilde{e}_R \tilde{\nu}_e$ $\tilde{\mu}_L \tilde{\mu}_R \tilde{\nu}_\mu$ $\tilde{\tau}_L \tilde{\tau}_R \tilde{\nu}_\tau$	(same) (same) $\tilde{\tau}_1 \tilde{\tau}_2 \tilde{\nu}_\tau$
neutralinos	1/2	-1	$\tilde{B}^0 \tilde{W}^0 \tilde{H}_u^0 \tilde{H}_d^0$	$\tilde{\chi}_1^0 \tilde{\chi}_2^0 \tilde{\chi}_3^0 \tilde{\chi}_4^0$
charginos	1/2	-1	$\tilde{W}^\pm \tilde{H}_u^\pm \tilde{H}_d^\pm$	$\tilde{\chi}_1^\pm \tilde{\chi}_2^\pm$
gluino	1/2	-1	\tilde{g}	(same)
goldstino (gravitino)	1/2 (3/2)	-1	\tilde{G}	(same)

TABLE 2.3: Masses eigenstates in the MSSM assuming that the mixing of the first- and second-family squarks and sleptons is negligible. The numerical index goes from the lightest to the heaviest particle. This table has been adapted from [14].

unnaturally large compared with the electroweak breaking scale. The largest contribution is proportional to the logarithmic of the top and squark top mass ratio $\ln(m_{\tilde{t}_1} m_{\tilde{t}_2} / m_t^2)$ leading to the problem called "the little hierarchy problem" of supersymmetry. In order to reduce this effect, some theoretical adjustments, under the so-called naturalness arguments, predict that at least the masses of the lightest few superpartners should not be much greater than the TeV scale. These would give top squark masses close to the top mass avoiding the fine-tuning or unnatural parameters in the theory. Despite the fact that these assumptions are a subject of theoretical discussion, they are a useful guide for the experimental searches.

The phenomenological MSSM

After introducing the soft supersymmetry breaking term, the MSSM Lagrangian incorporates more than 100 free parameters which include masses, mixing and phase angles. For direct LHC searches with limited range of energy and a finite precision, this set of parameters must be reduced to only those relevant for the experiments. This is known as the phenomenological adaptation of MSSM and it is denoted as pMSSM. The assumptions of the pMSSM, besides of avoiding a large number of parameters, provide a candidate for dark matter. The general characteristics are:

- no R-parity violation
- no new sources of CP violation
- mass-degenerated first and second generation of scalar
- no flavour-changing in neutral currents

In this context, the supersymmetric particles are produced in pairs. Theoretical predictions for the production cross section of various supersymmetric processes in pp collisions at $\sqrt{s} = 13$ TeV with an integrated luminosity of 36 fb^{-1} are shown in Figure 2.5. There are two weakly interacting massive particles (WIMP) in the pMSSM that are good candidates for dark matter: the lightest neutralino and the sneutrino. While the first one generally correspond to the LSP of the model, the last one escapes from the experiment leading the former as the only possible candidate.

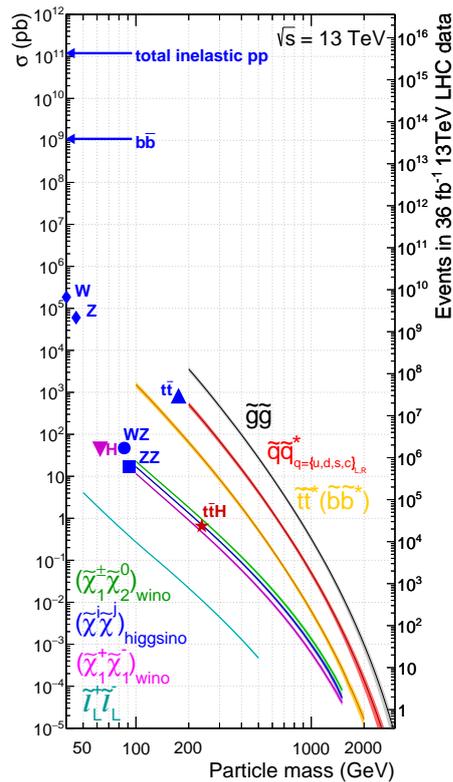


FIGURE 2.5: Theoretical production cross sections for selected supersymmetric processes as a function of the particle mass in pp collisions at $\sqrt{s} = 13$ TeV with an integrated luminosity of 36 fb^{-1} . On the right axis the expected number of events, on the left axis, the cross section value. Taken from the "LHC SUSY Cross Section Working Group" and [47].

In spite of the simplifications of the pMSSM, it is still a large number of additional particles, with different production and decay modes that leave a wide spectrum of new physics scenarios. In order to provide a more accessible description from the experimental point of view, all those properties and effects can be explicitly quantified in a given simplified model.

2.2.2 Simplified Supersymmetric Models

A simplified model [48–50], also known as simplified model spectra (SMS), is a simple model of new physics defined by an effective Lagrangian that contains few number of particles and their interactions. The number of free parameters in the theory is drastically reduced (absorbed) leaving only those directly related to collider physics observables: production cross sections, the branching fractions (\mathcal{B} 's) and the masses of the involved particles.

The simplified models give a general topology that covers a wide region of signal phase space and it is usual to interpret the result of one analysis in the context of several SMSs with similar final states. The following section presents the new physics search, objective of this thesis, highlighting the used SMS's for the results interpretation.

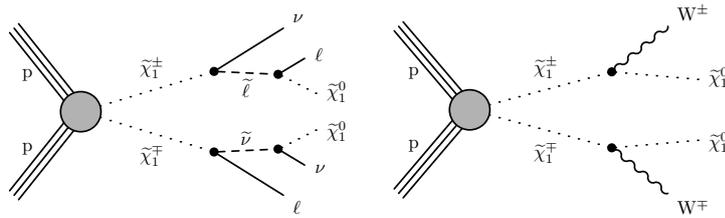


FIGURE 2.6: Feynman diagrams of the chargino pair production with two benchmark decay modes: the left plot shows decays through intermediate sleptons or sneutrinos, while the right one displays prompt decays into a W boson and the lightest neutralino.

2.2.3 The direct search for supersymmetric particles in this thesis

This thesis presents results on supersymmetric particles searches in final states with two oppositely charged (OC) leptons (ℓ) and missing transverse momentum stemming from the two LSPs. Only electrons (e) and muons (μ) are considered. The analysis targets two specific signal scenarios with the lightest chargino ($\tilde{\chi}_1^\pm$) and top squark (\tilde{t}_1) pair production using data from proton-proton (pp) collisions at a centre-of-mass energy (\sqrt{s}) equal to 13 TeV collected by the CMS experiment[51] at CERN LHC in 2016, and corresponding to an integrated luminosity of 35.9 fb^{-1} [19].

The results are interpreted in terms of the SMS scenarios shown in Figures 2.6 and 2.7 assuming R -parity conservation. The search for chargino pair production considers, as a reference, a model where the charginos decay into a lepton, a neutrino (ν), and the lightest neutralino ($\tilde{\chi}_1^0$) via an intermediate charged slepton ($\tilde{\chi}_1^\pm \rightarrow \tilde{\ell} \nu \rightarrow \ell \nu \tilde{\chi}_1^0$) or sneutrino ($\tilde{\chi}_1^\pm \rightarrow \tilde{\nu} \ell \rightarrow \nu \ell \tilde{\chi}_1^0$), Figure 2.6 (left). The three generations of sleptons are assumed to be degenerate, with a mass equal to the average of the chargino and neutralino masses. The B 's of the chargino decays into charged sleptons or sneutrinos are assumed to be equal.

Results are also interpreted in terms of a second model, Figure 2.6 (right), where each chargino decays into the lightest neutralino and a W boson. Previous searches for chargino pair production have been published by the CMS Collaboration using 8 TeV collision data in the context of the former scenario [52] and by the ATLAS Collaboration in the context of both scenarios using 8 TeV [53–55] and 13 TeV [56–58] collision data.

The search for top squark pair production focuses on an SMS in which the top squarks decays into a top quark and the lightest neutralino as shown in Figure 2.7 (left). The analysis targets the compressed spectrum scenario where the mass difference (Δm) between the top squark and the LSP lies between the top quark and W boson masses $m_W < \Delta m \lesssim m_t$. In this regime, the top quarks are produced off-shell, giving rise to final states with low-momentum bottom quarks (b) which often fail to be identified. Further interpretations of the results are given in terms of an additional model, where each of the pair-produced top squarks decays into a bottom quark and a chargino, which in turn decays into a W boson and the lightest neutralino, as shown in Figure 2.7 (right).

This work is complementary to another OC dilepton search published by the CMS Collaboration [59], which aims to test models where $\Delta m > m_t$ that results in signatures with on-shell top quarks and higher momentum particles. With respect to that

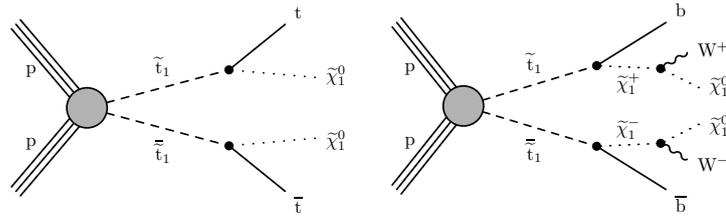


FIGURE 2.7: Feynman diagrams of the top squark pair production with two benchmark decay modes of the top squark: the left plot shows decays into a top quark and the lightest neutralino, while the right one displays prompt decays into a bottom quark and a chargino, further decaying into a neutralino and a W boson.

analysis, this search gains sensitivity in the compressed mass region by loosening the requirements on the jets from bottom quark hadronization and optimizing the signal event selection for the lower momentum carried by the neutralino LSPs.

At the time of the analysis here presented, other search has been carried out in the very compressed region $\Delta m \approx m_t$ being recently published by the CMS Collaboration [60]. This search targets a degenerate and nearly degenerated top squark region at low neutralino masses where the signal reproduces the top quark-antiquark pair production kinematics leading to a similar phase space. The strategy depends strongly on Δm where for the smallest $m_{\tilde{\chi}_1^0}$ the signal behaviour is very similar to the $t\bar{t}$ process and it is expected as an excess in the top background prediction.

The CMS Collaboration has also published other searches targeting the same signal models in the final states with exactly one lepton [61] and with no leptons [62], with the latter also covering the four-body-decay of the top squark in the region $\Delta m < 80$ GeV. The ATLAS Collaboration published several searches addressing these signal models using all three final states [63–65].

2.3 Physics at LHC

The underlying physics of a pp collision in the LHC is lead by the strong interaction between the quarks and gluons (partons) that compose each proton. Any hadron can be characterized as a bound state of a definite number of quarks called *valence* quarks, which confer the properties and quantum numbers of the whole system, and a non definite number of gluons. In the case of the proton, two up quarks and one down quark constitute the valence set. These quarks are continuously interacting by the exchange of gluons and those gluons can also self-interact producing more gluons or splitting into additional quark-antiquark pairs, the denominated *sea of quarks*. Therefore, the internal dynamics of the hadron will result in a momentum distribution among the valence quarks, the massless gluons and the sea quarks.

At the high energy of the LHC collisions, the interaction involves large momentum transfer ($Q^2 \gg 1$ GeV) causing a deep inelastic scattering between the two protons. In this regime of perturbative QCD, the hard scattering can be considered as elastic scatterings from the quasi-free (see *asymptotic freedom* in Section 2.1) point-like

constituents. However, neither the structure functions of the bounded system involved in the collision nor the cross section of the interaction can be calculated with perturbative QCD techniques.

In the regime of non-perturbative QCD, when the transferred momentum between partons is low ($Q^2 \lesssim 1 \text{ GeV}$), the vast majority of emissions that modify each parton's momentum in the proton are collinear to that parton. Technically, these emissions can be seen as modifying the proton's structure rather than being part of the parton interaction process itself. In this way, the proton internal structure can be determined in terms of the parton distribution functions (PDFs) that determine the probability density of finding a parton i carrying a momentum $p_i = x_i P$, being x_i the fraction of the total momentum P of the proton. These PDFs encapsulate the non-perturbative QCD behaviour of the partons.

The procedure known as *collinear factorization* is performed to systematically distinguish between the soft and hard aspects of hadron collider collisions. In the context of the $\overline{\text{MS}}$ factorization scheme² it is involved an arbitrary choice of the *factorization scale*, μ_F , understood as the limit above which the emissions with larger transverse momenta are treated in perturbative terms of proton structure function and below which the emissions are considered in the non-perturbative regime. The factorization theorem [66] establish that the cross section of a pp collision can be expressed as a function of the hard partonic cross section convoluted with the corresponding PDFs, in order to incorporate the probability of finding the necessary partons and its energy for the hard scattering. In Figure 2.8 there is an illustrative representation of the interaction process of the incoming particles.

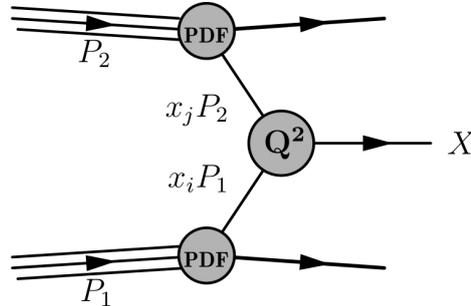


FIGURE 2.8: A representative schema of a hadron collision in the LHC. Two partons i, j , carrying a fraction of each hadron momenta $x_i P_1, x_j P_2$, interact exchanging the momentum Q^2 to produce a new particle X . The PDFs encapsulates the internal structure of each proton P_1, P_2 .

At present, the parton distributions cannot be calculated from first principles due to the large coupling constant of the partons. However, the QCD theory gives quantitative predictions about the rate of change of the PDFs with the Q^2 energy scale variation. These distributions are obtained by a fit to experimental data at one scale and then evolved to to different scales using the QCD evolution equations for parton densities (DGLAP) [67–70]. An example of PDFs at $Q = \sqrt{10} \text{ GeV}$ and $Q = 100 \text{ GeV}$ from the MSTW group can be found in Figure 2.9.

² $\overline{\text{MS}}$ squema stands for *minimal subtraction squema* defined in the context of dimensional regulation, where the scale parameter has no manifest physical interpretation.

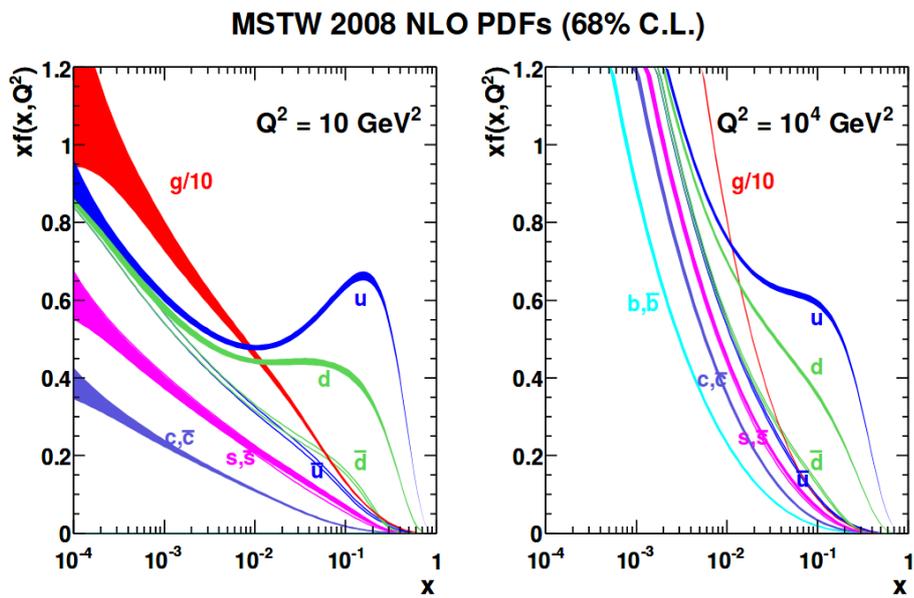


FIGURE 2.9: NLO PDFs at $Q^2 = 10 \text{ GeV}^2$ (right) and $Q^2 = 100^4 \text{ GeV}^2$ (left) calculated by the MSTW group [71].

Chapter 3

Experimental environment

The comprehension of the matter and the laws that govern the Universe have been intensively investigated by the mankind during all times. From the Greek conception of the atom to the discovery of radioactivity by Henry Becquerel (1896), the knowledge about the world that surrounds us has been evolving up to ours days.

The beginning of the twentieth century has represented a revolutionary epoch for the physics science. The birth of the Quantum Mechanics or General Relativity meant a change of paradigm that opened new fields in the theoretical and experimental exploration. But it would not be until the late forties, after the Second World War, that the experimental efforts combined with the new theoretical advances would allow the technological developments that we know today. In particular, the creation of the great experiments capable of exploring the smaller and largest scales expanding the boundaries of the known Universe began.

3.1 The European Organization for Nuclear Research

The European Organization for Nuclear Research usually known by the acronym of its French name, Conseil Européen pour la Recherche Nucléaire, CERN, was conceived as a laboratory to unite European scientists in the aim of performing fundamental physics research in a world-class nuclear physics facility. Located at Geneva canton, Switzerland, it started its mission in 1954 with the purpose of exploring what the Universe is made of and how.

Impulsed by the ideas of visionary scientists such as Pierre Auger and Louis de Broglie in France, and Niels Bohr in Denmark, it has been a reference centre for uniting people from all over the world to push the frontiers of science and technology development. The CERN 's convention states (6 July 1953):

"The Organization shall have no concern with work for military requirements and the results of its experimental and theoretical work shall be published or otherwise made generally available."

The CERN's facilities harbour a series of accelerator machines, particle detectors and computing infrastructure pioneers in history. From the first accelerator, the 600 MeV Synchrocyclotron (SC), built in 1957, passing through the Super Proton Synchrotron (SPS) working since 1971, to the Large Hadron Collider (LHC), which starts up its work in 2008, CERN has evolved, adapting its technology to the times.

Some of the most interesting results from the physics point of view are related with the artificial production of antimatter (1965), the quark-gluon plasma (2000) existing at about $10\mu\text{s}$ after the Big Bang, the W and Z bosons existence (1983), or the long-sought Higgs boson (2012) observation, as well as the discovery of the pentaquark hadrons (2015).

In general, its mission can be summarized in four words: "Research. Innovation. Collaboration. Inspiration."

This chapter introduces the experimental apparatus used to produce and collect the data of this search. It is organized in two main sections: the acceleration machine, Section 3.2, and the detection instrument, Section 3.3.

3.2 The Large Hadron Collider

The LHC [72] is the most recent acquisition in the CERN accelerator complex. The CERN accelerator complex is a succession of machines with increasingly higher energies. Each machine injects the beam into the next one, which takes over to bring the beam to an even higher energy. The LHC is the latest part in the acceleration chain and it represents the largest and most powerful particle accelerator in the world up to today.

The main physics purpose of LHC is to disentangle the current open questions related to the description of Nature provided by the SM, such as, the asymmetry between matter and antimatter, what the gravity is, the existence of dark matter and dark energy as well as provide a window onto the state of matter that existed in the early Universe, called *quark-gluon plasma*. Besides of the study of the properties and decay modes of the recently discovered Higgs boson (2012).

The LHC is a circular proton-proton (pp) and lead-lead ion collider. It has been constructed crossing the Franco-Swiss border in a tunnel of 26.7 km long and 3.0 m of diameter, which was the former location for the previous big accelerator of CERN, the Large Electron-Positron collider (LEP) [73] dismantled in 2000. The tunnel was built at a mean depth of 100 m and at a slight gradient of 1.4%, its depth varies between 175 m and 50 m.

In Figure 3.1 we can see an schema of the LHC. It is divided in eight arcs of circumference of 2.45 km length and eight insertions. The arcs contain the dipole magnets (154 each) that bend the charged particles trajectories. The insertions consist of a 545 m long straight section plus two (one at each end) transition regions (the 'dispersion suppressors'). Each insertion is designed for a specific use: physics (beam collisions within an experiment), injection, beam dumping, and beam cleaning. The part of the machine between two insertions is a sector and it is an autonomous working unit of the LHC. Each of the eight octants starts and ends in the middle of an insertion in such a way that each octant contains a complete insertion defining a different area depending on whether the magnets guide the beams into collision or through the injection, dumping, or cleaning sections.

The four main experiments around the LHC are ALICE [74], ATLAS [75], LHCb [76], and CMS [51, 77]. Along the way, three smaller experiments, TOTEM [78], LHCf [79] and MoEDAL [80], have joined them: TOTEM, installed next to CMS, LHCf next to

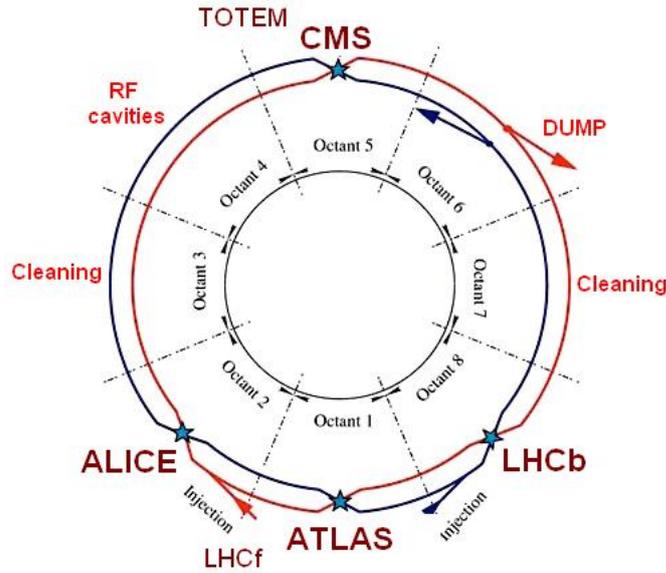


FIGURE 3.1: LHC schematic layout.

ATLAS and MoEDAL, deployed around the same intersection region as the LHCb detector.

In the LHC, two beams of particles, protons or lead ions (hadrons), are accelerated in opposite directions, along a circumference of 27 km length. Under the nominal conditions, each proton beam reaches the energy of $E = 7$ TeV at the moment of collision, giving, in the centre of mass frame, a total energy of 14 TeV ($\sqrt{s} = 2E$). The lead-ion beams have a maximum collision energy of 1150 TeV. These energies have not been reached by any other experiment of particle physics.

The data of this analysis has been produced in collisions of protons accelerated up to 6.5 TeV giving an energy of collision of 13 TeV. This data has been collected by the CMS experiment during the LHC operating time in 2016.

3.2.1 Proton acceleration process

Inside the LHC, the protons travel in opposite directions through separate beam pipes, two tubes kept at ultrahigh vacuum (pressure is 10^{-13} atm). They are guided along the ring by a strong magnetic field produced by a wide variety of superconducting electromagnets kept at a temperature of 1.9 K (-271.3°C) thanks to a distribution system of liquid helium that cools the magnets and other supply services.

The electronic devices incorporate thousand of magnets to control the beams, including dipoles, for bending the particle trajectory keeping it in a circle, combination of different multipoles to focus and squeeze the beam increasing the chance of collision, and accelerating cavities which are electromagnetic resonators whose main role is accelerate and keep the energy constant by compensating the energy losses through radiation. In a proton accelerator like the LHC, the maximum energy that can be achieved is directly proportional to the strength of the dipole field, given a

specific acceleration circumference. For this reason the design of the 1232 superconducting dipoles that constantly operate generating a constant magnetic field of 8.33 T at 14 TeV is the most technological challenge for LHC design.

The protons are extracted from a source of hydrogen gas. The hydrogen atoms are injected into a metal cylinder surrounded by an electromagnetic field that separates the atoms in their constituents, protons and electrons. These protons are introduced in the linear accelerator LINAC 2, which accelerates them up to 50 MeV after sending them to the Proton Synchrotron Booster (PSB) where they are accelerated to reach 1.4 GeV. The beam is then fed to the Proton Synchrotron (PS) where it is accelerated to 25 GeV, to be sent to the Super Proton Synchrotron (SPS), a ring of 7 km where the protons circulates increasing their energy up to 450 GeV. At this moment the protons are finally transferred to the two LHC beam pipes. The beams are accelerated to 6.5 TeV during around 20 minutes, both in a clockwise and an anticlockwise direction. The beams can be circulating for hours under normal operating conditions. The schema of the accelerator complex is shown in Figure 3.2. In Table 3.1 are shown the different energies reached by the protons in the acceleration process.

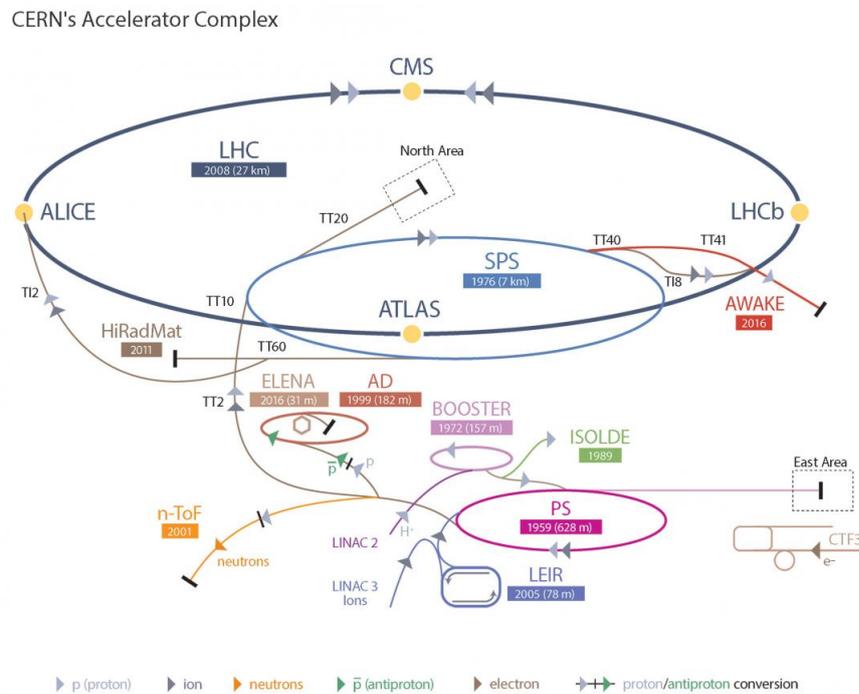


FIGURE 3.2: CERN accelerator complex.

The protons are grouped in well defined bunches. Each bunch has not a constant size, it can be squeezed or expanded. When they are far from the collision point, the bunches measure around 7 cm long and 1 mm wide, while when approaching the interaction point they are compressed to a section of about $16 \mu\text{m}^2$ which reduces the area of interaction increasing the chance of proton-proton interacting.

The beams are made of these bunches of protons. At the design values, the bunch spacing in a beam is 25 ns or 7.5 m (considering the protons travel at 99.9% the speed of light). So, with a circumference of 26659 m (27 km) we should have $26659/7.5 \sim 3550$ bunches, but, in order to get a correct sequence of bunches and be able to

Max. Kinetic energy	speed (%c)	Accelerator
50 MeV	31.4	Linac 2
1.4 GeV	91.6	PS Booster
25 GeV	99.93	PS
450 GeV	99.9998	SPS
7 TeV	99.9999991	LHC

TABLE 3.1: Maximum kinetic energy and speed of a proton in the different steps of the CERN accelerator complex. The rest mass of the proton is $0.938 \text{ GeV}/c^2$.

insert new ones when the non-useful ones are taken out, it is necessary to leave some free space, giving that the effective nominal (design) number of bunches per beam is 2880. Each bunch has around $N_p = 1.15 \cdot 10^{11}$ protons. Travelling close to the speed light the bunches cross at some point of space 11245 times per second, this is a revolution frequency of 11.25 kHz. Because the two beams are in the same conditions and the crossing point in the ring is just one at time, the average crossing rate can be calculated as $f = 11245 \times 2880 = 31.6 \cdot 10^6$ crosses/s (31.6 MHz).

The important parameters for an accelerator are the beam energy and the number of interesting collisions, specifically the number of collision per second per cm^2 that can be produced. The last is called luminosity \mathcal{L} and it determines the event rate of a collider; the bigger \mathcal{L} the bigger number of collisions and more likely is to find new and rare processes. The LHC is designed to get an instantaneous luminosity $\mathcal{L} = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$.

The total cross section of a proton-proton interaction represents the upper bound on cross section of any other process produced at the LHC. It is the sum of the inelastic collision cross section and the elastic collisions cross sections ($\sigma_{total} = \sigma_{el} + \sigma_{inel}$). At the design parameters is $\sigma_{total} = 100 \text{ mb}$, where the barn, $1\text{b} = 10^{-28} \text{ m}^2$, is the unit of area generally used in high energy particle physics.

Only the 60% of collisions give inelastic scattering which produce new particles with high enough angles with respect to the beam axis to be measured. The average number total number of event from inelastic collisions $N = \sigma_{inel} \cdot \mathcal{L} = 600 \cdot 10^6$ collisions/s. Considering number of crosses per second f , the average number of effective collisions per cross is $f \cdot N \sim 20$ (inelastic pp interactions / bunch cross). This number is called pileup.

The amount of data collected by experiments like CMS or ATLAS for each event or collision or interaction is around 1Mb which means 1PB for second. Collectively, the LHC experiments produce about 15 petabytes of raw data each year.

In addition to the proton-proton collisions, the LHC design program includes lead-lead ion collisions. The beams of fully stripped ($^{280}\text{Pb}^{82+}$) ions can reach an energy of 2.56 TeV/u (energy per nucleon) with a nominal magnetic field of 8.33 T in the dipoles, yielding a total energy of collision of 1.15 PeV and a nominal luminosity of $\mathcal{L} = 10^{27} \text{ cm}^{-2} \text{ s}^{-1}$. Since 2012 it has been included also lead-proton collisions. The LHC's lead-ions beams are accelerated in steps passing through Low Energy Ion Ring (LEIR), PS and SPS to finally be sent to the LHC where they are accelerated to the highest energy.

One of the main objectives for the lead-ion running is to recreate the condition at early stages of the Universe, few millionths of a second, shortly after the Big Bang, where the quarks and gluons were weakly interacting in the quark-gluon plasma. The study of the evolution of such structures of matter (hot nuclear matter) into the matter that makes up the Universe today is one of the roles of the LHC's experiments.

3.2.2 LHC performance

Although the SM has so far been tested with outstanding precision, it is considered an effective theory up to the \approx TeV scale. The wide discovery potential of the studies in the experimental frontier of the TeV energy domains and the engineering challenge motivated the choice of the 14 TeV as nominal energy produced in a pp collision in LHC.

However, before reaching the nominal values of centre of mass energy, the LHC has been increasing the beam energy gradually in order to optimize the delivery of particle collisions for physics research. It was first operated with a beam for short periods in 2008 and 2009. Then, in 2010, the first physics results could be produced under the collision energy of 7 TeV, rising to 8 TeV in 2012. This period is known as Run1. After that, a technical stop of 2 years, the first long shutdown (LS1), was used to consolidate the magnet interconnections and prepare all the systems to go towards the design beam energy.

The Run2 started in 2015, and corresponds to the data taking of 2015, 2016, 2017 and 2018. The beam energy for this period is 6.5 TeV giving an energy of centre of mass of 13 TeV with a bunch spacing of 25 ns, instead of the 50 ns of the 2011-2012 runs. These were preferences of the experiments in order to decrease the too high number of inelastic collisions per crossing (pileup) that would result. While 2015 was considered a re-commissioning and learning year, dedicated to prepare the machine for the full luminosity production, the 2016 was the first year, 2017 and 2018 have been years with outstanding beam operation (no data production).

Luminosity and LHC parameters

The event rate dN/dt of a physical process with a cross section σ_p is proportional to the collider instantaneous luminosity \mathcal{L} ,

$$\frac{dN}{dt} = \mathcal{L}\sigma_p. \quad (3.1)$$

This can be expressed in terms of machine and beam parameters [81] as follows:

$$\mathcal{L} = \frac{kN^2f}{4\pi\sigma_x^*\sigma_y^*}F = \frac{kN_p^2f\gamma}{4\pi\beta^*\epsilon_n}F, \quad (3.2)$$

where k is the maximum number of colliding bunches, N_p is the number of protons in each bunch, f is the LHC revolution frequency. The product $\sigma_x^*\sigma_y^*$ is the transverse size of the bunch at interaction point (IP) that can be expressed in terms of γ , the Lorentz factor, the normalized emittance ϵ_n (independent of energy) and the betatron function at the interaction point β^* , both quantities reflect geometrical characteristics of the beam and strongly depend on the LHC operation. Finally, the

LHC Parameter	2012	2016	Design
Energy per beam [TeV]	4	6.5	7
Bunch spacing [ns]	50	25	25
Bunch population N_p [10^{11} ppb]	1.65	1.1	1.15
Maximum number of bunches per ring k	1374	2220	2808
Crossing angle [μ rad]	290	370 / 280	285
β^* CMS/ATLAS [cm]	60	40	55
Normalized emittance ϵ_n [μ m]	2.5	2.2	3.75
Peak luminosity \mathcal{L} [10^{34} $\text{cm}^{-2}\text{s}^{-1}$]	0.75	1.4	1
Peak average event pileup μ	35	50	20
Peak stored energy [MJ]	145	270	360

TABLE 3.2: LHC parameters for 2012 and 2016 compared to the nominal values of design.

reduction factor F (≤ 1 ; $F = 0.65$ in 2016) that takes account of geometric luminosity reductions due to the presence of crossing angles between the beams at the IP.

A summary the main LHC parameter is found in Table 3.2 where the different values for 2012, 2016 and the design values can be compared. The luminosity performance is presented in Figure 3.3. During Run1 the maximum instantaneous luminosity was $7.6 \cdot 10^{33} \text{cm}^{-2}\text{s}^{-1}$, while in Run 2 the value exceeded the nominal value ($1 \cdot 10^{34} \text{cm}^{-2}\text{s}^{-1}$) with the luminosity peak in $1.4 \cdot 10^{34} \text{cm}^{-2}\text{s}^{-1}$.

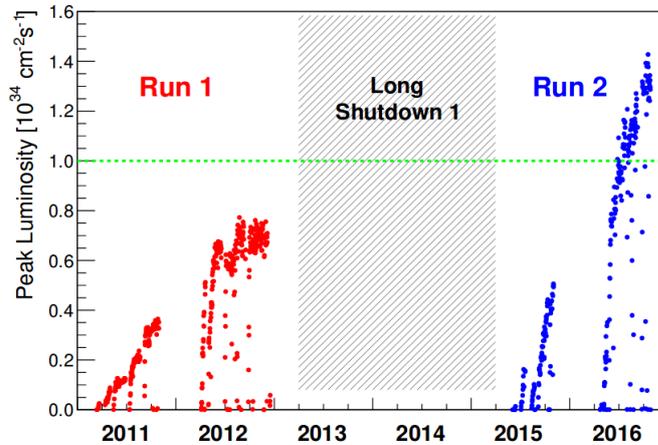


FIGURE 3.3: Evolution of peak luminosity between 2011 and 2016. The green line correspond to the design value $10^{34} \text{cm}^{-2}\text{s}^{-1}$ [82].

The integrated luminosity, denoted by L , is the total luminosity achieved by the collider during a certain amount of time and it is a measurement of a collider efficiency.

$$L = \int \mathcal{L} dt. \quad (3.3)$$

By means of the Equation 3.1 it gives a straight relationship between the total number of observed events of a particular type and the total cross section of the process.

L is measured with the inverse units of the cross section (SI of units). For proton-proton collisions in LHC typically ranges between several 1/picobarns (pb^{-1}) and 1/femtobarns (fb^{-1}).

The integrated luminosity of LHC with proton-proton collisions compared with the previous years and the expected projection is shown in Figure 3.5. The integrated luminosity registered by ATLAS and CMS experiments during 2016 reached the 40 fb^{-1} exceeding the planned 25 fb^{-1} . This is mainly thanks to the great availability of the operational time in LHC for the collision mode, more time than the previous years.

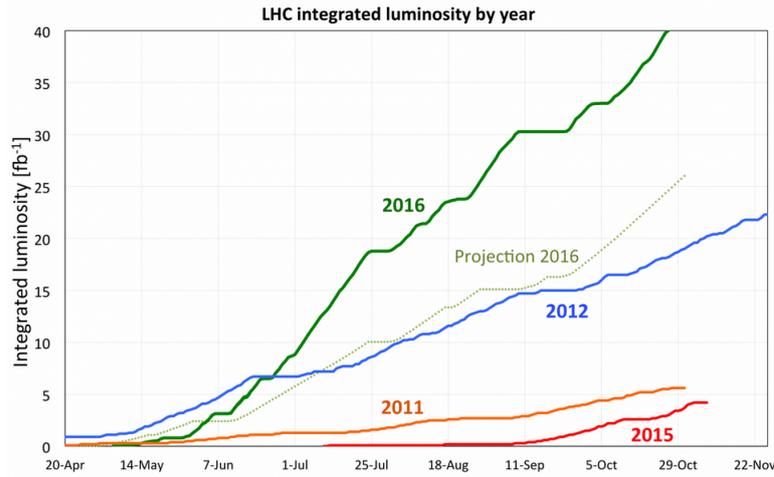


FIGURE 3.4: Evolution of integrated luminosity between 2011 and 2016. The dashed line correspond to the projection for 2016 [82].

Figure 3.5 shows the total integrated luminosity recorded by CMS at $\sqrt{s} = 13 \text{ TeV}$ in 2016 during the Run2.

3.3 The Compact Muon Solenoid

The data produced in the LHC is collected by the different experiments around the ring and is studied by their respective collaborations. In particular, this analysis studies the data produced by proton-proton collisions at 13 TeV collected in 2016 by the Compact Muon Solenoid (CMS) experiment [51, 84, 85].

The CMS detector is a general purpose experiment for which the main goals are to explore the physics at the TeV scale and study the mechanism for the electroweak symmetry breaking. The instrument has been prepared to search for new particles such as the Higgs boson, the supersymmetric partners of the SM particles or the Z' massive vector boson and the B_s mesons from heavy ions collisions, besides of the study of the SM processes. These processes include QCD, B-physics, detailed studies of the top quark properties or electroweak physics topics such as the W^\pm and Z^0 boson properties among others.

The detector receives its name from its overall characteristics. The total weight of 12500 tons divided up in a cylindrical structure of 21.6 m long with a diameter of

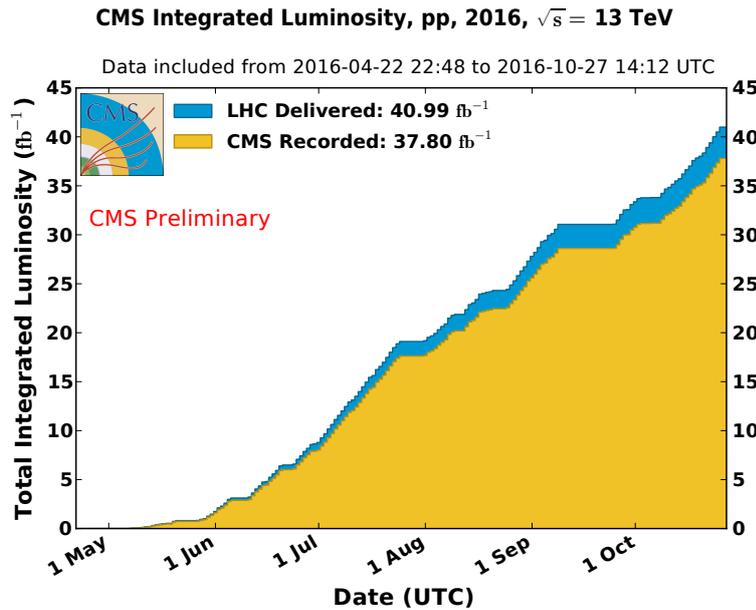


FIGURE 3.5: Integrated luminosity recorded by CMS versus the LHC delivered per day during the year 2016. This measurement uses the best available offline measurement and calibrations [83].

14.6 m high gives a *compact* structure. The design configuration was chosen to measure precisely the momentum charged particles. In particular the momentum of the centrally produced muons is measured very precisely (more than any other experiment) using different components of the so called *muon* system. The right performance of the muon system detection requires a high and constant magnetic field which is produced by a large superconducting *solenoid*. The choice of this kind of magnet is consistent with the cylindrical design of the apparatus.

The CMS detector follows the common shape of a high-luminosity experiment for general purpose in high energy physics (HEP). It is an hermetic cylindrical architecture made up by several concentric layers or subdetectors (grouped together like in a cylindrical onion), each of them with a particular role in the detection. A magnet is used to separate the different particles according to their charge and to measure their momentum. Two regions can be distinguished: the central region or barrel, and the forward region or the endcaps. Each of the two beams approach the detector by one of the endcaps and travel along the cylinder's axis to the interaction point, IP, that is in the centre of the detector. Figure 3.6 shows a sectional view of the CMS detector with the different detection systems.

These subdetectors can be grouped in two categories:

- The tracking devices: used to identify the tracks that electrically charged particles leave through the trails of ionized matter. In CMS there are 2 types: the vertex detector located close to the interaction point in the tracker and the muon chambers, located at the outer layers of the detector assembly (muons can travel meters through a dense material with small energy loss).
- The calorimeter devices: made of very dense material able to stop energetic particles. They are used to quantify the energy of particles by measuring the amount of energy released on the material. In CMS there are also 2 types:

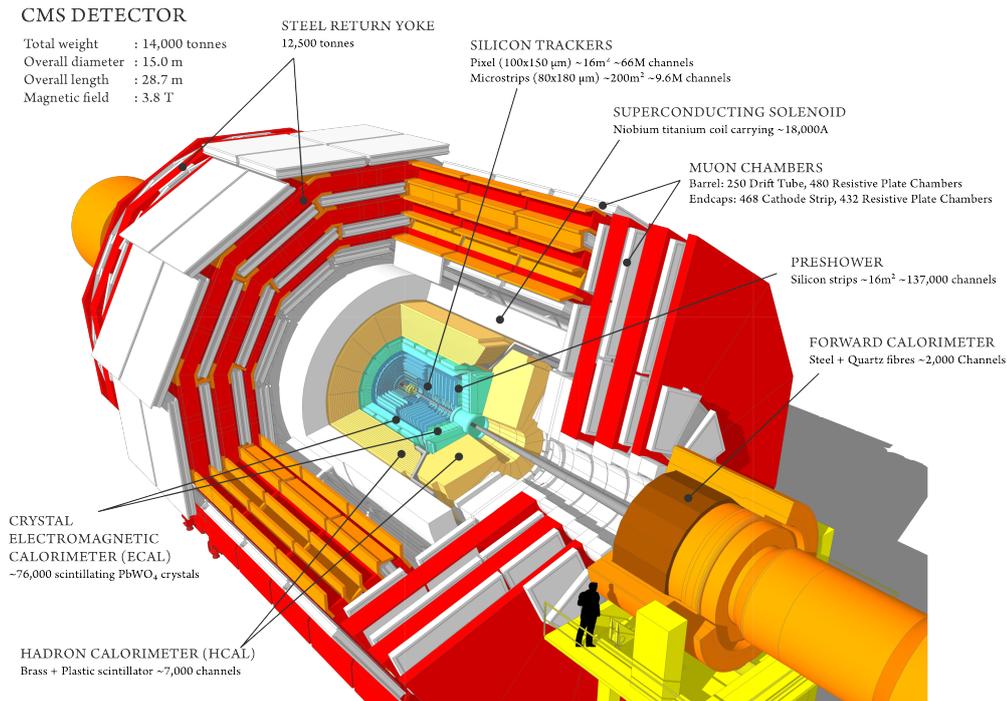


FIGURE 3.6: Sectional perspective view of the Compact Muon Solenoid (CMS) detector. The beams travel in opposite directions along the detector axis and collide at the center. The tracking system is closest to the collision, followed by two different calorimeters. In the outer most are the muon chambers.

the electromagnetic calorimeter (ECAL), generally fully absorbs electrons and photons and the hadronic calorimeter (HCAL), dedicated to measure the energy of strong interacting particles.

For each pp collision, the goal is to count, track and characterize all the elementary secondary particles produced and to measure their physics quantities (position, charge, speed, mass, energy) fully reconstructing the interaction. The particles that can be detected by the CMS experiment are photons (γ), electrons (e^+ , e^-), muons (μ^+ , μ^-), charged hadrons such as protons (p) and pions (π), or neutral hadrons such as neutrons (n). In Figure 3.7 it can be seen how the different particles interact with the different subsystems of the CMS detector.

From the beam line to outward, CMS is composed of the silicon tracker. Detector able to identify the first point of particle's track giving the exactly spatial point on which the interaction took place. The calorimeter (ECAL and HCAL) system, where the vast majority of particles are absorbed, fits the space under the superconducting solenoid which encloses the structure. And around the coil, embedded in the iron yoke of the magnet, there is the muon spectrometer dedicated to muon detection.

The CMS detector design is in many aspects complementary to the ATLAS detector design, both multipurpose detectors at LHC. The most important advantage of CMS over ATLAS is the higher momentum resolution of the whole tracking system, instead of the better energy resolution of the ATLAS hadronic calorimeter. The other

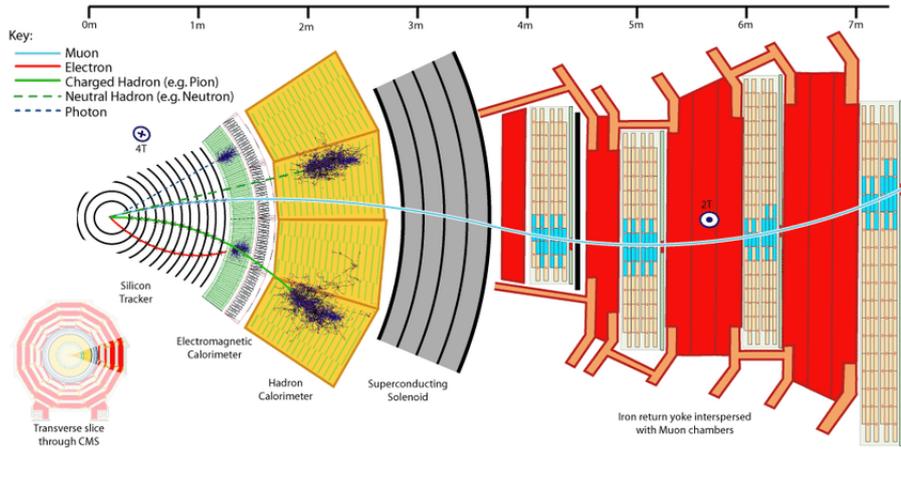


FIGURE 3.7: Transverse section of Compact Muon Solenoid (CMS) detector. The different subdetectors and their interaction with the particles susceptible to be detected.

two experiments of LHC such as ALICE or LHC have an asymmetric shape designed to focus the experiment on specific physics purposes.

The CMS collaboration is made up of more than 3500 scientists, engineers, and students, from 201 institutes in 36 countries (January 2017). Along the next sections the different characteristic of the subdetectors and the coordinate system is briefly described.

3.3.1 Coordinate system convention

The coordinate system adopted by CMS has the origin centered at the nominal collision point inside the experiment, the y -axis pointing vertically upward, and the x -axis pointing radially inward toward the centre of the LHC; the z -axis points along the beam direction toward a geographic point (by convention: towards the Jura mountains from LHC Point 5). The azimuthal angle ϕ is measured from x -axis in the XY -plane and the polar angle θ is measured from the z -axis to the XY -plane.

For convenience a different variable is used to measure the deflection of the prompt particles trajectories with respect to the beam axis, the *pseudorapidity*:

$$\eta = -\ln \tan \left(\frac{\theta}{2} \right). \quad (3.4)$$

This variable will range from 0 for transverse trajectories ($\theta \sim 90^\circ$) of particles arising from lower energy proton collisions and measured in the barrel and ∞ ($\theta \sim 0, 180^\circ$) for particles coming up from high energetic proton collision and travelling down the z -axis in the forward direction. In the limit of the highly relativistic particles typically produced in the LHC's collisions, the pseudorapidity is similar to the *rapidity*, which is defined as:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z c}{E - p_z c} \right) = \tanh^{-1} \left(\frac{p_z c}{E} \right), \quad (3.5)$$

where the difference between the two particles rapidities is invariant with respect to the Lorentz boosts along the z -axis $y'_1 - y'_2 = y_1 - y_2$.

The angle of emission of a particle from an interaction point can be given by the (η, ϕ) coordinates. The Lorentz invariant magnitude dR constructed as:

$$dR = \sqrt{(y^2 + \phi^2)}, \quad (3.6)$$

can be then written in the relativistic limit of the particle momenta as:

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

Equation 3.7 is widely used to measure the angular separation of two particles in the laboratory frame.

The momentum and energy measured in the transverse plane to the beam direction are denoted by p_T and E_T , respectively. They are computed from the x and y components of the three dimensional \vec{p} and \vec{E} measurements. The momentum imbalance measured in the transverse plane is denoted as p_T^{miss} .

The detector requirements for CMS to meet the goals of the LHC physics program can be summarized as follows:

- Good muon identification and momentum resolution over a wide range of momenta in the region $|\eta| < 2.5$, good dimuon mass resolution ($\approx 1\%$ at $100 \text{ GeV}/c^2$), and the ability to determine unambiguously the charge of muons with $p < 1 \text{ TeV}/c$.
- Good charged particle momentum resolution and reconstruction efficiency in the inner tracker. Efficient triggering and offline tagging of τ and b jets, requiring pixel detectors close to the interaction region.
- Good electromagnetic energy resolution, good diphoton and dielectron mass resolution ($\approx 1\%$ at $100 \text{ GeV}/c^2$), wide geometric coverage ($|\eta| < 2.5$), measurement of the direction of photons and correct localization of the primary interaction vertex, π^0 rejection and efficient photon and lepton isolation at high luminosities.
- Good p_T^{miss} and dijet mass resolution, requiring hadron calorimeters with a large hermetic geometric coverage ($|\eta| < 5$) and with fine lateral segmentation ($\Delta\eta \times \Delta\phi < 0.1 \times 0.1$).

The main distinguishing features of CMS are a high-field solenoid, a full silicon-based inner tracking system, and a fully active scintillating crystals-based electromagnetic calorimeter.

3.3.2 Magnet

The solenoid magnet located in the central part of CMS is in terms of size, weight and structural rigidity the major element of the detector and it is used as the principal support structure for all the subdetectors [86].

It is formed by a cylindrical superconducting coil of high-purity stabilised aluminium, a vacuum tank and the magnet yoke. The solenoid has been designed to produce an axial uniform magnetic field of 4 T along the z -axis direction, but after some tests, and taking into account considerations regarding the lifetime of the system, it was decided that the magnet would operate at 3.8 T as the nominal value.

The magnetic flux generated by the superconducting coil is returned via a 1.5 m thick saturated iron yoke. This return yoke is designed as a 12-sided structure and reaches out 14 m in diameter. It also acts as a filter, allowing through only muons and weakly interacting particles such as neutrinos. Due to the general design of the CMS detector, the yoke is split into a cylindrical central part, the barrel, and at the extremities, two endcaps.

The tracker and calorimeters (ECAL and HCAL) detectors are contained inside the magnet whilst the muon chambers are interleaved with the 12-sided iron structure that surrounds the magnet coils and contains and guides the field. The muon chambers have 2T of uniform axial magnetic field. The vacuum tank, made of stainless steel, is cantilevered from the central ring of the barrel yoke. It houses and supports the superconducting coil. The main parameters of the magnet are shown in Table 3.3.

Parameter	Value
Magnetic field	4 T
Inner diameter	5.9 m
Length	12.9 m
Number of turns	2168
Current	19.5 kA
Stored energy	2.7 GJ

TABLE 3.3: Main parameters of the CMS superconducting solenoid.

The high magnetic field produced by CMS is able to bend the trajectory of the high-energy particles arising from a pp collision just starting from the beginning of the curve, at the primary vertex (primary interaction point). This makes the magnet an indispensable part for the identification of the charge and the measurement of the momentum of those prompt particles.

Depending on its momentum, the trajectory of a charged particle is bent inside a magnetic field and depending on whether its charge is positive or negative, this bending is in opposite direction. The relation between the transverse momentum p_T of a particle with charge q moving inside an uniform magnetic field in the z -direction, B_z , and the curvature radius of the particle's trajectory, r , can be inferred from the classic mechanics:

$$p_T = qB_z r, \quad (3.8)$$

which shows that the highest momentum particles have straighter trajectories than the lowest momentum particles that describe trajectories with smaller r and higher curvature (q/p_T).

The magnetic field at CMS has been designed to achieve a momentum resolution of $\Delta p/p \approx 10\%$ at $p = 1 \text{ TeV}/c$, as part of the required unambiguous identification of

the high energy muons with momentum ≈ 1 TeV and the measurement of narrow states decaying into muons.

3.3.3 Inner tracker detector

The particle identification relies on the efficient reconstruction of the tracks inside the detector. This implies among other aspects, the correct identification of the interaction vertex, the impact parameter or the transverse momentum associated to the track. For that reason, a robust tracking and vertex recognition within an intense magnetic field plays an important role for the particles detection.

The inner tracking system is located in the heart of CMS, and it is compound by a silicon pixel and a silicon strip detector systems [87]. In terms of the amount of received flux, they are distributed in three radial regions:

- Close to the interaction vertex, where the particle flux is the highest ($\approx 10^7$ /s at $r = 10$ cm from the vertex), there are placed the pixel detectors. The size of a pixel is $\approx 100 \times 150 \mu\text{m}^2$ giving an occupancy of about $10^{-4}\%$ per pixel per LHC crossing.
- An intermediate region ($20 < r < 55$ cm) that receives low enough flux to enable the use of silicon microstrip detectors with a minimum cell's size of $\approx 10 \text{ cm} \times 80 \mu\text{m}$. The occupancy is of about 2-3 % per LHC crossing.
- And the outermost region ($r > 55$ cm) where the flux has dropped sufficiently to allow the use of larger-pitch silicon microstrips, with a maximum size of $\approx 25 \text{ cm} \times 180 \mu\text{m}$ keeping the 1% of occupancy.

Even in heavy-ion (*Pb-Pb*) collisions, the occupancy is expected to be at the level of 1% in the pixel detectors and less than 20% in the outer silicon strip detectors, permitting track reconstruction in the high density environment.

The pixel tracker detector is the part of the inner tracker system closest to the beam pipe, where the collision occurs. It consists of 3 barrel layers with 2 endcap disks on each side on them. The barrel layers are located at mean radio of 4.4 cm, 7.3 cm and 10.2 cm, extending in length of 53 cm along the z -axis. The pixel shape is "almost" square with a size of $100 \times 150 \mu\text{m}^2$ in (r, ϕ) and z coordinates.

The two end disks with radius of 6 to 15 cm, are placed on each side at $|z| = 34.5$ cm and 46.5 cm. The endcap disks are assembled in a turbine-like geometry with blades rotated by 20° to (see Figure 3.8) to benefit from the Lorentz effect (Lorentz angle $\sim 23^\circ$).

Charged particles passing through the silicon atoms of the sensor ionize the material, creating electron-hole pairs. Each pixel uses an electric current to collect these charges on the surface as a small electric signal. The presence of the magnetic field of the solenoid magnet, perpendicular to this electric current, produces the known "Hall effect" resulting in a transverse force that makes the charges and holes deflecting from their track along the pixel. This shift is called "Lorentz effect", and usually is expressed as an inclined angle. In this case, it determines the shape of the whole sub detector.

To reconstruct the track, it is needed to know which pixels have been touched. And because the detector is made of 2D tiles, rather than strips, and has a number of layers, it can reconstruct a three-dimensional path. The spatial resolution is measured to be about $10\ \mu\text{m}$ for the r - ϕ measurement and about $20\ \mu\text{m}$ for the z measurement.

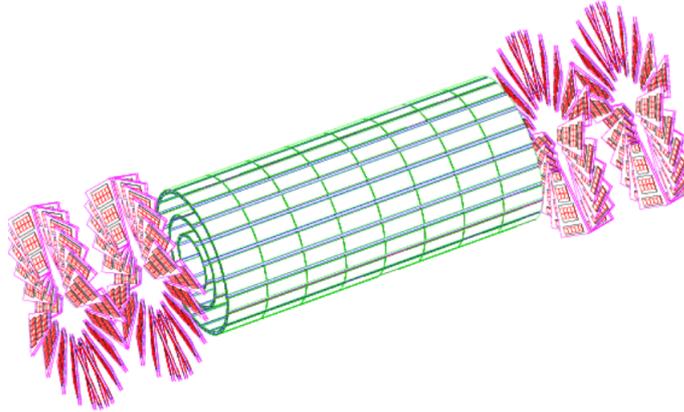


FIGURE 3.8: 3D Layout of CMS pixel tracker detector.

Next to the pixel detector system and outwards from the beam pipe it is situated the tracker silicon strip detector. It is again divided in two regions: the barrel and the endcaps. The entire silicon strip detector consists of almost 15400 modules, with a total of 10 million detector strips, working under the nominal operating temperature of -20°C to mitigate the radiation damage of the materials. The layout of the CMS inner tracker system is shown in Figure 3.9.

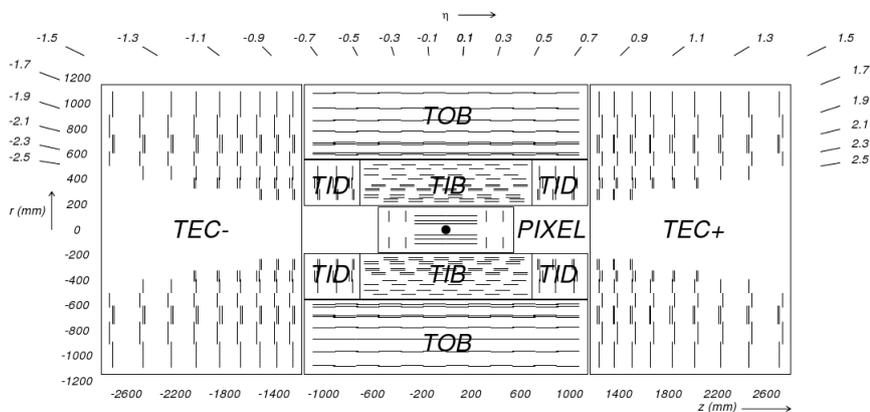


FIGURE 3.9: Transverse view of the CMS inner tracker system. Each line represents a detector module.

The barrel tracker region is divided in two parts: the Tracker Inner Barrel (TIB) and the Tracker Outer Barrel (TOB). The TIB is composed of four layers and covers up to $|z| < 65\ \text{cm}$. The sensors thickness is $320\ \mu\text{m}$ and the strips pitch variate from 80 to $120\ \mu\text{m}$. The TOB comprises 6 layers with a half-length of $|z| < 110\ \text{cm}$. Because the radiation levels are smaller in this region, the thickness of the silicon sensors can be enlarged, $500\ \mu\text{m}$, for longer strip length and wider pitch, between 120 and $180\ \mu\text{m}$.

In both cases the measurements are provided in 3 dimensions (r, ϕ) and z coordinates) and the single point resolution varies from 23 - $34\ \mu\text{m}$ in the (r, ϕ) plane and

230 μm in the z direction for the TIB and from 32-52 μm in the (r, ϕ) plane and 530 μm in the z direction for the TOB.

Finally the Tracker Endcaps (TEC) and the Tracker Inner Disk (TID) complete the tracker design. Each TEC comprises 9 disks that extend into the region $12\text{ cm} < |z| < 280\text{ cm}$, and each TID comprises 3 small disks that fill the gap between the TIB and the TEC. The TEC and TID modules are arranged in rings, centred on the beam line, and have strips that point towards the beam line, therefore having a variable pitch. The thickness of the sensors is 320 μm for the TID and the 3 innermost rings of the TEC and 500 μm for the rest of the TEC.

The inner tracker comprises 66 million pixels and 9.6 million silicon strips, what gives accuracy description of the particle's path. The total active area of the pixel detector is $\approx 1\text{ m}^2$, whilst that of the silicon strip detectors is 200 m^2 , providing coverage up to $|\eta| < 2.5$. Figure 3.10 shows the performance of the tracker system in events with reconstructed muons coming from a Z boson decay.

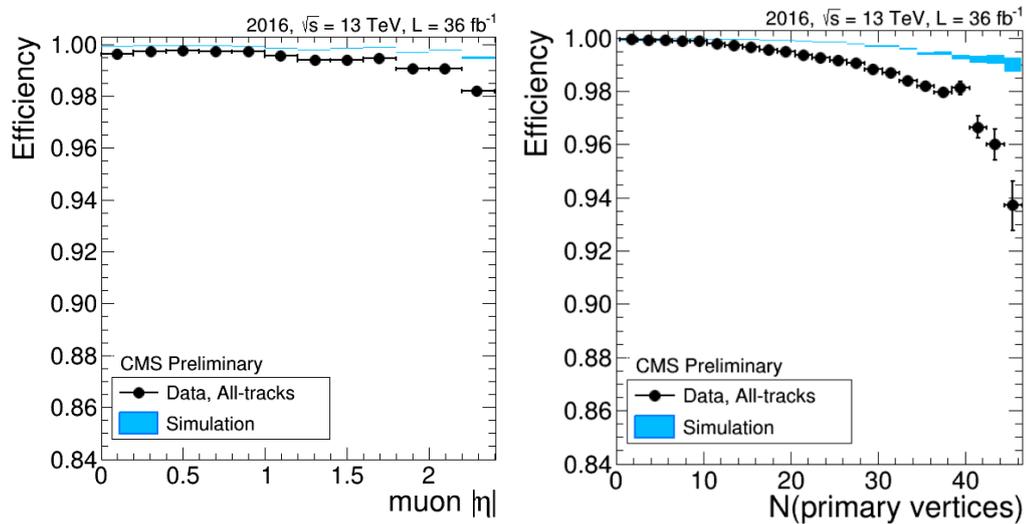


FIGURE 3.10: Tracking efficiency for reconstructed muons coming from the Z boson decay using the tag-and-probe technique [88] for data (black dots) and simulation (light blue rectangles) as a function of the absolute pseudorapidity of the probe muon (left) and the number of primary vertices in the event (right) [89].

3.3.4 Electromagnetic calorimeter

The electromagnetic calorimeter [90] is the essential component of the CMS detector dedicated to measure the energy of the electrons and photons arising out of the LHC. Surrounding the tracking system and fitting within the hadronic calorimeter and the magnet coil, the ECAL is the first of the two calorimeters in the CMS detector. It consists of an hermetic, homogeneous scintillating-crystal calorimeter, divided into the barrel region and two endcaps.

Due to the difficult experimental environment that the LHC imposes, the requirements on the detector specifications have been largely studied. The material finally chosen for the ECAL crystals has been the lead tungstate (PbWO_4). This material is radiation hard and assures good performance in the strict conditions at LHC. It

has short radiation length ($\chi_0 = 0.89$ cm) and a small Molière radius (2.2 cm) that allows fast scintillators with high energy resolution of each particle in the event. These high-density crystals create a compact detector.

When the high-energy electrons or photons going through the tracker arrive at ECAL they interact with the atoms of the crystals generating an electromagnetic shower of photons and electrons inside the material. The amount of light produced in this process is proportional to the energy that was deposited in these crystals by the incoming electron or photon.

Because the yield of the scintillation light is relatively low, it is required the use of photodetectors that convert the light measurement to electrical signal and then amplify it. These devices are directly placed over the back of each calorimeter crystal and will operate in high radiation conditions within a high magnetic field. Those chosen types are the silicon avalanche photodiodes or APDs for the barrel region and the vacuum phototriodes (VPTs) for the endcaps, because here the radiation is too high for using silicon photodiodes. Both the crystals and the APDs are very sensitive to temperature changes and require a high thermal stability. The cooling system maintains the temperature of 100 tonnes of crystal to within 0.1°C for good performance.

The barrel section (EB) has an inner radius of 129 cm. It is compounded by 61200 lead tungstate crystals assembled in 36 larger structures known as “supermodules”. Each of them covers half the barrel length and corresponding to a pseudorapidity interval of $0 < |\eta| < 1.479$. The crystals are quasi-projective (the axes are tilted at 3° with respect to the line from the nominal vertex position) and cover about 1°C in $\Delta\phi$ and $\Delta\eta$. The crystals have a front face cross section of about 22×22 mm² and a length of 230 mm, corresponding to $25.8\chi_0$.

The endcaps (EE) are at a distance of 314 cm from the vertex and covering a pseudorapidity range of $1.479 < |\eta| < 3.0$. Each endcap consists of two semi-circular aluminium plates from which are cantilevered structural units of 5×5 crystals, known as “supercrystals”. The 7324 crystals of the endcap are arranged in the x - y grid, instead of the η - ϕ grid used in the barrel. They are all identical and have a front face cross section of 28.6×28.6 mm² and a length of 220 mm ($24.7\chi_0$).

To complete the assembly, the CMS’s ECAL poses a preshower detector. Its function is mainly to identify energetic long live particles such as neutral pions (π^0) decaying into two low energy photons that leave tracks with smaller angular separations than those distinguished by the other systems in the ECAL. The preshower system has been included in front of the crystals in both endcaps covering a pseudorapidity range of $1.7 < |\eta| < 2.6$. The preshower has a much finer granularity than the rest of the ECAL devices and in general can give better spatial resolution for high energy forward particles.

The active elements of this device are two planes of lead (absorber at depths, of a total thickness, of $2\chi_0$ and $1\chi_0$), each followed by one detector plane of silicon strips with a pitch of 1.9 mm. The impact position of the electromagnetic shower is determined by the centre-of-gravity of the deposited energy. In order to correct the energy deposited in the lead converter, the energy measured in the silicon is used to apply corrections to the energy measurement in the crystal. Combining the energy measurement of the two detector planes the particle’s path can be accurately defined.

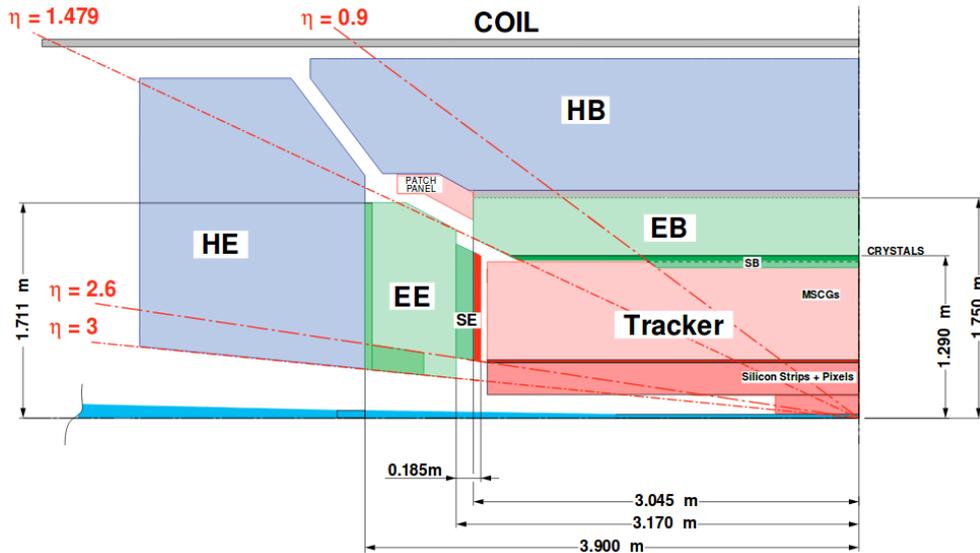


FIGURE 3.11: Layout in the (r,z) plane of the calorimeter system and the inner tracker system of CMS. In the case of the HCAL, it is shown only the central eta region (< 3.0) under the coil.

The complete preshower system forms a disc, about 2.5 m in circumference with a 50 cm diameter hole in the middle (where the beam pipe passes through). As in the tracker, the silicon detectors must be kept at a temperature of between -10°C and -15°C . However, the nearby ECAL is very sensitive to the temperature and must be kept within precisely 0.1°C , so, the preshower must therefore be cold on the inside but warm on the outside, achieved using both heating and cooling systems. In the Figure 3.11 can be seen a schema of the calorimeter system and the tracker.

The reconstruction algorithms starts combining the energy deposits of individual crystals into energy clusters which are then grouped together into superclusters, serving as origin point for photons and electrons identification. The energy resolution is measured by fitting a Gaussian function to the reconstructed energy distributions. It has been parametrized as a function of energy,

$$\frac{\sigma^2}{E} = \left(\frac{S}{\sqrt{E}} \right)^2 + \left(\frac{N}{E} \right)^2 + C^2, \quad (3.9)$$

where S is the stochastic term, N the noise and C the constant term (due to detector non-uniformity and calibration uncertainties). Figure 3.12 shows the ECAL performance in 2016 through the invariant mass of photon pairs reconstructed in one crystal of the ECAL barrel (left) and endcaps (right). The events has been selected in the mass range of π^0 , and correspond to an integrated luminosity of approximately 100 pb^{-1} . These events are used as prompt feedback to monitor the calibration of ECAL crystals.

3.3.5 Hadronic calorimeter

The hadronic calorimeter [92] is the last step for most of the secondary particles travelling outwards the CMS detector. The exception are the muons which can travel

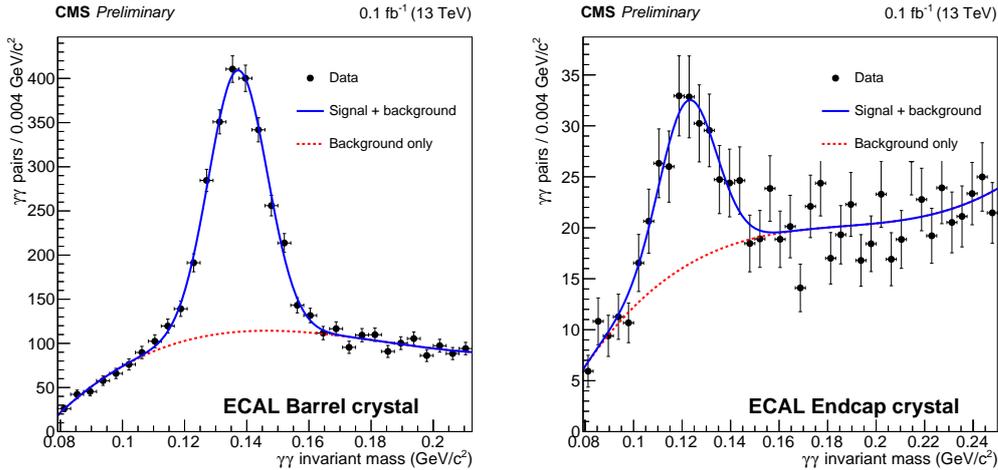


FIGURE 3.12: Invariant mass of photon pairs reconstructed in one crystal of the ECAL barrel (left) and endcaps (right). Simulation (color lines) and data (black dots) is compared [91].

several meters through dense material, and the invisible particles such as the neutrinos that, extremely soft interacting, cannot be directly detected by the CMS experiment, but which leave in the event a momentum and energy imbalance in the transverse plane of the event.

The HCAL is in charge of measuring the energy and direction of hadrons and the missing transverse momentum flow. The missing transverse energy flow measurement will form a crucial signature for new physics searches such as supersymmetry searches, where the new particles would appear as an imbalance in the transverse energy. The HCAL will also help with the identification of electrons, photons and muons, together with the tracker, ECAL and the muons spectrometer.

The HCAL system surrounds the ECAL and it is placed just before the magnetic coil. It can be considered in two pieces: the central calorimeter region ($|\eta| < 3.0$) in which it is required an excellent jet identification and moderate single particle and jet resolution, and the forward/backward calorimeter (HF) ($3.0 < |\eta| < 5.0$) with modest hadron energy resolution but with good jet identification capability. The forward calorimeter is physically separated from the central calorimeter with its front face located at ± 11 m from the interaction point.

The central calorimeter consists of the Hadron Barrel (HB) and Hadron Endcap (HE) calorimeters, both located inside the CMS magnet cryostat. An additional layer of scintillators, referred to as the hadron outer (HO) detector, lines the outside of the coil with the aim of measuring late shower development and ensuring the total shower energy containment. It covers the region $|\eta| < 1.26$.

The central hadron calorimeter (HB, HE) is a sampling calorimeter consisting of alternating, parallel to the beam axis, plates of brass (or stainless steel) absorber and plastic scintillator tiles, as active material, embedded with Wavelength Shifting (WLS) optical fibers for the readout. The absorber plates are 5 cm thick in the barrel and 8 cm thick in the endcap. This material has a reasonably short interaction length, it is easy to machine, and it is non-magnetic. The plastic scintillator tiles are 4 mm thick.

The HCAL has been built to identify the particle's position, energy and arrival time throughout its deposits along the layers. When a hadron hits the absorber plate, both charge or neutral particles, strong interact with nucleons of the atoms loosing its energy and generating new particles in cascade. The charged secondary particles can excite the atoms in the tiles of plastic scintillator which recover their normal state emitting light of certain wavelength. This light is detected by the tiny optical fibres WLS that carry the signal away to readout boxes where is amplified using photodetectors.

The fibers from the different layers of tiles comprising a ϕ, η depth segment are re-organized into towers, and the light from all the tiles making up a *tower* is optically mixed and sent to an optical transducer (multi-channel Hybrid Photodiode HPD). This photodetection device is specially configured to operate in high magnetic field, and give amplified response in proportion to the original signal for a large range of particles energies. The amount of light in a given region is summed up over many layers of tiles in depth, called a "tower", this total amount of light is a measurement of a particle's energy.

The HB is constructed in two half barrels, it covers the pseudorapidity region $|\eta| < 1.4$ while each part of the covers the region $1.3 < |\eta| < 3.0$. The HO detector contains scintillators with a thickness of 10 mm which are grouped in tiles along the $-1.26 < \eta < 1.26$ matching the ϕ segmentation of the DT chambers. It increases the effective thickness of the hadron calorimeter reducing the energy resolution function and improving the p_T^{miss} resolution.

To finish the design, the HF calorimeter consists of a large steel structure with embedded hard quartz fibers which constitute the active material of the detector. The quartz fibers run longitudinally through the absorber and parallel to the beam axis providing a fast collection of the Cherenkov light produced by the incident particles. This light is then channeled to the photodetectors and forms the basis of the energy measurement. The HF calorimeters are arranged with a cylindrical symmetry around the beam line. The radius of the active part of the HF is 1.4 meters. The length, along the beam, is 1.65 m, or about 10 nuclear interaction lengths. This is largely sufficient to longitudinally contain the Cherenkov signal produced by hadrons of up to 1 TeV. The central region is open (25 cm in diameter) to allow for the beam pipe (20 cm in diameter).

The overall assembly enables the HCAL to be built with essentially no instrumented cracks or dead areas in ϕ . The gap between the barrel and the endcap HCAL, through which the services of the ECAL and the inner tracker pass, is inclined at 53 degrees and points away from the centre of the detector. In figure 3.13 can be seen an schema of the whole system.

Because of the structure and non-compensating nature of the CMS calorimeters, there are different sources of non-Gaussian tails in the energy distributions. An important requirement of HCAL is to minimize the non-Gaussian tails in the energy resolution and to provide good containment and hermeticity for assuring the right p_T^{miss} measurement. Hence, to capture as extent as possible every particle emerging from the collisions, the HCAL design maximizes the material inside the magnet coil in terms of interaction lengths becoming a very compact detector.

In Figure 3.14 the HCAL performance in 2016 data is shown in terms of the p_T^{miss} resolution using $Z \rightarrow e^+e^-$, $Z \rightarrow \mu^+\mu^-$ and photon events. The transverse momentum of the boson is denoted by q_T , and the transverse and longitudinal momenta of the

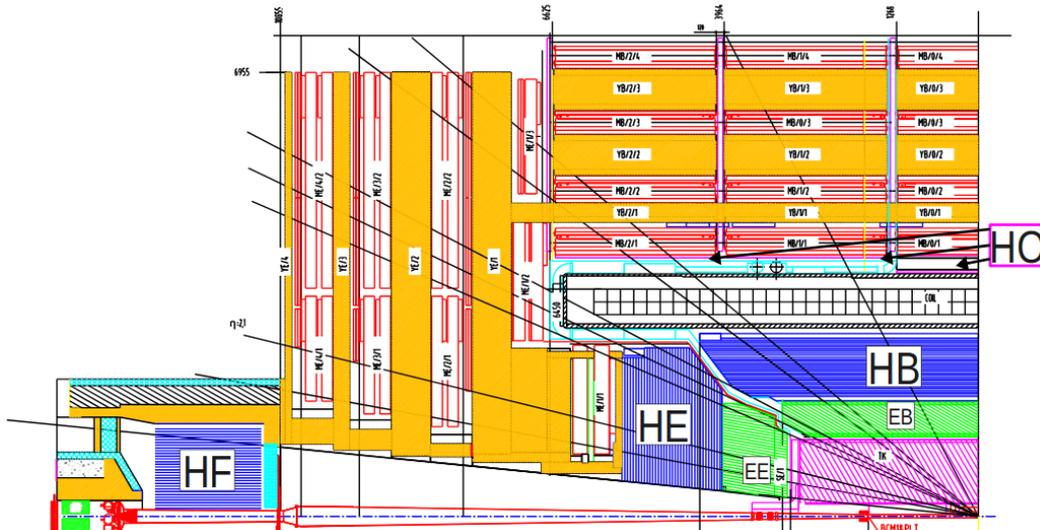


FIGURE 3.13: Transverse view of the hadronic calorimeter distribution in the CMS detector.

hadronic recoil system are denoted respectively by u_T , and $u_{||}$. Please refer to the following reference for more information [93].

3.3.6 Muons system

The outermost detection system of CMS is dedicated to the muons identification. These particles interact with the matter softer than the other charged and/or massive particles and can penetrate several metres of dense material without stop. While most of the particles end their trajectory in the calorimeters, the muons can go beyond and leave clean records of their paths in the CMS muon detection system [94].

The tracks of centrally produced muons are measured three times: in the inner tracker, after the coil, and in the return flux. The precise muon momentum is measured combining the inner tracker segments with the track segments reconstructed in the muon system. This relation requires to place stringent constraints on the alignment of the tracker relative to muon chambers.

The momentum of the muons reconstructed by the muon system alone (*standalone muons*) is measured essentially by the muon bending angle after the coil, but only for high momentum (> 200 GeV) the result is not spoiled by the multiple scattering interactions in the material before the first muon station. On the opposite, for low momentum muons reconstructed using only the silicon (inner) tracker measurements (*tracker muons*) the momentum resolution is by an order of magnitude better than for low momentum standalone muons. In the case of high momentum muons, the extrapolation of the muon trajectory back to beam-pipe is possible due to the compensation of the bend before and after the coil when multiple scattering and energy loss can be neglected. This improves the momentum resolution for these muons as can be seen in Figure 3.15.

The muon system is compound by three types of gaseous detectors: the drift tubes (DT), the cathode strip chambers (CSC), and the resistive plate chambers (RPC).

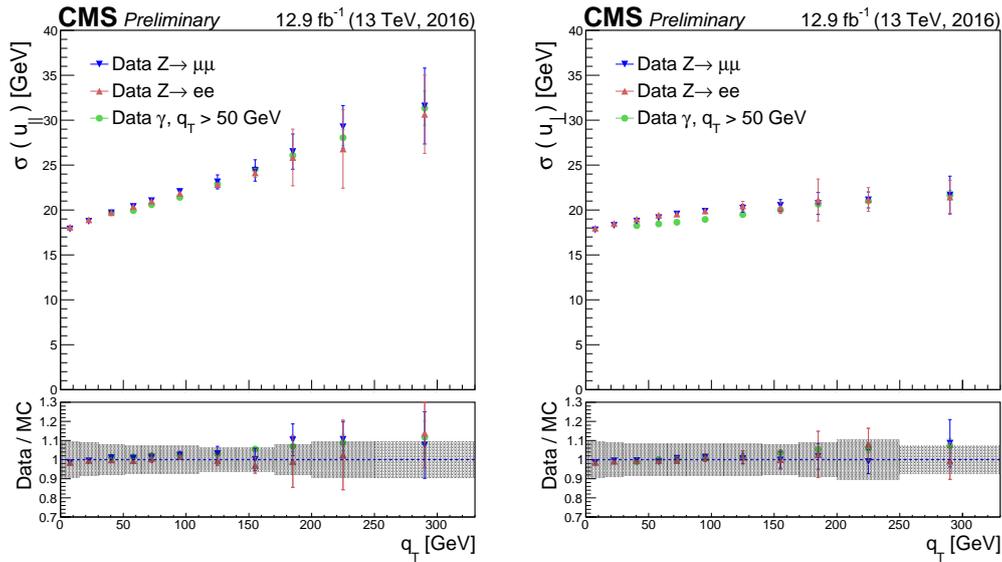


FIGURE 3.14: Observed resolution curves of p_T^{miss} as a function of q_T , u_T , and $u_{||}$ for $Z \rightarrow e^+e^-$, $Z \rightarrow \mu^+\mu^-$ and photon events. The lower panel shows the ratio of data to simulation values. The error band displays the systematic uncertainty of the simulation, estimated as the $Z \rightarrow e^+e^-$ channel systematic uncertainty [93].

These systems are distributed around the barrel and the endcaps. The choice of the different detector technologies has been driven by the very large surface they have to cover and by the different radiation environments. The layout of one quarter of CMS muon detection system according to η , r and z coordinates can be seen in Figure 3.16.

The Muon Barrel (MB) region, taking up to $|\eta| < 1.2$, shelters the DTs. Each DT is a $13 \times 42 \text{ mm}^2$ tube filled out with the nominal gas mixture of Ar (85%) and CO_2 (15%). The maximum drift distance is 2.1 cm, approximately 400 ns maximum drift time. The whole system consist on 250 chambers organized in four stations (MB1, MB2, MB3 and MB4 with the last being the outermost) arranged in cylinders interleaved with the iron yoke. In this central part of the detector the muon rate is low and the residual magnetic field in the chambers is also low. Each station is designed to give a muon vector in 2D space, with a ϕ precision better than $100 \mu\text{m}$ in position and approximately $1 \mu\text{rad}$ in direction.

In the Muon Endcaps (ME) region where the muon rate and the residual magnetic field are high, the CSCs are deployed covering the region up to $|\eta| < 2.4$ on each of the endcaps. Every CSC is a trapezoidal shape chamber of maximum length of 3.4 m and maximum width of 1.5 m, containing six detector layers inside. Each layer is composed by an anode wire plane stretched between two planar copper cathodes, one continuous, the other segmented in strips. That strips run almost perpendicularly to the wires giving a 2D position measurement. The spatial resolution provided by each CSC is typically about $200 \mu\text{m}$ with an angular resolution in ϕ of order $10 \mu\text{rad}$. The nominal gas mixture used to fill the gaps between the layers is 40% Ar , 50% CO_2 and 10% CF_4 .

The total ME system comprises 468 CSCs in the two endcaps arranged again in four

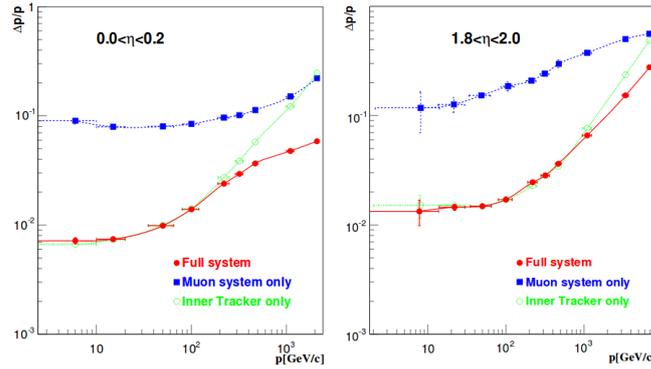


FIGURE 3.15: Muon momentum resolution as a function of the momentum value p . The measurements have been done with simulated data using the muon system only, the inner tracker only or both combined (full system). The barrel corresponds to $|\eta| < 0.2$ (left) and the endcaps to $1.8 < \eta < 2.0$ [84].

stations of chamber (ME1, ME2, ME3 and ME4) which are mounted on disks enclosing the CMS magnet in the perpendicular plane to the beam. In each disk, the chamber are divide into concentric rings, three rings in the innermost station, and two in the others. In addition to this, the RPCs are used in both the barrel and the endcap regions. They are gaseous parallel-plate detectors made out of phenolic resin (bakelite) with a bulk resistivity of $10^{10} - 10^{11} \Omega\text{cm}$, separated by a gas gap of a few millimeters stuffed with a gas construction mixture of 95% $\text{C}_2\text{H}_2\text{F}_4$ (Freon), 5% $i\text{C}_4\text{H}_{10}$ (Isobutan). They are fast gaseous detector that provides muon trigger system parallel to those in the DTs and the CSCs systems. They combine a good spatial resolution with a time resolution of just one nanosecond (one billionth of a second).

The DTs or CSCs and the RPCs operate within the first level trigger system, providing two independent and complementary sources of information. The complete system results in a robust, precise and flexible trigger device. In total, the muon system contains of order 25000 m^2 of active detection planes, and nearly one million electronic channels.

The performance of muon reconstruction as a function of the alignment position error for 2016 is in Figure 3.17. Here, it can be seen that the poorest measurements comes from the transition part between the barrel and endcaps region.

3.3.7 Trigger and data acquisition system

In LHC, every second, a great number of collisions take place, producing more data than those can be stored. The trigger and data acquisition systems (DAQ) [96–99] are in charge of selecting, processing and storing only the relevant information for the CMS collaboration, besides of controlling the performance of the detector.

At nominal luminosity, for a beam crossing frequency of 25 ns, the average number of effective pp interactions per bunch crossing is about of 20. If it is considered that the LHC's crossing rate is of about 40 MHz, however, for practical reasons (see Section 3.2.1), the average bunch crossing frequency is about 30 MHz, with a huge amount of $\sim 10^8$ proton-proton collisions per second. But, the maximum input rate

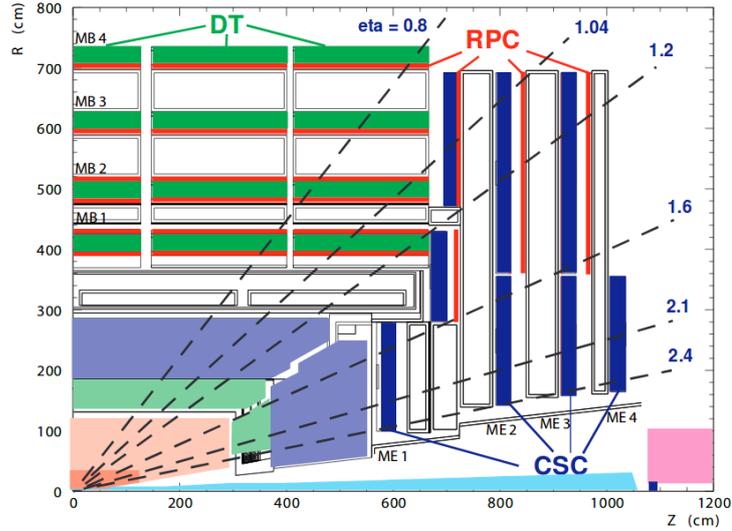


FIGURE 3.16: Quarter of CMS muons detection system transverse view. The Muon Barrel region corresponding to $|\eta| < 1.2$ includes the whole DT system. The ME region is designated to the RPC and CSC systems.

that can be stored by the on-line computer farm is 100Hz, what means that from the nominal LHC rate a rejection factor of at least $O(10^5)$ must be applied.

Although all this volume of events seems to be potentially interesting, not all the collision give enough information for physics purposes, events from low energy glancing collisions, for instance, can be discarded with minimum loss of physical content, keeping only the most essential data.

The trigger and data acquisition system is designed to inspect the detector information looking for remarkable features of each event. Based on physics selection algorithms it can make quick decisions in order to accept or refuse to store the event.

The data acquisition system is a very large and complex computing system. It is designed flexible and modular enough to be adapted to new technologies and different machine operating conditions. It controls the detector elements readout, the selection and the means of archiving the events in mass storage. Another crucial function of the DAQ system is the operation of the Detector Control System (DCS) which is in charge of the performance of the whole CMS experiment and guarantees a CMS safe operation to obtain high quality physics data.

Because of the rejection power of $O(10^5)$ is high enough to be achieved efficiently in a single processing step, the full selection task is split into two steps. The first step, Level-1 (L1) trigger, is designed to reduce the rate of events accepted to less than 100 kHz and keep them for further processing. The second step, High-Level Trigger (HLT), is designed to reduce this maximum Level-1 accepting rate of 100 kHz to a final output rate of 100 Hz.

The CMS trigger and data acquisition system consists of four parts: the detector electronics, L1 trigger processors (calorimeter, muon and global), the readout network, and an online event filter system (processor farm) that executes the algorithms for the HLT processing.

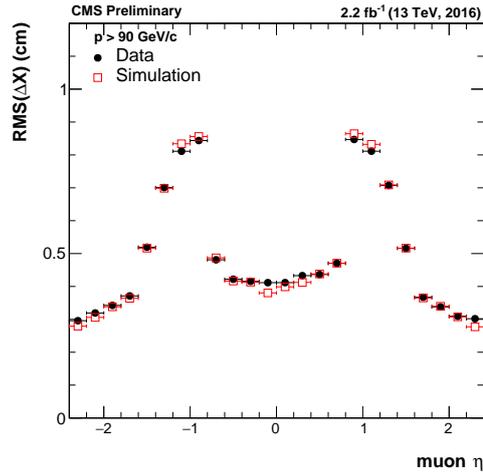


FIGURE 3.17: Root mean square width of residuals of the local- x position for the track-to-segment match in the first muon station as a function of the muon pseudorapidity, with a requirement on momentum $p > 90$ GeV. Data (with asymptotic alignment) are compared with Monte Carlo expectations [95].

Level-1 trigger

The Level-1 trigger is implemented in a customized hardware to read out the information from the calorimeters and the muon system also taking into account the correlation of the information between these systems. Its decisions are based on the presence of the *trigger primitive* objects, such as electrons, muons, photons and jets above a set of E_T or p_T thresholds, besides of, the global sums of E_T or p_T^{miss} of the event. Figure 3.18 shows the L1 trigger schema and data flow.

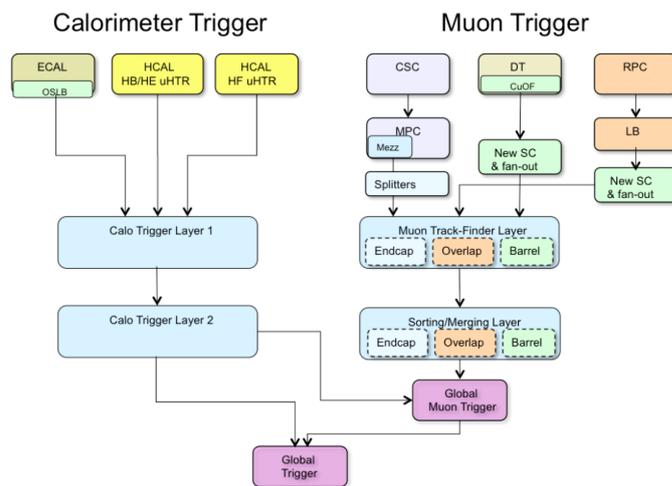


FIGURE 3.18: Structure of the Level-1 trigger of CMS. The arrows corresponds to the data flow [99].

The data flow starts in the detector, where the trigger primitive information is sent to the corresponding sub detector L1 trigger processors. The system spends about $1\mu\text{s}$ in reaching the decision of discard or keep one event. Because of the size of the CMS detector and the underground caverns, a characteristic transit time is imposed

in the signal propagation which, combined with the electronics, gives a latency time of $3.2\mu\text{s}$. This is the time the system needs to send the signal from the front-end readout detector electronics to the services hosting the L1 trigger logic, make the decision and return back to the detector front-end electronics.

While small portion of events are accepted by the L1 trigger, they are still too many for the storage capability. The data of an accepted event is kept in temporary memory devices and sent to the HLT system, which will reduce by 1/1000 this output rate for mass storage.

To assure the high efficiency in the trigger performance, the physics requirements on the L1 trigger are:

- The CMS trigger system should be capable of selecting leptons and jets over the pseudorapidity range $|\eta| < 2.5$ with an efficiency which is very high, above a selected threshold in transverse momentum.
- For the single lepton triggers it is required that the trigger is fully efficient ($> 95\%$) in the pseudorapidity range $|\eta| < 2.5$ with a threshold of $p_T > 40$ GeV.
- For the dilepton trigger, it is required that the trigger is fully efficient ($> 95\%$) in the pseudorapidity range $|\eta| < 2.5$ with thresholds of $p_T > 20$ and 15 GeV for the first and second leptons respectively.
- Single photon and diphoton triggers are required to have thresholds similar to those of the leptons.
- Single and multiple jet triggers are required with a well defined efficiency over the entire rapidity range $|\eta| < 5$ in order to reconstruct jet spectra that overlap with data attainable at lower energy colliders such as the Tevatron. For higher transverse momenta the jet trigger should also be fully efficient.
- A missing transverse energy trigger with a threshold of about 100 GeV is required.

High Level Trigger

After L1 trigger selection, the data is forwarded from the end-front electronic to the HLT system implemented in software. The further selection is based on better granularity and resolution information than that available for the L1 trigger, including also information from the tracking detectors. It will receive on average one event every $10\mu\text{s}$. Each event has a size of about 1.5 MB (in proton-proton interactions). The HLT system uses on-line, physics decision algorithms almost as sophisticated as those used in the offline final reconstruction. The processing time required by the HLT for analyzing one event can go up to 1s . This implies the use of a processor farm, where each processor is fully programmable and runs the same HLT software code. The flexibility of its environment provides the HLT algorithms with the necessary adaptability to the changes in the experiment conditions and physics requirements.

The CMS event selection is kept as broad and inclusive as possible in order to the general purpose of the detector. The number of “trigger levels”, that will be used to achieve the rejection factor from Level-1 output rate of 10^5 Hz to the final storage rate of 10^2 Hz, is related to the diversity on the type of events to kept. In CMS detector two distinct trigger systems, Level-2 trigger and Level-3 trigger, are used.

The data processing of the HLT is structured around the concept of *HLT path*. The HLT path is a set of algorithm processing steps run in a sequence of increasing complexity in physics reconstruction and selection. Each HLT path is seeded by L1 objects with thresholds lower than those imposed at higher levels.

Usually, even after the HLT trigger selection, the stream of data is still high enough to saturate the data taking bandwidth. In that cases, the trigger algorithm can be prescaled, or blocked once a defined number of events of a defined data range have been recorded, i.e., 1 event over 10, 10 over 1000, ..., avoiding the saturation.

The on-line selected datasets are determined by CMS physics priority and usually should contain one lepton events, multi-lepton events, photons, leptons plus jets or only jets, and/or events with determined magnitude of p_T^{miss} . The data is naturally arranged in Primary Datasets (PDs) based on trigger paths.

3.3.8 Luminosity measurement

The integrated luminosity measurement during the 2016 LHC data taking is based on the pixel cluster counting method, and the absolute luminosity scale calibration is derived from Van der Meer Scans [100].

The silicon pixel detector is used to measure the instantaneous luminosity evaluating the number of pixel clusters occurring on average in a zero-bias triggered event as:

$$\mathcal{L} = \frac{\nu \langle n \rangle}{\sigma_{vis}}, \quad (3.10)$$

where ν is the beam revolution frequency, $\langle n \rangle$ is the average number clusters per event and σ_{vis} is the visible inelastic cross section, defined as the average number of clusters per inelastic collision times the total inelastic cross section. Finally, σ_{vis} is calibrated by measuring the interaction rate (with the same $\langle n \rangle$) as a function of the the transverse beam separation. For more information see [101] and [102].

The minimal range of time considered for the estimation of the integrated luminosity is called *luminosity section* (LS), defined as the time necessitated by each beam to complete 2^{18} LHC orbits which corresponds to $t_{LS} = 23.31$ s. For each LS, the value of $\langle n \rangle$ is measured, then the instantaneous luminosity \mathcal{L} resulting from the Equation 3.10 is multiplied by t_{LS} to obtain the integrated luminosity for that LS. The final integrated luminosity for the analysis is computed summing the integrated luminosity of each LS recorded by the CMS experiment during the corresponding data taking. In the case of this analysis the precise period can be seen in Figure 3.5.

Chapter 4

Analysis framework

One of the most important parts in the experimental procedure, from the physics point of view, is the data taking and the data management. The experimental information has to be precisely recorded, correctly stored and carefully manipulated. In the case of CMS, due to the high complexity of the experiment, the computing environment development represents a great challenge .

The CMS computing model must support the storage, the transfer and processing of all collected data. It has to provide the framework for the online CMS Trigger and DAQ operations, as well as ensure the offline physics analyses related activities and the production and distribution of simulated data. Among others requirements, the considerable amount of information to take care of, the complex reconstruction algorithms, the large number of users, or the expected longevity of the experiment, are all of them demanding conditions for a large scale system with high flexibility and very manageable.

The CMS computing model has been designed as a distributed system based on Grid middleware, under a hierarchical architecture of centers located around the world. The Grid infrastructure for both data storage and analysis is built and administrated by the Worldwide LHC Computing Grid (WLCG) project [103], for the entire HEP community that uses the LHC.

In Section 4.1 of this chapter an overview of the CMS event data model and the distributed database system will be introduced. The CMS physics event reconstruction algorithm and the objects reconstruction procedure can be found in Section 4.2. Finally, in Section 4.3, the Monte Carlo simulated data production will be described.

4.1 Event data model and data tiers

The physics analyses combine a wide variety of reconstructed information from the recorded detector data and the Monte Carlo generated data. This information can be separated in two classes: "event data" and "non-event data".

The CMS event data model is centered around the concept of "Event", a computational object that holds all the physical information of the triggered physics event as well as all the information derived from the data taking itself.

The Event provides access to the whole event description, such as raw digitised data, reconstructed products, simulation products or high-level analysis objects, relating to experimental or simulated pp interactions. It also contains the reconstructed data

conditions like calibration and alignment information, detector status, or the software configuration used to create each data product. The former information, included in the category of "non-event data", is needed to fully understand the physics data collected from the detector.

The Event content is accessed by a sequence of independent modules, each of them covering a particular event-processing functionality. Those include: event data producers, which add new data products into the Event, filters, that work in online triggering and offline selections, the analyzers, that study the features of the event, or input and output modules for both DAQ and data storage. Such modules are allowed to communicate to each other only through the object Event. The exact number and the sequence of modules to use is defined by the user. This modular framework is as easily adaptable to new developments and conditions as is required for the complex experiment.

The CMS Trigger and Data Acquisition system controls the event filtering, the data flow and the storage management. It uses real-time detector data from the front-end electronics to generate a primary set of reconstructed data that will be further used by the physicists.

The action of data reconstruction is referred to all the operations of constructing physics entities from the detector readouts. The process of reconstruction is performed in different steps that successively create higher-level reconstructed units, at the end being suitable for high-level triggering or physics analysis. That is therefore a procedure of data reduction whose main objective is extract the underlying physics objects such as tracks, particles, vertices,..., emerged from the collision.

Due to the flexibility and level of data reduction required by CMS, several data formats with different levels of detail and precision are constructed. These data are arranged into a hierarchy of data tiers. Each data tier represents an event size and a step in the reduction process, which is typically carried out in the Grid. The main data path are summarized next:

- DAQ-RAW: the primary record of a physics event. Detector data from front-end electronics and L1 trigger results. It will be the input for the HLT.
- RAW: this is the primary archive of events. It would contain the detector data after online formatting, the L1 trigger result, the result of the HLT selections, and potentially some of the higher level quantities calculated during HLT processing.
- RECO (Reconstructed): contains high-level selected objects and a full record of hits and clusters used to produce them. Enough information for subsequent reconstruction is retained.
- AOD (Analysis Object Data): a subset of RECO format is destined for analysis. The events contain only high-level physics objects plus additional information for kinematics refitting.

The CMS computing environment, constructed as a distributed system, makes use of different computing centres spread throughout the world interacting one to each other through the Grid services. CMS uses this connectivity of resources to perform the data processing, data archiving, Monte Carlo event generation, and all kind of

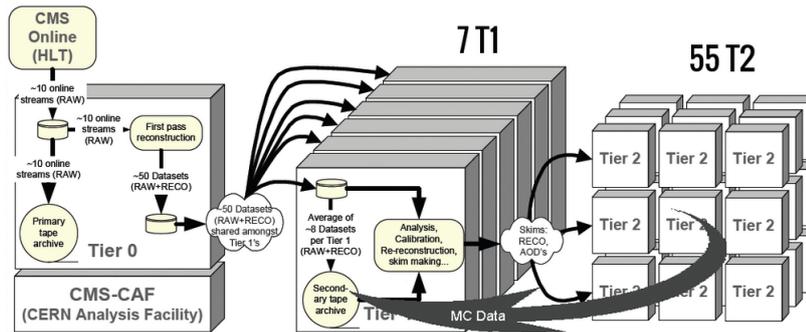


FIGURE 4.1: Experimental data stream. At CERN, Tier-0 has access to new data, producing and storing the RAW and first-pass RECO datasets. A copy of them is sent to long-term storage in Tiers-1 centres in national labs. The final reconstructed and simplified data formats are sent to Tier-2 centres. The Monte Carlo generated data is typically produced in Tier-2 centres and archived in their associated Tier-1 centres.

computing-related activities. This network of centres is organized using a hierarchical structure related to data tiers. The detector data flow through the hardware tiers is following itemized and next shown in Figure 4.1:

- Tier-0 centre hosted at CERN: accept data from the online system and carry out the prompt reconstruction of the raw data to produce a first pass of RECO dataset. Export a copy of RAW and RECO data to Tier-1 centres.
- Tier-1 centres hosted by CMS collaborating national labs: provide of long term storage of RAW data and holds a fraction of CMS simulated and RECO data and a complete copy of AOD data. Carry out the second-pass RECO reconstruction. Store and serve to Tier-2 centres.
- Tier-2 centres hosted at CMS institutes: support the local analysis or specialized activities an MC production for the whole experiment.

4.2 Event reconstruction

The reconstruction of high-level objects requires the combination of different levels of information. The algorithms first use real or simulated data for a local reconstruction in the subdetector modules. The outputs typically are the position measurement in tracking-type detectors (Tracker and Muon spectrometer) and clusters of energy deposition in the calorimeter-type detectors (ECAL and HCAL).

During a more evolved step, all this material from several modules of the same sub-detector is thus combined providing the reconstructed entities for a physics candidate. For example, the reconstructed track hits in the muon system are connected to form a candidate muon track.

Finally, the last stage of the reconstruction process combines the object constructed in each subdetector to create objects based on the complete CMS detector. For example, a candidate muon track can be extrapolated into the Tracker system, improving the measurement of the track parameters and the candidate characterization. Other

example is the matching between the clusters of energy found in the ECAL and the HCAL and their combination into a jet candidate.

In the following sections, the algorithm employed for reconstructing the physical events used in this analysis is described. Pointing to the main elements of the reconstruction, the procedure to restore the physics object arising out of the collision from the detector information is illustrated.

4.2.1 Particle-flow event reconstruction

Each pp collision produces an extensive list of stable particles in an intricacy of final-states. The particle-flow (PF) event reconstruction algorithm [104] combines the information of CMS subdetectors to identify and reconstruct each individual particle in the events. The fine spatial granularity and hermetic assembly of the detection instruments are key pieces to make the PF reconstruction algorithm be suited to individually identify particles, even in very complex environments, such as inside a high-energy jet, in events with high number of pileup interactions, or with secondary interactions within the materials.

The PF algorithm is used as much for the online HLT reconstruction as the offline reconstruction at analysis level for real and simulated collisions. The global event description is accomplished by the complete characterization of charged and neutral hadrons, leptons (electron and muons) and photons. In a higher level of reconstruction more specialized algorithms make use of these objects to reconstruct jets of particles, calculate the missing transverse momentum of the event, provide a precise identification of electrons, muons and tau leptons, tag the jets, etc. In the same way, the collection of reconstructed particles can be used to identify pileup interaction events and enable the development of efficient pileup mitigation methods.

The PF reconstruction parts from the basic PF elements: reconstructed tracks of charged particles in the inner tracker, specifically the electron and muon tracks, and the reconstruction and calibration of calorimeter's clusters in the preshower, the ECAL and the HCAL. All of them connected by the "link" algorithm to create a particle candidate. All the information is available for the analyzers or the user dedicated high level algorithms.

4.2.2 Electrons and isolated photons reconstruction

The electron reconstruction procedure combines measurements from the ECAL and the Tracker [105]. Due to the high density of the Tracker material, most of electrons radiate a sizeable fraction of energy in form of bremsstrahlung photons, that can be converted into e^+e^- pairs which in turn emit more bremsstrahlung photons. For this reason, to achieve the best electron momentum reconstruction, the identification algorithm performance is based on the ability to gather all the radiated energy and propagate it to the corresponding linked energy cluster in the ECAL with low mismeasurement rate. The large probability for electrons to radiate in the Tracker material can be also exploited to disentangle electrons from charged hadrons.

Two approaches are carried out in order to determine the seed of electron's track candidate: the ECAL-based algorithm and the Tracker-based. Their results are then combined to overall increases the seeding efficiency in the whole Tracker acceptance.

The ECAL-based algorithm that makes use of the ECAL energy clusters and positions to extrapolate the electron trajectory towards the collision vertex, is highly efficient for well-isolated electrons. And the Tracker-based algorithm, developed in the context of the PF algorithm, that overall increases the seeding efficiency, specially for electrons missed by the ECAL-based approach. Those are mainly low p_T electrons whose tracks are significantly bent by the magnetic field and the radiated energy is spread over such an extended region and electrons in jets for which the energy and position of the associated cluster are often biased by the overlapping contributions from other particle deposits, leading to large inefficiencies.

Once the electron seed has been selected, it is used to initiate electron-track building, which is followed by the track fitting. The track building uses a combinatorial track finding algorithm based on the Kalman Filter method [106]. For each electron seed, it proceeds iteratively with the track parameters provided in each layer including, one-by-one, the information from each successive layer. Once the hits are collected, a Gaussian-sum filter (GSF) [107] fit is performed to estimate the track parameters.

The procedure of track building and fitting provides electron tracks that can be followed up the ECAL and thereby extract the track parameters at the surface of the ECAL.

Finally the electron candidates are reconstructed from the combination of GSF tracks and energy clusters in the ECAL. The electron momentum is very sensitive to the bremsstrahlung and photon conversion so that it is estimated from different electron observables and energy corrections.

The basic properties and techniques used for the tracking and the recognition of the energy deposition patterns of electrons and photons are similar. Isolated photons are therefore reconstructed together with the electrons, where photon candidates are seeded from an ECAL cluster if it is provided that they do not have link to a GSF track. The distinction between electrons and photons in the PF global event description can change for different specialized analyses. To deal with these specific event interpretations, the complete history of the electron and photon reconstruction is tracked and saved.

The reconstruction requires loose identification criteria so as to ensure high identification efficiency for genuine electrons with low rate of misidentification probability, further quality criteria are applied in CMS.

4.2.3 Muon reconstruction

The muons spectrometer allows muons to be identified with high efficiency over the pseudorapidity interval $|\eta| < 2.4$ without acceptance gaps. The capability of the calorimeter system to absorb other particles (except neutrinos) grants a high purity in the signal. Additional measurements from the inner Tracker provide a more precise momentum reconstruction of these muons. Further information can be derived out of muon energy deposits in ECAL, HCAL, and HO which associated to the muon track improves the muon identification performance.

The muon tracking is not specified in the PF algorithm. The high level muon objects are reconstructed from primitive muons in a multifaceted way [108]. The final collection incorporates three different types of muon candidates:

- *standalone muons*: the reconstructed hits within the DT or CSC detectors are clustered to form track segments, which are used as seeds for the pattern recognition in the muon spectrometer, to gather then all DT, CSC, and RPC reconstructed hits along the muon trajectory. The result of the final fitting is called a standalone-muon track.
- *global muons*: each standalone-muon track is extrapolated into the Tracker detector to match with a tracker track, referred to as an inner track. If the two tracks are compatible, the hits from the inner track and from the standalone-muon track are combined and fit to form a global-muon track. At large transverse momenta ($p_T \geq 200$ GeV) the global-muon fit improves the momentum resolution with respect to the tracker-only fit. See Figure 3.15 in Section 3.3.6.
- *tracker muon*: each inner track with $p_T > 0.5$ GeV and a total momentum $p > 2.5$ GeV is extrapolated to the muon system. If at least one muon segment matches the extrapolated track, the inner track qualifies as a tracker muon track.

The global muon reconstruction is designed to have high efficiency for muon penetrating more than only one detector plane. It typically requires track segments associated in at least two muon detector planes. The muons with momenta below 10 GeV fail more often this requirement due to the large multiple scattering before reaching the muons chambers. For these muons, the tracker muon reconstruction is more efficient because it only requires one track segment in the muon system.

About 99% of the muons produced within the geometrical acceptance of the muon system are reconstructed either as a global muon or a tracker muon and very often as both. Global muons and tracker muons that share the same inner track are merged into a single candidate. On the other hand, the muons reconstructed only based on standalone-muon tracks have worse momentum resolution and a higher admixture of cosmic muons than global and tracker muons.

The PF muon identification is conceived to retain prompt muons (from e.g. decays of W and Z bosons or quarkonia states), muons from heavy hadrons (from decays of beauty or charm hadrons), and muons from light hadrons (from decays in flight of π or κ mesons), with the highest possible efficiency minimizing the probability to misidentify a charged hadron as a muon. A charged hadron may be misidentified as a muon when, coming from the HCAL reaches the muon system, e.g. some of the hadron shower remnants (punch-through).

The isolated global muons are first selected by considering an isolation criterion, to reject sufficiently these misidentified muons. The sum of all inner tracks p_T and calorimeter energy deposits E_T within a distance $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.3$, is required not to exceed 10 % of the muon p_T . For muons inside jets or non isolated muons there are other stringent criteria.

In order to obtain the desired balance between efficiency and purity, different identification and isolation criteria can be applied. The final reconstructed objects classified as muons by the PF algorithm will be further used in physics analyses usually after a higher customized selection.

4.2.4 Hadrons and non-isolated photons reconstruction

Once the electrons, muons and isolated photons have been identified, the remaining information in the event concerns to the hadron reconstruction. These particles can be detected as charged hadrons (e.g. K^\pm , π^\pm , protons), neutral hadrons (e.g. neutrons) or non-isolated photons (e.g. from π^0 decays) from jet fragmentation and/or hadronization process.

The neutral hadrons and photon candidates are reconstructed when the ECAL and HCAL energy clusters do not link to any track in the Tracker System. The clusters in the HCAL are typically turned into neutral hadrons and the clusters in the ECAL are turned into photons. This association is justified by the observation that, within the Tracker acceptance ($|\eta| < 2.5$), about 25% of the energy in hadronic jets is deposited in the ECAL by photons, while the neutral hadrons leave only the 3%. Beyond the Tracker acceptance, the charged and neutral hadrons are indistinguishable leaving in total 25% of the jet energy in the ECAL. In this case, the ECAL clusters linked to a given HCAL cluster are assumed to arise from the same hadron shower and the particle is reconstructed as a (charged or neutral) hadron. Otherwise, when such a link does not exist, the ECAL cluster are classified as a photon.

To distinguish the energy deposits from neutral hadrons when overlapping with charged hadron deposits, the particle identification relies on an accurate calibration of the calorimeter system. If the calibrated calorimetric energy E_{CALO} results in excess with respect to the sum of all associated charged particle tracks' momenta by an amount larger than the expected calorimeter energy resolution for hadrons ($E_{CALO} > P_{tracks} + \sigma_{CALO}$), the energy is associated to a neutral hadron or a photon. Otherwise, if the calorimetric energy is compatible with the sum of all charged-particle's tracks momenta ($E_{CALO} \sim P_{tracks} + \sigma_{CALO}$), it is interpreted that no neutral particles have been produced.

Finally, it may happen that the calibrated calorimetric energy is smaller than the sum of all tracks' momentum ($E_{CALO} - \sigma_{CALO} < P_{tracks}$). When this difference is larger than few standard deviations, it can indicate the presence of muons not identified in the previous steps and whose small deposits of energy are contributing to the total calorimeter energy. Thus, the redundancy of measurements in Tracker and calorimeters allows few more muons to be found at this stage of reconstruction, without increasing the misidentified muon rate. Alternatively, if this difference is significantly larger, there is an indication of residual misreconstructed tracks. The algorithm solves this situation storing the tracks in decreasing order of p_T uncertainty to sequentially removing them from the total sum of track momenta until the compatibility of total calibrated energy is restored or a calorimeter energy excess appears, recovering thus one of the two former situations described above.

4.2.5 Lepton isolation

Lepton isolation is the main handle for selecting prompt muons and electrons produced in the electroweak decay of massive particles such as Z or W bosons and for rejecting the large number of nonprompt leptons produced in jets through the decay of heavy-flavour hadrons or the decay in flight of charged pions and kaons. The isolation is quantified by estimating the total p_T of the particles emitted around the direction of the lepton. The PF-based isolation relative to the lepton p_T^l is defined as:

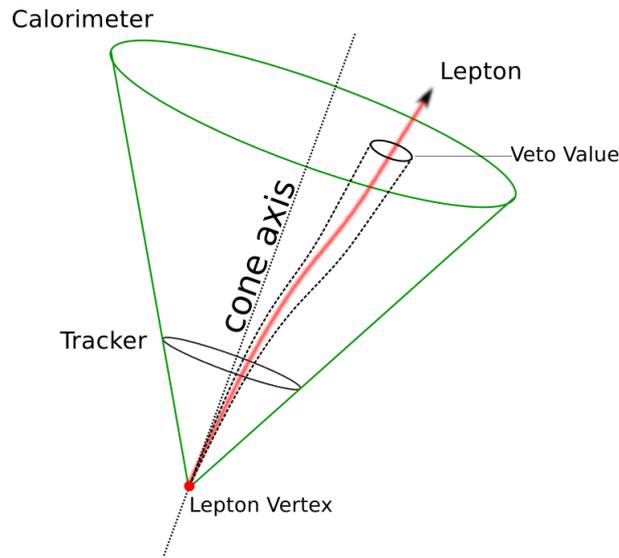


FIGURE 4.2: Schematic illustration of the lepton isolation cone. The lepton direction defines the cone axis. The lepton energy contribution is subtracted from the total energy sum by excluding a small area around the lepton (veto value). The threshold on the cone size determines the lepton isolation.

$$I_{PF_{lepton}} = \frac{1}{p_T^l} \left(\sum_{h^\pm} p_T^{h^\pm} + \sum_{\gamma} p_T^\gamma + \sum_{h^0} p_T^{h^0} \right), \quad (4.1)$$

where the sums run over the charged hadrons (h^\pm), the photons (γ) and the neutral hadrons (h^0) is computed for a distance ΔR to the lepton in general smaller than 0.3 or 0.4 in the (η, ϕ) .

An illustration of the isolation cone is represented in Figure 4.2. The geometrical construction of the cone starts on the lepton trajectory choosing its axis direction according to the lepton direction at this point. The lepton contribution to the energy inside the cone is usually subtracted by excluding a small area around the lepton (called veto value), in order to improve the discriminating power of the isolation algorithm.

The particle isolation can be also computed using a detector-based algorithm in a similar way but with less performance. As the analysis described in this thesis uses PF reconstructed particles, only PF-based isolation is going to be taking into account. From now on, along the whole text, the reference to PF will be omitted for simplicity.

Since the calculation of lepton isolation involves summing the p_T values of charged hadrons, photons, and neutral hadrons, lepton isolation is sensitive to pileup interactions, which give rise to additional reconstructed particles inside the isolation cone. To mitigate the deterioration of the isolation efficiency due to pileup energy contributions, the isolation as defined in Equation 4.1 needs to be modified. In similar way for electrons and muons, firstly, only charged hadrons associated with the hard-scatter vertex (PV) are considered. Secondly, the expected contributions from pileup are subtracted from the p_T sums of neutral hadrons and photons (see Section 4.2.6).

In the case of electrons, the standard 0.3 value uses in this analysis gives the minimal pileup dependence and reduced probability of other objects overlapping with the cone. But a further correction is needed to remove the effect of pileup in the isolation sum. It counts for an effective area of the isolation cone ($A_{eff}(\gamma, h^{0,\pm})$) and the average energy expected by particles from pileup (ρ):

$$I_e^{abs} = \sum_{h^\pm, HS} p_T^{h^\pm} + \max(0.0, \sum_{\gamma} p_T^\gamma + \sum_{h^0} p_T^{h^0} - (\rho \times A_{eff})). \quad (4.2)$$

In the case of muons, the expected contribution of photons and neutral hadrons from pileup is estimated from the scalar sum of the transverse momenta of charged hadrons in the cone that are identified as coming from pileup vertices, $\sum_{h^\pm, pileup} p_T$. This sum is multiplied by the factor $\Delta\beta = 0.5$ which corresponds approximately to the ratio of neutral particle to charged hadron production in inelastic proton-proton collisions, as estimated from simulation. Equation 4.1 becomes:

$$I_\mu^{abs} = \sum_{h^\pm, HS} p_T^{h^\pm} + \max(0, \sum_{\gamma} p_T^\gamma + \sum_{h^0} p_T^{h^0} - \Delta\beta \sum_{h^\pm, pileup} p_T^{h^\pm}). \quad (4.3)$$

In order to simplify the isolation requirement, the relative isolation is often used:

$$I_l^{rel} = \frac{I_l^{abs}}{p_T^l}. \quad (4.4)$$

The performance of isolation algorithms is studied for electrons and muons in 2016 data and Monte Carlo simulated sample of Drell-Yan + jets events using the tag-and-probe technique [88]. In Figure 4.3 is shown the efficiency to select signal prompt muons as a function of η , p_T and the number of primary vertices in the event. The isolation efficiency is computed for single prompt muons in a cone of radius $\Delta R < 0.4$, with $I_l^{rel} < 0.15$ applying the $\Delta\beta$ correction, and the tight muons identification selection. The observed efficiency has an excellent agreement with the expected efficiency. For muons with $p_T > 40$ GeV, the isolation efficiency is greater than 95% through the overall detector acceptance.

The performance of isolation algorithm depends on the accurate understanding of the involved variables. In Figures 4.4 to 4.6 the relative charged hadron isolation, electromagnetic hadron isolation and the relative neutral hadron isolation for 2016 data and $Z \rightarrow e^+e^-$ simulated events are shown. These variables have been reconstructed in a $\Delta R = 0.3$ cone around the electron with none isolation cut applied.

The expected distributions reproduce accurately the observations giving a reliable measurement of the hadron and γ energy contribution around the electron. The electron isolation involves these variables.

Even when the data is faithfully reproduced by the simulation, the manipulation of physics objects in the event may lead to differences in identification and isolation efficiency on data and simulation. This effect can be mitigated by the use of scale factors derived from the data and simulation ratio. In this case, the ratio of the efficiency in both data and Monte Carlo simulation is computed separately for identification and isolation algorithms.

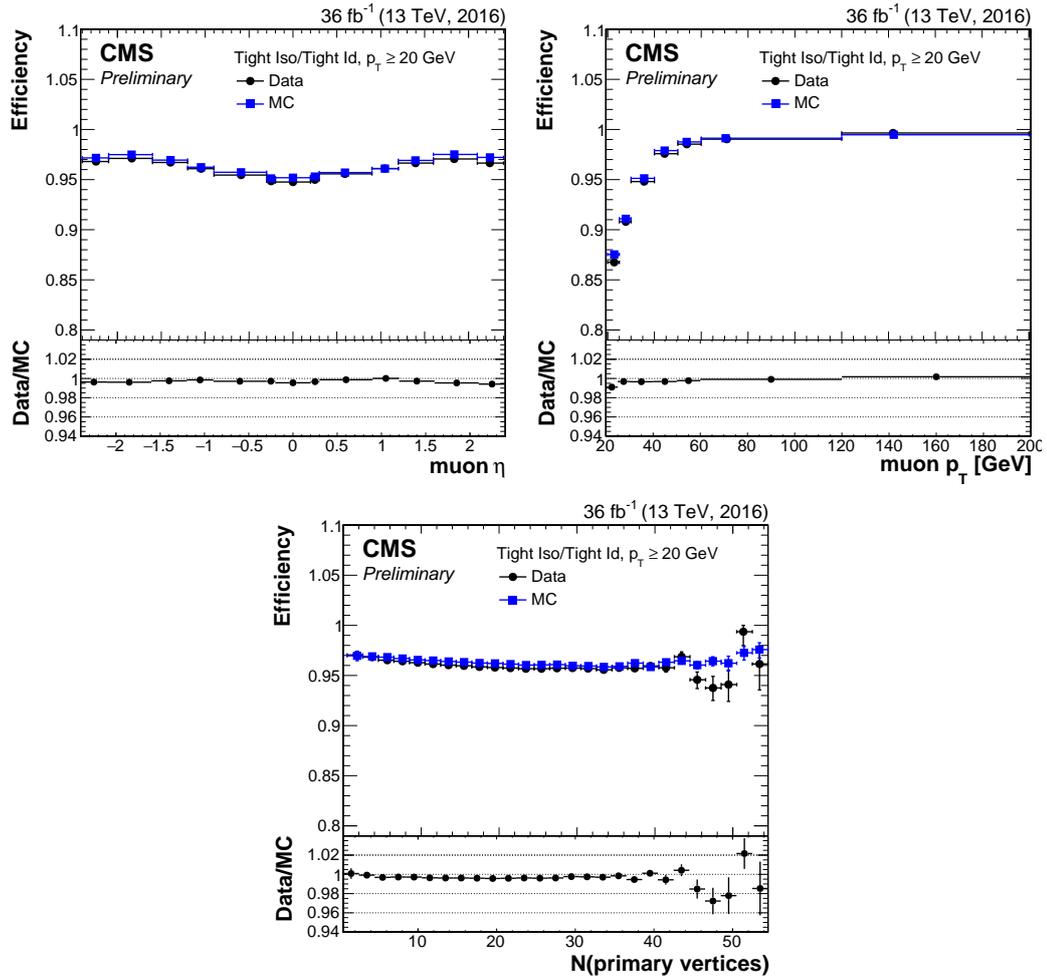


FIGURE 4.3: Tight muons isolation efficiency as a function of η (top left), p_T (top right) and the number of primary vertices in the event (bottom). The relative isolation cut ($I_1^{rel} < 0.15$) corresponds to the tightest selection criteria. The correction $\Delta\beta$ has been applied. Error bars include only statistical uncertainty [109].

4.2.6 Pileup interactions and primary vertex

During the data taking, under the 2016 LHC running conditions, an average number of 20 pp inelastic interaction per bunch crossing take place in the centre of CMS experiment (see Section 3.2.1). This mean number of interactions per bunch crossing is called pileup.

In the time of a bunch crossing, each hard collision originate a primary vertex. These interaction vertices are spread along the beam axis around the centre of CMS coordinate system, over a luminous region known as beam spot. The average number of pileup interactions μ in 2016 can be seen in Figure 4.7. The two plots shown here use the same data but with different values assumed for the minimum bias cross section.

Each primary vertex creates a number of new particles (quarks, leptons, hadrons, ...). These new particles can successively interact one each other, with the detector

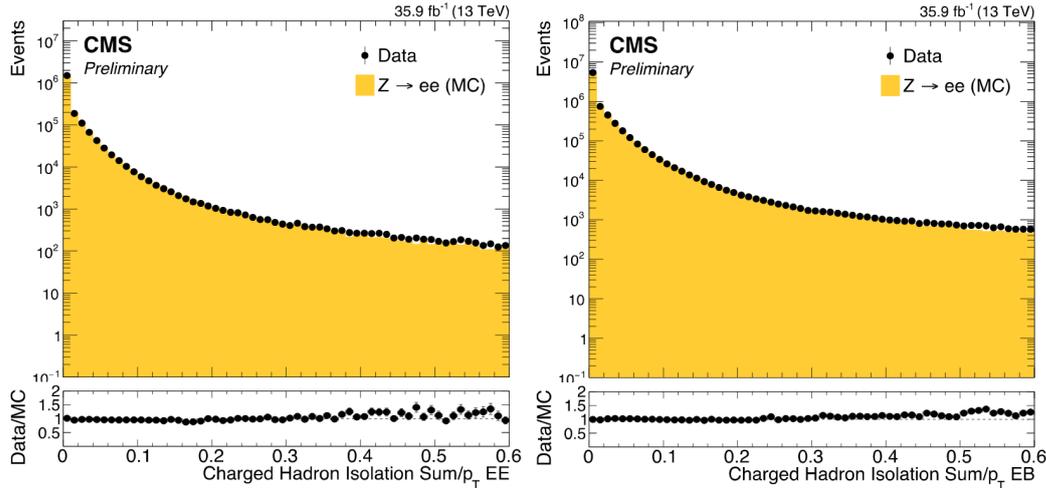


FIGURE 4.4: For electrons in the ECAL barrel (left) and ECAL endcaps (right), sum of transverse momenta of all the charged hadrons in a cone $\Delta R = 0.3$ around the electron divided by the transverse momentum (p_T) of such electron. No isolation cuts are included [110].

material or simply decay in flight producing new secondary particles whose trajectories start from a secondary vertex located at some displaced point with respect to the primary interaction vertices.

In the reconstructed event, only one primary vertex is expected. But, because of the number of simultaneous collisions, not only the vertex of the primary interaction is reconstructed, but also the ones from additional secondary collisions. Moreover, it may happen that the products from hard interactions of other "out-of-time" bunch crossing appears superimposed to the event contributing to the particle multiplicity.

The PF algorithm has been designed without taking pileup into account. As consequence, the extra contribution of particles produced in pileup interactions, such as charged or neutral hadrons and photons, results in an average additional p_T approximately constant in the (ϕ, η) plane that affect the jets, and p_T^{miss} reconstruction as well as lepton isolation and the identification of hadronic τ decays. The impact of this extra contribution can be mitigated using subtraction techniques:

- When charged hadrons are reconstructed within the Tracker acceptance, they can be identified as coming from a pileup vertex by associating their track with a reconstructed pileup vertex. If identify as coming from pileup they are subtracted from the list of the reconstructed particles used to form physics objects. This algorithm is called pileup charged-hadron subtraction (CHS).
- Photons and neutral hadrons, that are not reconstructed by the Tracker, cannot be associated to with one of the reconstructed vertices. In that case, to mitigate the impact of these particle, the uniformity of the p_T density of pileup interactions ρ in the (ϕ, η) plane allows that the average p_T contribution expected from pileup to be subtracted. That p_T density can be calculated with jet clustering techniques. Alternatively, this pileup contribution can be calculated as the ratio between the charged to neutral energy from pileup interaction in a local area, typically 0.5, for example around a lepton (see Section 4.2.5).

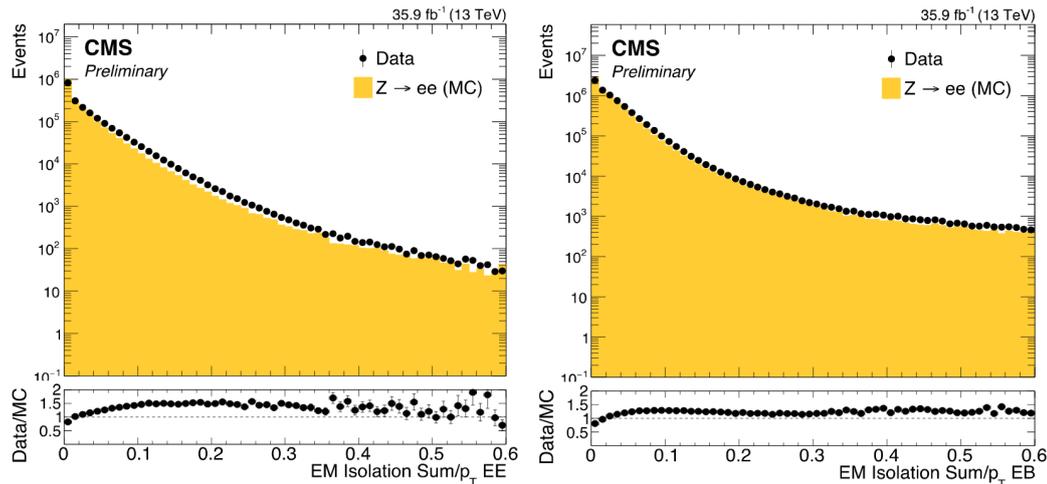


FIGURE 4.5: For electrons in the ECAL barrel (left) and ECAL endcaps (right), sum of transverse energies of all the neutral electromagnetic candidates in a cone $\Delta R = 0.3$ around the electron divided by the transverse momentum (p_T) of such electron. No isolation cuts are included [110].

The beam spot and vertex reconstruction is intimately connected with the particles identification since without an approximate estimation of the vertex position the particle tracking algorithms cannot begin. But, on the other side, the vertex finding algorithms are seeded with the tracks of the particle that come out from the vertex.

The primary vertex reconstruction starts at HLT processing using a quick finder algorithm based on the pixel Tracker information to locate each vertex position on the z -axis. The offline primary vertex reconstruction [111] proceeds in a similar way to the HLT reconstruction but with higher quality ending with best estimate of vertex parameters. The reconstructed vertex with the largest value of summed physics-object p_T^2 is taken to be the primary pp interaction vertex that addressed the trigger activation. The physics objects are the jets, clustered using a jet finding algorithm [112, 113] with the tracks assigned to the vertex as inputs, and the associated momentum imbalance in the transverse plane, taken as the negative vector p_T sum of those jets.

This secondary vertices are also reconstructed and taking into account for particle identification. An excellent impact parameter resolution is needed for a precise measurement of their positions with respect to the primary pp interaction vertices. In particular, in Section 5.1.4 will be shown that the presence of secondary vertices in an event is related in the heavy-flavour jets identification.

4.3 Monte Carlo simulation

The Monte Carlo (MC) method is based on repeated random (pseudo-random) sampling from a probability distribution to estimate the numerical solution of some stochastic or deterministic process. Wide range of problems in mathematics and physics can be approached using algorithms based on this technique. The MC algorithms are likewise used to develop MC simulations of probabilistic phenomena

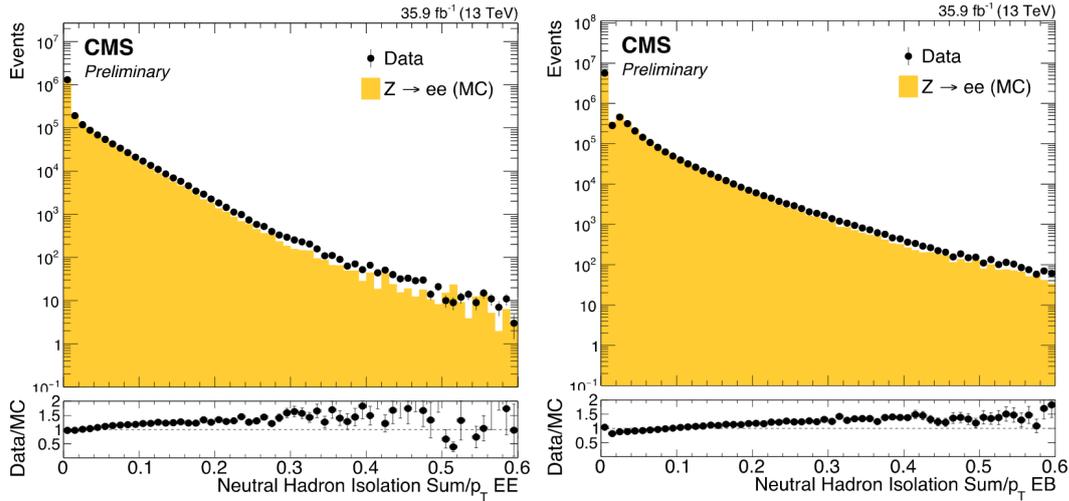


FIGURE 4.6: For electrons in the ECAL barrel (left) and ECAL endcaps (right), sum of transverse momenta of all the neutral hadron candidates in a cone $\Delta R = 0.3$ around the electron divided by the transverse momentum (p_T) of such electron. No isolation cuts are included [110].

arising in nature like for instance, the high-energy particle collisions in LHC. Although some authors show divergences on its birthday date, the method has been officially dated from 1949 when the paper of Metropolis and Ulam [114] was published for first time.

The full description of quantum physics processes occurring during and right after a hadron-hadron collision is significantly complex. In addition to the QCD interactions, the electroweak effects owing to the presence of charged particles will complicate the structure of the event. The computation of the full final states will necessitate to involve multi-particle calculations. This high-dimensional phase space leaves the MC integration the most suitable option to give a good approximation to the problem. In such a case, MC simulation is used as a faithful tool for generating a large number of simulated collision. The probability to produce a simulated event is kept approximately proportional to the probability of appearing this event in a real collision.

4.3.1 Monte Carlo event generator

The MC event simulation usually begins with a highly energetic collision between the constituent partons of the incoming protons and ends with a list of stable final-state particles. Each simulated event contains detailed predictions of the final state, specifying all the present particles and their momenta. The different stages of a hadron-hadron collision as simulated by a MC event generator are shown in Figure 4.8.

The generation process commences with some final X arising from the collision of two protons p_1 and p_2 at high energy. The process occurs at short distance where the

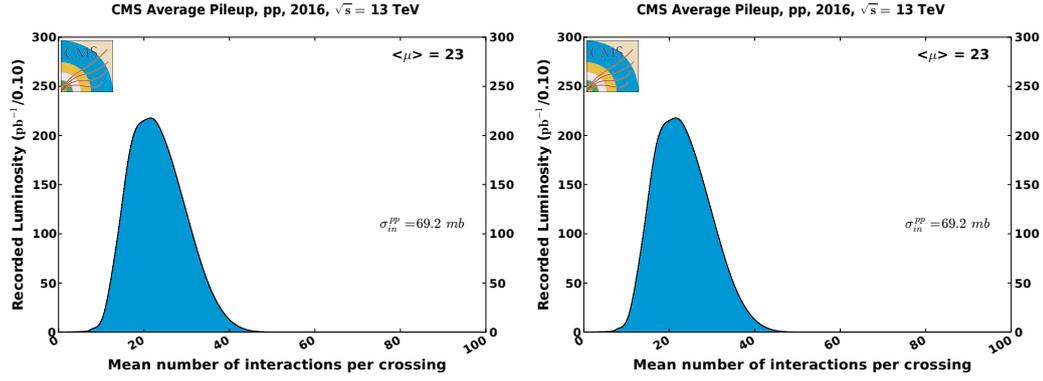


FIGURE 4.7: Mean number of interactions per bunch crossing during the 2016 data taking at $\sqrt{s}=13\text{TeV}$. On the left it is assumed the minimum bias cross section "CMS recommended" value, $\sigma_{pp}^{inel} = 69.2$ mb whereas on the right the "LHC standard" value of 80 mb is used. This value is calculated for comparisons with other LHC experiments [83].

QCD is weakly interacting and the MC calculations can be done upon the perturbation theory. The algorithm starts with the computation of the hard scattering cross section of the process at some order of perturbation theory can be written as [115]:

$$\sigma_{p_1 p_2 \rightarrow X} = \sum_{a,b \in \{q,g\}} \int dx_a dx_b f_{a,p_1}(x_a, \mu_F^2) f_{b,p_2}(x_b, \mu_F^2) \hat{\sigma}_{ab \rightarrow X}(x_a x_b s, \mu_R^2, \mu_F^2). \quad (4.5)$$

Here the sum runs over all possible partons participating in the hard interaction, with longitudinal momentum fractions x_a, x_b , that can give rise to a final state X at a centre-of-mass energy of $\sqrt{x_a x_b s}$, where s is the squared center-of-mass energy of the collision. The functions $f_a^{p_1}(x_a, \mu_F^2), f_b^{p_2}(x_b, \mu_F^2)$ are the PDFs of that partons in collinear factorization at the factorization scale μ_F . The short distance cross section for the scattering of partons of the type a and b is denoted as $\hat{\sigma}_{ab \rightarrow X}$.

In the framework of perturbative QCD, the predictions for observables are expressed in terms of the renormalized coupling $\alpha_s(\mu_R^2)$, a function of the (unphysical) renormalization scale μ_R usually named *running coupling*. In this case, the inclusive cross section of the pp collision, $\sigma_{p_1 p_2 \rightarrow X}$, can be expressed in perturbation series in the running coupling:

$$\sigma_{p_1 p_2 \rightarrow X} = \sum_{n=0}^{\infty} \alpha_s^n(\mu_R^2) \sum_{a,b \in \{q,g\}} \int dx_a dx_b f_{a,p_1}(x_a, \mu_F^2) f_{b,p_2}(x_b, \mu_F^2) \times \hat{\sigma}_{ab \rightarrow X}^n(x_a x_b s, \mu_R^2, \mu_F^2) + \mathcal{O}\left(\frac{\Lambda^2}{M_X^4}\right), \quad (4.6)$$

where the correction term $\mathcal{O}\left(\frac{\Lambda^2}{M_X^4}\right)$ accounts for contributions that are fundamentally non-perturbative. Traditionally, the fully inclusive hard cross sections has been calculated at the leading order (LO) of the series, $n=0$, but nowadays, due to the improvement of the computational tools, it is often to have corrections up to higher orders terms, normally next-to-leading-order (NLO) or next-to-next-to-leading-order (NNLO). Once the cross section for the initial hard process is calculate at some given

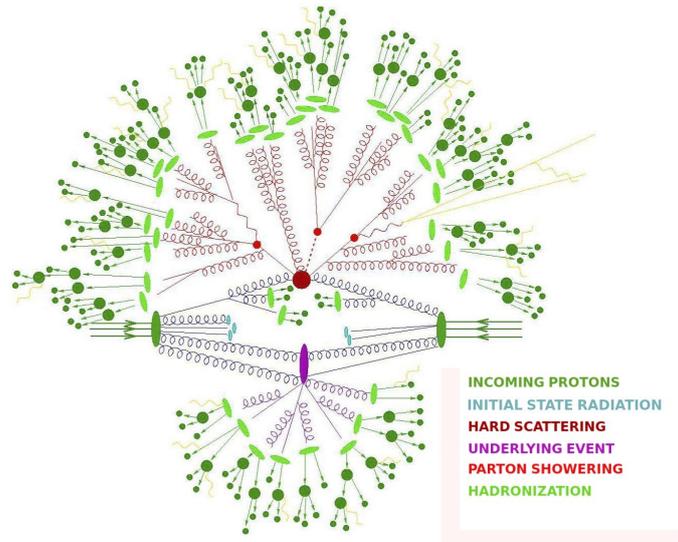


FIGURE 4.8: Schematic picture of a high energy hadron-hadron collision simulated by a MC event generator. Initial state parton showers are shown in light blue. The hard scattering producing the signal process is shown in dark red. Parton showering is presented in light red. Parton hadronization and hadron decays in green. Underlying events in purple. QED radiation in yellow.

order in perturbation theory, the higher-order real-emission terms of the perturbation series are a sequence of parton splittings understood as correction terms to the hard process. The dominant contributions are led by the collinear parton splitting and soft (low-energy) gluon emission. The algorithm develops an iterative process of parton splitting for each parton involved in the hard process called *parton shower generation*. Therefore, the initial cross section is partitioned into the cross sections of multitude of final states of arbitrary multiplicity whose sum is equal to the cross section of the primary process. This can be physically interpreted as high multiplicity of final state particles emerging alongside of the original process.

In the soft and collinear limits, due to the real or virtual emission of soft gluons, non-integrable divergences arise in the QCD matrix element. The singularities appear when the soft gluons become collinear (parallel) to another interacting parton and its momentum tends to zero. This divergent behavior can be present also for the massive quarks, but in this case, the quark mass acts as a cut-off on collinear singularities. This effect is smaller for massive quarks, such as, c, b or t than for lighter ones, thus less collinear activity is expected for heavy quarks. Which in turn is the reason why heavy quarks carry a larger fraction of the momentum acquired in the hard production process.

In the experimental measurements the tight constraints on emissions in the final state can veto a significant part of the inclusive integral, $\sigma_{p_1 p_2 \rightarrow X}$ over the soft and collinear divergences. This results in partial cancellation between real emission terms (subject to the constraint) and loop (virtual) contributions (not subject to the constraint), causing each order of α_s to be accompanied by a logarithmic coefficient $L = \ln\left(\frac{M_X}{p_T^2}\right)$. The terms of perturbative series in Equation 4.6 go as $\sim (\alpha_s L^2)^n$, where normally $\alpha_s L^2 \gg 1$ that makes the series converge very poorly if at all. In that case a process called *resummation* is carried out to account for the dominant logarithmically

enhanced terms to all orders in α_s .

The prediction of shower MC algorithms are independent of the specific observable under study and normally performs much better than strict leading logarithmic resummation. But it may not perform well for all the observables of the process, instead, it is common to specify what kind of corrections are include in the simulated sample.

At longer distances, the infrared cutoff of the parton shower leads to the hadronization process. In this context, the hadronization is the mechanism by which a set of partons from the shower is transformed into a set of primary hadrons. The transition take places, by construction, at the hadronization scale Λ_{QCD} (~ 1 GeV), where the perturbative regime gives way to the non-perturbative one. In that situation, the dynamics is not resolved by first principles, instead of that, QCD-inspired models such a *string model* or *cluster model* are used to describe QCD behaviour.

The new particles are found predominantly in the vicinity of the original ones creating clusters of radiation called QCD jets. If there are hadrons inside, these jets are preserved and will be identified further in physics analysis.

Besides of the main QCD process, additional activity may be present in the event. All the interactions that do not come from a processes ascribed to the hard interaction are included in the concept of the underlying event. It is associated to the initial/final state radiation (ISR/FSR) activity.

Before entering the hard-scattering process, the incoming particles may undergo collinear radiation generating an initial state shower. The hard process recoils and the event gains extra transverse momentum conferring extra boost to the particles produced in the interaction. It is usual that the partons emitted in ISR showers initiate secondary showers that behaves as the final state radiation (FSR) showers, the ones developed from an outgoing parton of the hard subprocess. These processes are very interesting in physics analyses that look for particles with high momentum, whose decay products overlap in a narrow region of space (boosted signatures).

The presence of ISR jets is widely used in searches for new physics targeting compressed scenarios where the invisible particles are soft leading to a relatively small p_T^{miss} signature. If an ISR jet is required in the event, the products will be boosted, and the soft invisible particles will acquire a higher momentum, leading, potentially, to a large p_T^{miss} signature. Since part of this analysis focuses on compressed regions, the use of ISR jets is considered to increase the sensitivity.

Additional contribution from QED radiation dresses the whole process. The non-hadronic resonance decays or bremsstrahlung photon emissions from charged particles are included in the MC event generation. Corrections of the order of $\alpha_{em} \ln(Q/m)$, where Q is the transferred momentum and m is the mass of the radiating particle, are taken into account.

4.3.2 The Monte Carlo simulation in CMS

Monte Carlo (MC) simulated events are used for different purposes in physics analyses. It is crucial, that the simulation fully reproduces the data collected by the detector. Besides of the hard pp interaction, the MC data samples include also the conditions of the detection machine, aiming to make them be comparable to the real

data samples. They include, a complete simulation of the pp hard scattering and the detector response. The generation of high-energy-physics events is centrally done by the CMS collaboration in three main steps:

- Generation of events at parton level: the particles coming from a physics process are generated by Matrix Elements (ME) calculators such as, POWHEG [116–118], MadGraph [119], ALPGEN [120], Sherpa [121], etc.
- Fully hadronized event: the generated events in the previous step are then run through a general-purpose generator which includes theory models for a number of physical aspects such as, the hard (hadronization) and soft interactions (underlying events), the parton distribution functions, parton showering, multiple interaction (pileup interactions), fragmentation and decays. Some the general-purpose generators used in CMS are PYTHIA [122, 123] and Herwig [124, 125].
- Detector simulation and digitization:
 - Full simulation (FullSim): the full simulation relies on GEANT4-based detector simulation [126]. It provides the modelling of CMS detectors geometry and behaviour (hits in the sensitive devices, configuration, calibration,...) including the field map from the 3.8 T solenoid, and reproduce the detector electronics response.
 - Fast simulation (FastSim): a parametric approach to simulate and reconstruct events in CMS detector. The concept of FastSim is to reduce the cpu time while still benefiting from an accurate simulation of the detector effects.

Once an event has been generated, it is processed with the same software and conditions as experimental data. The tier structure before the high-level event reconstruction is summarized as follows:

- GEN: generated Monte Carlo events.
- SIM: energy depositions of simulated particles in the detector.
- DIGI: SIM hits converted into detector response. Basically the same as the RAW output of the detector.

The full simulation is extensively validated by detailed comparisons with dedicated datasets. The derived corrections and the extracted uncertainties are taken into account in the subsequent physics analyses.

During its operation phase at high luminosity ($\mathcal{L} = 10^{34} \text{cm}^{-2}\text{s}^{-1}$), the LHC accelerator is producing an average of about 20 inelastic (hard-core) pp collisions per bunch crossing which will "pileup" on top of the signal collision firing the trigger.

Even more, in addition to the in-time pileup, it is necessary to account in the simulation for out-of-time pileup coming from bunch crossings before and after the triggered event. The number of crossings to consider before and after the nominal one depends on the front-end time response of the different subdetectors. Special cases such as bunch crossings with no pileup either before or after the nominal one are also considered for the simulation.

The luminosity evolution, or the number of hard interactions per bunch crossing, during a run cannot be precisely simulated a priori. Hence, when the MC file is

produced, the generated pileup profile is an approximate approach to the observed pileup distribution in data. Later in physics analysis, the MC events are reweighted to match the distribution of true interactions. The pileup in data is obtained per LS, so this value is a measurement of the averaged pileup during a single LS (see Section 3.3.8). The distribution of pileup for individual data events will therefore be a Poisson distribution around this average.

In MC samples, the pileup collisions are simulated separately from signal collisions. Both outputs are merged in a second step, using a luminosity dependent pileup contribution. This information is saved in a special collection for MC events, both in Fast and Full Simulation, for further uses. In the next chapter a summary of the MC samples of this analysis is detailed.

Chapter 5

Analysis physics objects

The previous chapters targeted the theoretical motivation for these searches, together with the experimental and computational environments, which allow to collect and reconstruct the data from the pp collisions at LHC. In this chapter a closer approach to the specific elements used in this data analysis is given. Section 5.1 covers the definition and characteristics of the physics objects. Section 5.2 is related to the online event selection. And, Section 5.3 describes the samples of simulated events used to study the different background and signal contributions to the data.

5.1 Physics objects identification

The dilepton final states of the top squark and chargino pair productions are characterized by the presence of two high- p_T isolated lepton (electron, muon) associated with missing transverse momentum, p_T^{miss} , and in the case of two top squark also two b-flavour jets. The physics object reconstruction introduced in Chapter 4 is now completed with the details of the identification and selection criteria for this analysis.

5.1.1 Electron identification

The electron candidates, after the reconstruction process described in Section 4.2.2, on the basis of are selected on basis a set of discriminating variables. The electron physics object is thus tightened in order to identify prompt isolated electrons, coming from a Z or W boson decay, and to separate them from the nonprompt electrons originating by background sources. The photon conversions and semileptonic decays of b and c quarks, or those charged hadrons, that interacting mostly in the ECAL, may be misidentified as electrons constitute the main cause of nonprompt leptons in the event.

Different strategies are carried on by CMS collaboration using simple and robust selection criteria based on HCAL, ECAL and Tracker measurements. The main observable discriminators used for the electron identification are described in Table 5.1. These can be grouped in three main categories:

- Calorimeter observables: used to separate genuine electrons (signal or electron from photon conversion) from misidentified electrons (jets with large electromagnetic contributions), exploiting the fact that the electromagnetic shower are narrower than hadronic showers, and the transverse shape of showers in

the ECAL. Also are utilized the energy fraction deposited in HCAL and the preshower in the endcaps.

- Tracking observables: employed mainly to improve the separation between the electrons and charged hadrons, exploit the GSF-fitted track information.
- Comparison observables: the ECAL and Tracker measurements are compared through the track-cluster matching information.

$\sigma(\eta, \eta)$	Shower width along η direction
$\Delta\eta$	Distance in the η plane between the track and the energy supercluster
$\Delta\phi$	Distance in the ϕ plane between the track and the energy supercluster
$E_{HCAL}/E_{ECAL} \equiv H/E$	Ratio between the energy deposited in the HCAL and the ECAL
$1/E_{SC} - 1/p_{trk}$	The difference between the inverse of the supercluster energy and the track momentum at the closest point to the primary vertex

TABLE 5.1: Observable ECAL and tracker variables used for electron identification.

Unlike the prompt electrons whose trajectories start from the beam-line, the secondary electrons, produced in photon conversion inside the volume of the Tracker, often leave tracks with missing hits in the innermost layers. To reject this background, CMS algorithms exploit the pattern of track hits identifying electron candidates from photon conversions. In the same way, the impact parameter (IP) of the electron's track, illustrated in Figure 5.1, can be also used to reject this source of background. Expressed in terms of the transverse (d_0) and longitudinal (d_z) distance to the vertex at the point of closest approach in transverse plane, or the three dimensional impact parameter ($IP(3D)$) and its significance ($S_{IP(3D)} = \frac{IP(3D)}{\sigma_{IP(3D)}}$), the IP variables are required to have small values assuring that the selected candidate electron tracks are those closest to the hard interaction PV.

Because of some electrons from photon conversion may still evade these cuts, they are vetoed by checking the matching between a given prompt electron candidate to at least one conversion candidate which also passes the selection cuts.

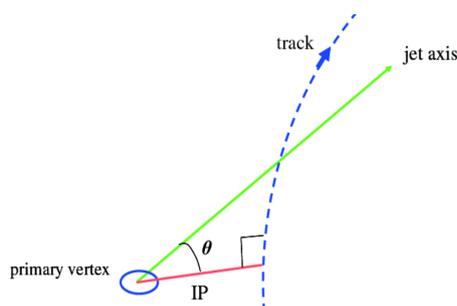


FIGURE 5.1: Illustration of a generic impact parameter track left by a particle produced inside LHC.

The other significant background to isolated primary electrons is due to genuine electrons within a jet originating from semileptonic decays of b or c quarks or to misidentifying of jets. In both cases, the electron has a significant amount of energy around its trajectory and the isolation requirement, as detailed in Section 4.2.5,

reduces these background sources. The isolation requirement is imposed independently on the identification algorithms, and the interplay between them tend to be analysis-dependent.

The CMS algorithms for electron identification have two approaches. A first one, that applies sequential requirements on the variables with a "cut-based" selection. And the second one, that combines the different variables in a multivariate (MVA) analysis [127] using more complex techniques, such a Boosted Tree Decision (BDT), to achieve better discrimination between signal and background.

The MVA-based electron identification is mainly used in analyses which require high efficiency down to low electron p_T with a good background rejection, for instance Higgs boson searches in leptonic final states [128]. This thesis will make use of high p_T electrons selected by the cut-based identification algorithms with the most tight cuts, since the purity of selected electrons can be relaxed in favour of a higher selection efficiency.

The cut-based approach has four basic working points (WP). These are defined in terms of the average identification efficiency against background rejection. The tight WP, with a standard efficiency of about 70%, is used for measurements where the background is important and the probability of electron misidentification is high. The medium and loose WPs, with an median efficiency of about 80% and 90% respectively, are generally used when the background rates are low. And the very loose with the average efficiency of 95% is usually reserved for third lepton veto requirement. In the MVA approach only two WPs of 80% and 90% of signal efficiency are defined.

Figure 5.2 presents the identification efficiency of the two methods measured by CMS collaboration in data collected during 2016 with a single electron trigger using the tag-and-probe technique with prompt electrons from $Z \rightarrow e^+e^-$ events [110]. It is shown in four η ranges as a function of the electron p_T . The "cut-based identification" efficiency corresponds to the tightest criteria selection and "MVA-based" identification efficiency corresponds to an average of 80% of MVA discriminant efficiency. The efficiency increases with the p_T increment to the 90% for the cut-based identification with very good agreement between data and simulation along the distribution getting better at high p_T values for the lower ranges of η which correspond to the central part of the detector. The electron isolation is described in Section 4.2.5.

The selection requirements for the tight and very loose WPs used in this analysis are summarized in Table 5.2. The threshold of the discriminant variables are derived as a function of the η region of reconstruction and depends on the running conditions. In addition, further dedicated selection to optimize the purity and the efficiency of the target signal has been done.

Electron selection

The final selection criteria used for electron candidates of this analysis are:

- $p_T > 20$ GeV and $|\eta| < 2.4$
- veto of the transition region in the ECAL: $1.4442 < |\eta| < 1.5660$
- cut-based tight ID

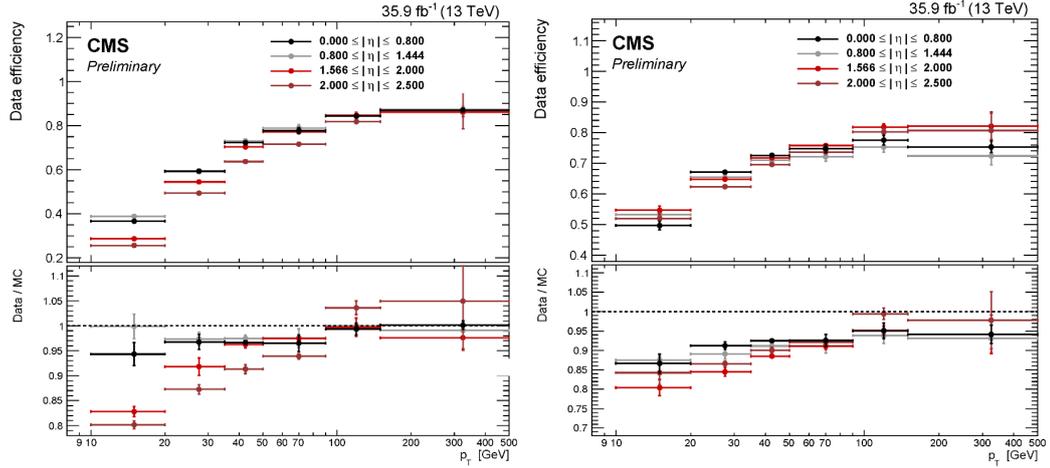


FIGURE 5.2: Electron cut-based identification with the tight selection criteria (left) and electron MVA-based identification efficiency with an average of 80% of MVA discriminator efficiency (right), for data events collected in 2016 using single electron trigger. Comparison with expected efficiency in simulated $Z \rightarrow e^+e^-$ events is shown in the low panel. The error bars include systematic and statistical uncertainties [110].

- $I_1^{rel} < 0.12$ (assuming $\Delta R < 0.3$ and effective area correction)
- n^o of lost hits in the track = 0 (rejecting the photon conversion)
- transverse impact parameter with respect to the beam spot $d_{xy} < 0.05$ cm and $d_z < 0.1$ cm, applied on the track from the inner tracker
- $|S_{IP(3D)}| < 4$

There is also defined a looser selection for vetoing events with a third electron passing these criteria:

- $p_T > 15$ GeV and $|\eta| < 2.4$
- veto of the transition region in the ECAL: $1.4442 < |\eta| < 1.5660$
- cut-based very loose ID
- $I_1^{rel} < 0.4$ (assuming $\Delta R < 0.3$ and effective area correction)
- $d_{xy} < 0.05$ cm and $d_z < 0.1$ cm
- $|S_{IP(3D)}| < 4$

The efficiency for the reconstruction, identification, isolation and selection of the electrons used in this analysis is shown in Figure 5.3. It is found to be 26–70% depending on their p_T and η .

5.1.2 Muon identification

The identification of the muons used in this analysis is based on PF reconstruction, as described in Section 4.2.3, with additional track-quality and muon-quality requirements. Specific criteria are designed to be highly efficient identifying prompt muons

	EB cuts ($ \eta_{SC} \leq 1.479$)		EE cuts ($ \eta_{SC} > 1.479$)	
	Tight	Veto	Tight	Veto
$\sigma(\eta, \eta) <$	0.0101	0.0114	0.0279	0.0352
$ \Delta\eta <$	0.00926	0.0152	0.00724	0.0113
$ \Delta\phi <$	0.0336	0.216	0.0918	0.237
$H/E <$	0.0597	0.181	0.0615	0.116
$ 1/E_{SC} - 1/p_{trk} <$	0.012	0.207	0.00999	0.174
photon conversion veto	yes	yes	yes	yes

TABLE 5.2: Selection requirements for tight and very loose (veto) electron identification in 2016 dataset used in this analysis. The electrons are split as a function of the ECAL barrel (EB, $|\eta_{SC}| \leq 1.479$) and the ECAL endcap (EE, $2.5 > |\eta_{SC}| > 1.479$) reconstruction.

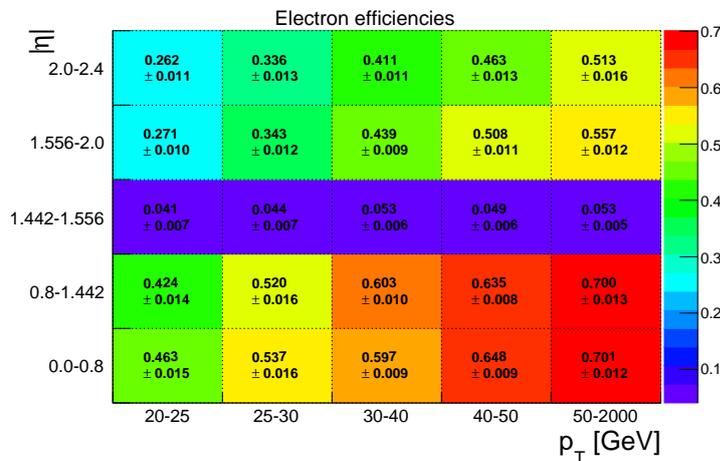


FIGURE 5.3: Efficiency of electron reconstruction, identification and selection (no veto selection) calculated for $t\bar{t}$ events with generator-level electron from W boson. The values are corrected by the data to SM simulation ratio.

from W and Z boson decays or heavy and light flavour quark decays, and to separate them from charged hadrons misidentified as muons.

The desired balance between identification efficiency and purity is achieved by applying a specialized selection based on various muon discriminatory variables. Depending on the type, number and threshold of these variables, three WPs are designed attending to this purpose. The Loose WP created to be highly efficient for prompt muons, as well as for muons from heavy and light quark decays, selects only either a global or PF tracker muon. The Medium WP includes the aforementioned track-quality and muon-quality requirements improving the background rejection with a small loss of selection efficiency. And the Tight WP with a lower efficiency only select global muons with a high purity criteria.

The compromise between the large background rejection and a high efficiency of selection makes the Medium WP as the most optimal choice for this analysis. The selection cuts of the variables used to define this WP are show in table 5.3. The muons are required to be reconstructed using the muon and tracker systems, i.e to be global muons, with extra requirements on the reconstructed track quality to better rejection of muons from mididentified punch-through hadrons, or at least satisfy the

tight criterion for the segment compatibility in muons stations. Further cuts on the IP are applied to avoid nonprompt muons from hadrons decays in flight, cosmic rays or pileup interactions. In addition an isolation requirement is introduced in this analysis to reject the non isolated muons produced inside jets.

	Medium
loose muon	yes
Fraction of valid tracker hits	> 0.8
Global muon	yes
Max.normalized χ^2 for global-muon track fit	< 3
Tracker-Standalone muon position matching	< 12
Kink finder in muon tracks	< 20
Segment compatibility in muon stations	> 0.303
Tight segment compatibility in muon stations	> 0.451

TABLE 5.3: Selection requirements for Medium WP muon identification in 2016 dataset. The cut value on the segment compatibility in muon stations will depend on whether the muon is global and satisfies the track requirements or not.

The performance of the identification procedure for the 2016 data taking period [109] has been probed in samples of prompt muons from Z boson decays using the tag-and-probe technique. The results are compared with Monte-Carlo predictions of a Drell-Yan + Jets sample in Figure 5.4. The data have been collected with a single muon trigger. The three different selection criteria in increasing order of purity have been tested.

The overall agreement between observation and simulation is excellent. The data confirm that prompt muons are identified by the PF algorithm with an efficiency close to 100%. For the *Medium* working point, the efficiency is more than 98% in the barrel and more than 95% in the most cut off part of the endcaps regions. The two points with lower efficiency in the middle of the figure correspond to the change of a muon chamber. In Section 4.2.5 the Figure 4.3 shows the isolation efficiency in 2016 dataset.

Muon selection

The final selection criteria for muon candidates used in this analysis are:

- $p_T > 20$ GeV and $|\eta| < 2.4$
- Medium ID
- $I_1^{rel} < 0.12$ (assuming $\Delta R < 0.3$ and $\Delta\beta$ correction)
- transverse impact parameter with respect to the beam spot $d_{xy} < 0.05$ cm and $d_z < 0.1$ cm, applied on the track from the inner tracker
- $|S_{IP(3D)}| < 4$

There is also defined a looser selection for vetoing events with a third muon passing these criteria:

- $p_T > 15$ GeV and $|\eta| < 2.4$

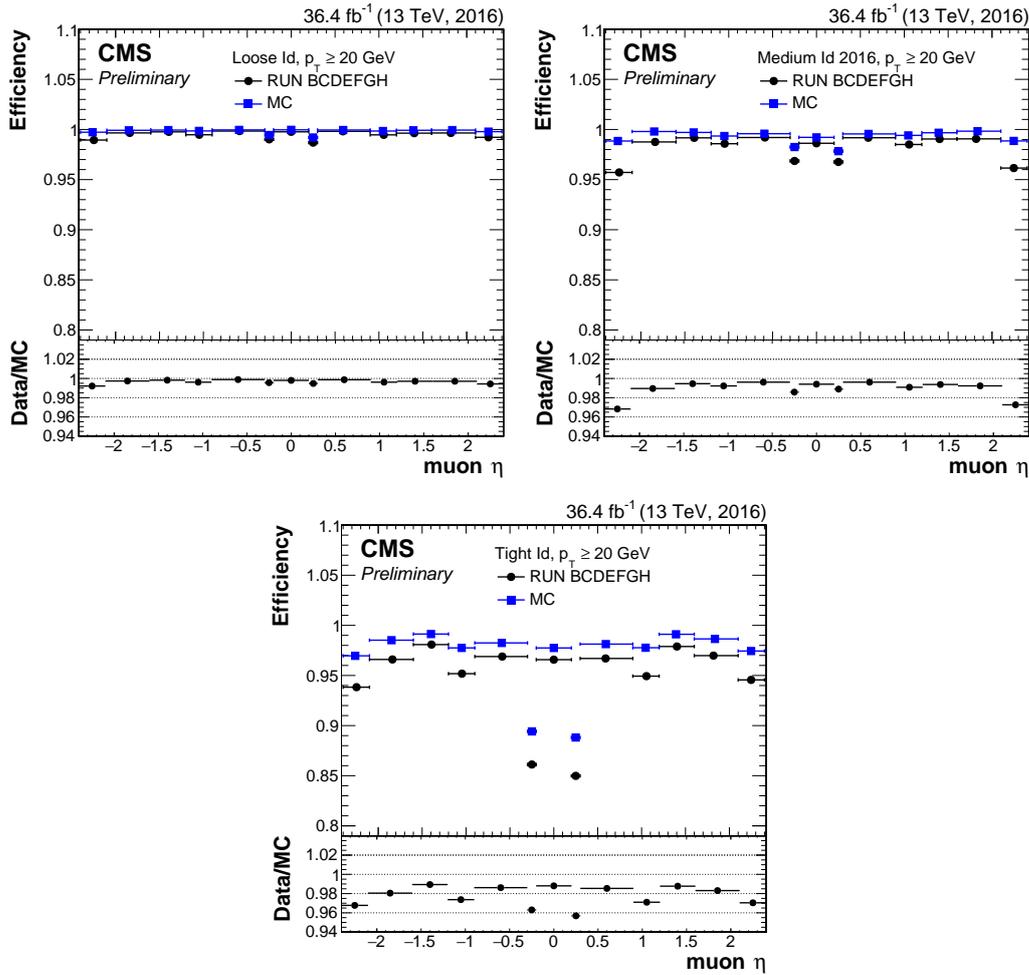


FIGURE 5.4: Muon identification efficiencies vs η for the full 2016 data in increasing order of tightened selection. Error bars include only statistical uncertainty [109].

- Loose ID
- $I_i^{rel} < 0.4$ (assuming $\Delta R < 0.3$ and $\Delta\beta$ correction)
- $d_{xy} < 0.05$ cm and $d_z < 0.1$ cm
- $|S_{IP(3D)}| < 4$

The efficiency for the reconstruction, identification, isolation and selection of the muons used in this analysis is shown in Figure 5.5. It is found to be 65–90% depending on their p_T and η .

5.1.3 Jet reconstruction

Jets are collimated streams of stable particles (lifetime $c\tau > 1$ cm) produced by the hadronization of an energetic quark or gluon arising from the hard scattering pp collisions. The hadronic jets of this analysis are reconstructed using the anti- k_T algorithm [112, 113], by clustering the four-momentum vectors of all the PF particles reconstructed in the event with a distance parameter of 0.4. This algorithm is infrared

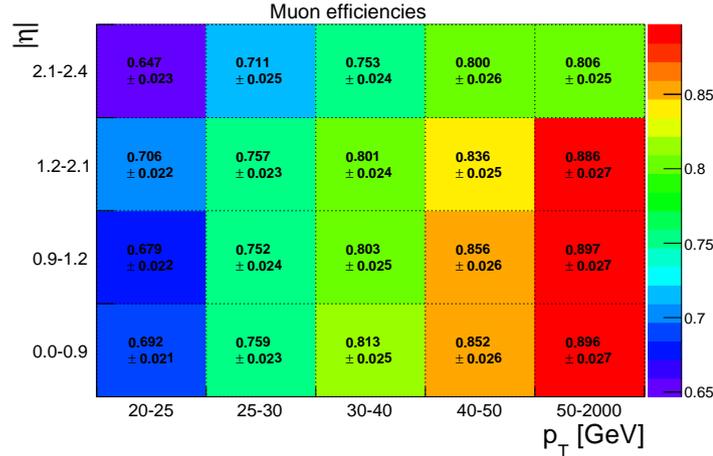


FIGURE 5.5: Efficiency of muon reconstruction, identification and selection (no veto selection) calculated for $t\bar{t}$ events with generator-level electron from W boson. The values are corrected by the data to SM simulation ratio.

and collinear safe, in the sense that the soft emissions from low energetic partons or collinear parton splitting inside the main cone do not affect the stability of the reconstruction and hence do not change the jet configuration.

Additional proton-proton interactions within the same or nearby bunch crossings, the pileup interactions, can contribute with additional tracks and calorimetric energy depositions to the jet momentum. In order to mitigate this effect, charged particles identified as originating at a pileup vertex are discarded during the jet clustering, and an offset correction is applied to the jet energy accounting for the remaining contributions.

The jet performance strongly depends on the particle reconstruction and the clustering algorithm. The PF jets benefit from the individual PF particle reconstruction, since dedicated calibration can be independently applied to each object, mitigating the misreconstruction, detector response and pileup effects in the jet energy response and jet energy resolution (JER).

The jet energy response is defined as the mean ratio of the reconstructed jet energy to the reference jet energy and it is measured in QCD MC simulation [129]. The PF jet response is softly dependent on the jet p_T and has a non-uniformity behaviour with detector acceptance, see Figure 5.6. As consequence, a set of jet energy corrections (JEC) are used to bring the jet energy response to unity, and remove the dependence on p_T and η variables.

The JEC accounts for pileup interaction effects, the detector response to hadrons, and residual differences between data and MC simulation jet energy scale (JES) as a function of the jet p_T and η . The jet energy corrections are derived from simulation to bring the measured response of jets to that of particle level jets on average. In situ measurements of the momentum balance in the dijet, multijet, photon + jet, and leptonically decaying Z + jet events are used to account for any residual difference in jet energy scale in data and simulation [129]. The uncertainties affecting the JES determination are taken as systematic uncertainties.

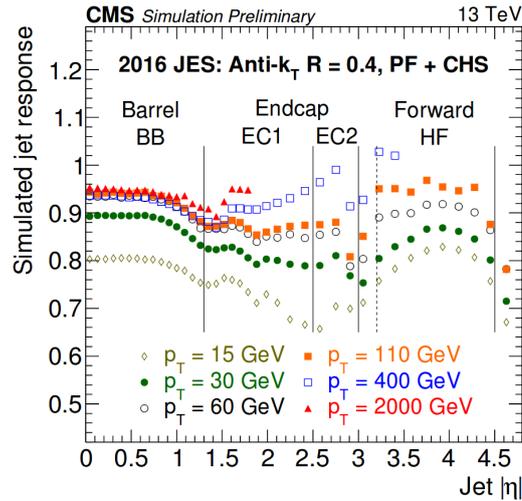


FIGURE 5.6: Simulated jet response as a function of jet η for different values of jet p_T . The jets have been reconstructed using the anti- k_T algorithm with a radius parameter $R=0.5$. The jets from charged hadrons associated to a pileup vertex has been subtracted (PF+CHS). The acceptance limit is shown for the Barrel, ECAL endcaps and Forward HCAL [130].

The CMS collaboration has adopted a factorized approach for the jet energy calibration [131] by applying the JEC in main three levels. These corrections are sequentially applied within a fixer order on the jet four momentum components. Each level correction covers essentially an specific effect:

- **Level 1** correction or the offset correction subtracts the energy not associated to the high- p_T scattering. The main sources of this energy are the pileup interactions and the electronic noise.
- **Level 2** correction or MC correction is used to correct for the p_T and η dependence of the jet response. This dependence is mainly caused by calorimeter nonlinearities, p_T thresholds and detector geometric effects.
- **Level 3** correction or the residual correction provides the final JEC factor. It takes into account the residual differences between the JES, i.e, the average response, in simulated events, and the JES measured in data. It is only applied to data.

Only corrected jets are used in this analysis. The jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be, on average, within 5 to 10% of the true momentum over the whole p_T spectrum and detector acceptance.

The PF jet composition studied in dijet events for fully corrected jets is shown as a comparison between 2016 data and multijet QCD simulation in Figure 5.7. The jets are reconstructed from all reconstructed particles, as PF jets, except charged hadrons associated with pileup vertices (PF+CHF jets) using the anti- k_T algorithm with a radius parameter $R = 0.4$. The improvements in tracking efficiency during the end of data taken (right) leads to better agreement between data and MC.

Additional selection criterion is applied to each jet in order to remove jets potentially dominated by anomalous contributions from various subdetector components or

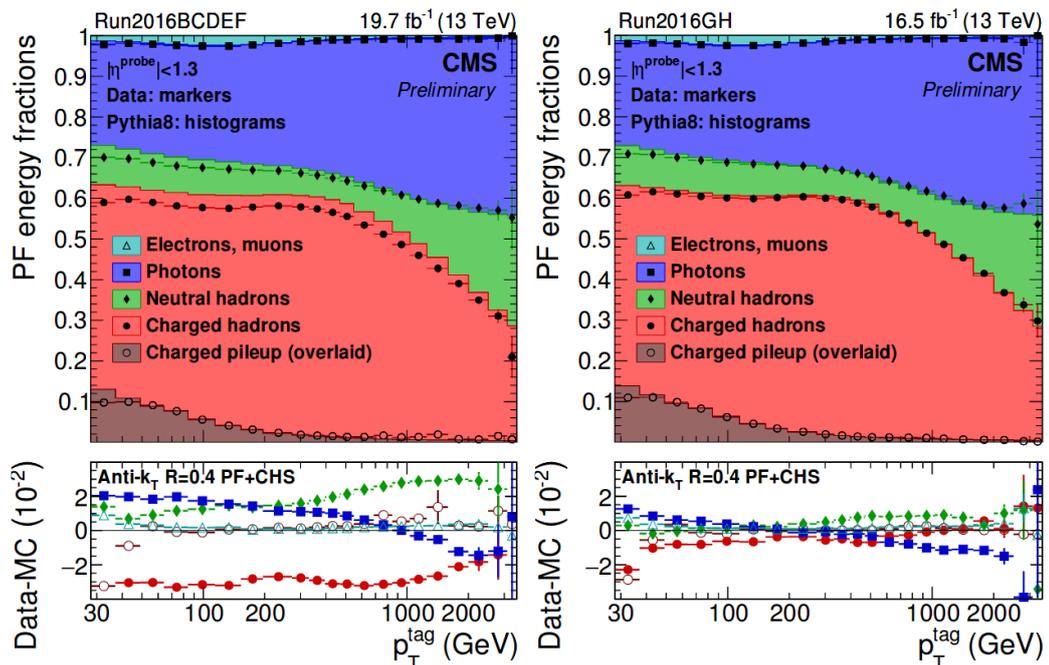


FIGURE 5.7: Jet energy composition as a function of jet p_T for simulated and observed events in different 2016 data taking periods: B to F (left), G to H (right). The simulated and measured PF energy fractions are stacked in the top panel, whereas the bottom panel shows the difference between observed and simulated events. The fraction of energy associated to charged hadrons linked to a pileup vertex (charged pileup) is removed before jet clustering, and it is only shown for visualization [130].

reconstruction failures. The reconstructed PF jets overlapping within a distance of radius $\Delta R < 0.4$ with the fully selected leptons (electron/muon) are excluded of the event. This procedure, known as jet lepton cleaning, is also applied in this analysis.

Finally, only jets with $p_T > 20$ GeV and $|\eta| < 2.4$ if they satisfy the loose identification criteria are selected. The jet identification criteria is a cut-based selection on the particles multiplicity and the composition on energy fractions.

In Figure 5.8 an event display during 2016 data taking is shown. The picture highlights the main properties of the physics objects composing an event registered by the CMS Experiment.

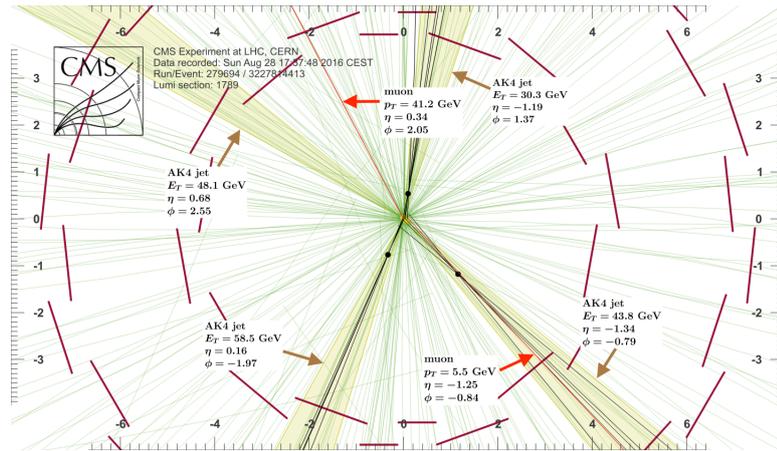


FIGURE 5.8: Recorded event (Rho-Z projection) with three jets, one isolated and one displaced muons in 2016 data collected at 13 TeV. The three jets have been reconstructed using the anti- k_T algorithm with a cone of $\Delta R = 0.4$. The jet with $p_T = 43.8$ GeV contains a muon. The p_T^{miss} of the event is characterized by a $p_T = 72.5$ GeV, $\phi = -0.32$ rad. The jet candidates for a b jet from top quark leptonic and hadronic decays are tagged by CSVv2T algorithm. One of the other two jets is tagged by CharmT algorithm. Tracks with $p_T = 0.5$ GeV are shown. The number of reconstructed primary vertices is 18. Reconstructed $m_t(W)$ is 101.8 GeV. Beam spot position correction is applied. Pixel detector is visualized. Reconstructed primary vertices are shown in yellow colour, while reconstructed displaced vertices and associated tracks are presented in black color. Dimensions are given in cm. CMS-PHO-EVENTS-2017-006-30.

5.1.4 Identification of b-flavour jets

The identification of the jets from the hadronization of quarks and gluons is a fundamental tool for the analysis of data in HEP. Excluding the top quark that has a very short lifetime and decays before the hadronization can occur, the rest of the gluons and quarks produced in pp collisions develop parton showers and eventually hadronization giving rise to jets of collimated particles that are then observed in CMS detector.

Based on the characteristic features of these particles, the jet identification algorithms exploit their properties to classify or tag each jet as a function of the original parton flavour. In particular, the algorithms for heavy-flavour jet identification exploit the

properties of the hadrons present in the jet. The heavy-flavour hadrons originating in jets from the radiation and hadronization of b or c quarks carry special properties that allow to discriminate these jets (b jet or c jet) from those jets originating from light-flavour quarks or gluons hadronization (light-flavour jets). The identification of jets from the hadronization of bottom quarks (b tagged jets) or charm quarks (c tagged jets), known as heavy-flavour jets, is essential both for the study of the standard model processes and the searches for new physics.

Algorithms for heavy-flavour jets identification

The hadrons containing a b quark (B hadrons) have a lifetime of the order of $\tau \approx 1.5$ ps, while the lifetime of hadrons containing a c quark (C hadrons) is about $\tau \leq 1$ ps. Depending on their momentum, the typical displacements from the PV can go from few mm to one cm for the B hadrons. This long lifetime of heavy-flavour hadrons in b or c jets leads to displaced decays that result in displaced tracks of particles with large impact parameter and originated in a secondary vertex (SV). In Figure 5.9 an illustration of a heavy-flavour jet with secondary vertex from a presumed B or C hadron decay is shown. Additionally, a recorded event by CMS with three jets, two SVs and displaced muon and electron tracks is shown in Figure 5.10. The two jets with the SV have been identified as b jets.

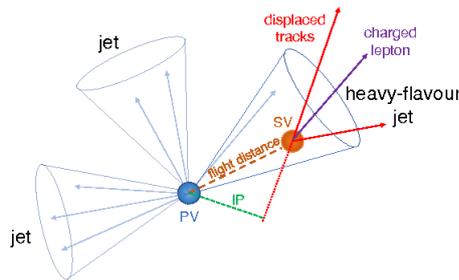


FIGURE 5.9: Illustration of a heavy-flavour jet with a B or C hadron displaced decay in a secondary vertex. The resulting charged-particle tracks presents a large impact parameter value with respect to the primary vertex.

In addition, the heavy-flavour quarks have larger mass and harder fragmentation than light quarks and massless gluons. As a consequence, the decay products of the heavy-flavour hadrons will have, on average, the p_T relative to the jet axis larger than the other jet constituents. The presence of soft electrons or muons in the heavy-flavour hadrons semi-leptonic decays is another important characteristic. Approximately in 20% (10%) of cases, a muon or an electron is present in the b (c) hadron decay. Hence, both the properties of secondary vertices and the large impact parameter of the displaced tracks, as well as the presence of charged leptons inside the jet or the p_T value of the jet constituents, are discriminant features for heavy flavours identification.

A variety of variables is constructed from such objects as tracks, primary and secondary vertices or isolated leptons. They constitute the base for the discriminating observables that then reflect the heavy-flavour jet properties. The identification algorithms exploit this information using one or more of these variables to provide a single discriminator value for each tagged jet. Depending on the threshold of this

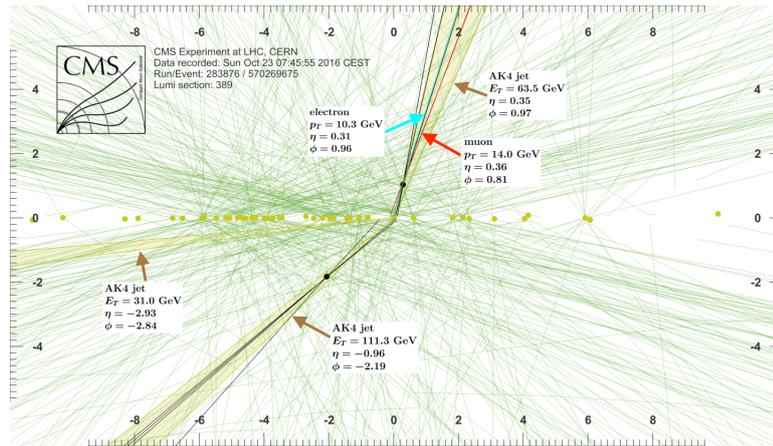


FIGURE 5.10: Recorded event (Rho-Z projection) with three jets, one displaced muon and one displaced electron tracks in 2016 data collected at 13 TeV. The three jets have been reconstructed using the anti- k_T algorithm with a cone of $\Delta R = 0.4$. The jet with $p_T = 63.5$ GeV contains muon and electron. Both jets have a reconstructed displaced vertex and are tagged by CSVv2T algorithm. Tracks with $p_T = 0.5$ GeV are shown. The number of reconstructed primary vertices is 44. Beam spot position correction is applied. Reconstructed primary vertices are shown in yellow colour, while reconstructed displaced vertices and associated tracks are presented in black color. Dimensions are given in cm. [CMS-PHO-EVENTS-2017-006-8](#).

value, different operating points with a certain misidentification probability for light jets can be defined.

In the CMS collaboration there are diverse techniques to develop sophisticated and robust tagging algorithms (taggers) that improves each time the identification efficiency for b and c tagged jets [132]. The newest tagging algorithms, based on machine learning procedures [133], have been implemented during the Run2 resulting in a better performance with respect to the simpler approaches of Run1. New developments and continuous updates are always ongoing due to the crucial importance that the b and c tagging has for the physics analysis. The most performant b taggers in the LHC Run1 and Run2 up to 2016 are summarized in the following items:

- The **Jet Probability (JP)**, used both in Run1 and Run2, sets its discriminating variable from the likelihood of all tracks (associated to the jet) to come from the primary vertex using the impact parameter significance values.
- The **Combined Secondary Vertex (CSV)**, used only in Run1, combines the SV properties such as the flight distance significance or the vertex mass with track-based lifetime information. Two likelihood ratios are built from these variables to discriminate against c and light- flavour jets.
- The **Combined Secondary Vertex version 2 (CSVvs2)**, used only in Run2, is based on the CSV algorithm of Run1. It combines larger number of variables using the secondary vertex and track-based lifetime information in a neural network (NN).
- The **combined multivariate algorithm version 2 (cMVAv2)**, based on cMVA

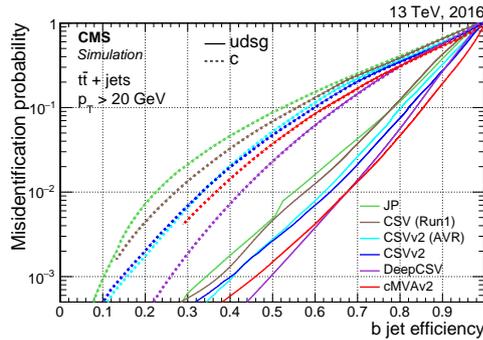


FIGURE 5.11: Misidentification probability for c and light-flavour jets versus b jet identification efficiency for various b -tagging algorithms of Run2. The Run 1 version of the CSV algorithm is also shown. These efficiencies are measured in $t\bar{t}$ simulated events [132].

algorithm of Run1, combines the information from six different b jet identification discriminators from JP, CSVv2 and other simple taggers with a Boosted Decision Tree (BDT). It is used only in Run2.

- The **DeepCSV** is an extension of CSVv2 algorithm from which it inherits the common features. It uses more charged particle tracks (up to 6) and combines all the variables with a deep neural network (DNN) instead of a simpler NN.

The heavy-flavour tagging efficiency is determined using simulated event samples from QCD processes and $t\bar{t}$ production. The efficiency to correctly identify a tag jet with a flavour f is defined as the number of jets with flavour f that pass the identification (tagging) requirements divided by the total number of jets with flavour f . In the same way, the misidentification probability is the probability for non- b jets to be misidentify as a b jet.

The misidentification probability for a c or light flavour jet versus the b jet identification efficiency at 13 TeV with the running condition of 2016 is shown in Figure 5.11. The jets are reconstructed from PF particles using the anti- k_T algorithm, with a distance parameter of 0.4. In simulated samples the jets are required to have $p_T \geq 20$ GeV and $|\eta| < 2.4$.

In general, the efficiency depends on the p_T , η and the number of pileup interactions in the event. In this figure the efficiency is integrated over the p_T and η distributions of the jets in the simulated sample. The DeepCSV discrimination against c and light-flavour jets outperforms all other algorithms for b -tagging efficiencies below 70%, while the cMVA2 tagger performs better against light jets for b -tagging efficiencies above 70%. Both taggers improve the CSVv2 performance by $\sim 4\%$ for a mistag rate for light jets of 1%.

Two variants of the CSVv2 algorithm exist according to the two secondary vertex reconstruction algorithms used by CMS, the adaptive vertex reconstruction (AVR) algorithm [134] and the inclusive vertex finding (IVF) algorithm [135]. For jets with $p_T > 20$ GeV in $t\bar{t}$ sample, the efficiency of secondary vertex reconstruction for b (udsg) jets using the IVF algorithm is about 75% (12%), compared to 65% (4%) reconstruction efficiency using the AVR algorithm. Since the beginning of LHC Run2 the IVF algorithm has become the standard algorithm adopted by CMS.

The absolute improvement in b jet identification efficiency of CSVv2 (AVR) with respect to the CSV algorithm of Run1 is about 2-4% for the same misidentification probability of light-flavour jets. Because the secondary vertex reconstruction based on the IVF algorithm is more efficient than the reconstruction using the AVR algorithm, an additional improvement of the order of 1-2% is appreciable in the CSVv2 algorithm when it uses IVF secondary vertices.

Three standard operating or working points have been defined for each b-tagging algorithm using jets with $p_T > 30$ GeV in a sample of simulated multijets with the average jet p_T of 75 GeV. These WP are categorized as loose ("L"), medium ("M") and tight ("T") corresponding to a misidentification probability for light jets of about 10%, 1% and 0.1%, respectively.

In this analysis, the jets originating from the hadronization of bottom quarks are identified by the CSVv2 algorithm using the medium operating point. This requirement provides an efficiency for identifying b jets that increases from 50 to 70% for jets with p_T from 20 to 100 GeV. The misidentification rate for jets originating from light quarks and gluons is about 1% in the same p_T range.

It is worth to mention that the τ leptons can decay into either charged leptons (e or μ) and neutrinos or few hadrons and one neutrino. The hadronic decays of τ s can be differentiated from quark and gluon jets by the multiplicity, the collimation and the isolation decay products.

5.1.5 Missing transverse momentum

In the lab frame, each pp collision is produced along the z-axis of the CMS coordinate system. Along the beam direction, as a function of the energy the longitudinal boost of interacting partons is significantly high, while in the perpendicular plane the boost is considered negligible. Assuming the law of momentum conservation, the measured transverse momentum before and after the hard collision should be the same quantity. However, when weakly interacting neutral particles are produced, the absence of signal through the detector layers cause a transverse momentum imbalance in the reconstructed event.

The transverse momentum imbalance is a physic observable that reflects the total transverse momentum of all those particles present in the event which do not leave any sign in the detector, such as neutrinos. It is known as missing transverse momentum (\vec{p}_T^{miss}) or missing transverse energy (\vec{E}_T^{miss}) and its magnitude is denoted as p_T^{miss} (E_T^{miss}). The \vec{p}_T^{miss} quantity is calculated as the negative sum of the transverse momenta of all final-state particles reconstructed in the detector.

The CMS Collaboration has developed several techniques to estimate the \vec{p}_T^{miss} in each event [136, 137]. The approach adopted in this analysis uses the PF reconstruction algorithm to compute the quantity \vec{p}_T^{miss} as the negative vectorial sum of all PF candidates transverse momentum in the event,

$$\vec{p}_T^{miss} = - \sum_{i=1}^{N_{PFcand}} \vec{p}_{T,i}. \quad (5.1)$$

These PF candidates correspond to well reconstructed physics objects, namely electrons, muons, photons, or jets, and the unclustered energy (E_U) of the event. The

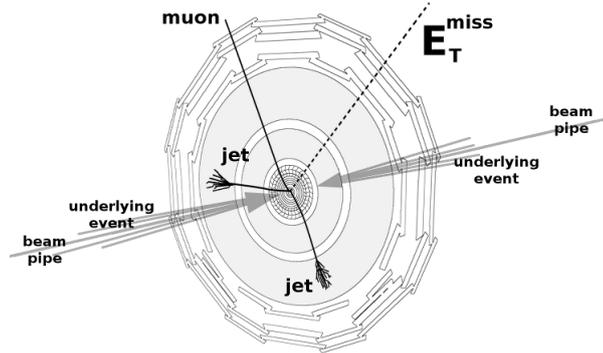


FIGURE 5.12: Transverse energy imbalance produced by a high energy collision at the center of a general detector similar to the CMS experiment.

last one, E_U , accounts for the contribution of all the PF candidates, usually with very low p_T , not associated with any of the previous physics objects. Thus, the accurate estimation of p_T^{miss} relies on the precise identification of the physics object and its uncertainty depends strongly on the topology of the event. A geometrical schema of a typical CMS event containing \vec{p}_T^{miss} is shown in Figure 5.12.

There is a variety of reasons that can affect the measurement of the genuine \vec{p}_T^{miss} overestimating or underestimating its magnitude. Among others, the sources include the nonlinearity response for hadrons and the minimum energy thresholds in the calorimeters, as well as p_T thresholds in the tracker and inefficiency in the track reconstruction. These effects can lead to important bias on the p_T^{miss} value when further physics analyses are carried out. However, they are usually reduced taking into account the energy corrections applied to jets 5.1.3. Hence, the Equation 5.1 becomes:

$$\vec{p}_{T,corr}^{miss} = \vec{p}_T^{miss} - \sum_{j=1}^{N_{PFjets}} (\vec{p}_{T,j}^{corr} - \vec{p}_{T,j}). \quad (5.2)$$

Despite the caution, some events with anomalous p_T^{miss} can turn up in the analysis dataset. In 2016 data spurious events with known p_T^{miss} problems were found and rejected by applying a set of filters algorithms developed during this run [93]. These events are affected by the electronic noise of calorimeters, the beam halo particles traveling trough the beam-pipe, or non-PF muons that are reconstructed as charged hadron candidate, contributing to a large anomalous p_T^{miss} .

One the other hand, the contribution from pileup interactions to the genuine \vec{p}_T^{miss} is close to zero, as the probability of producing weakly and neutral interacting particles like neutrinos in inelastic pp scattering interactions is small. However, the nonlinearity response and the energy thresholds can also produce on average, an extra momentum imbalance in the vectorial sum of observable particles momenta from pileup collisions. There are specific algorithms specially developed to mitigate effects from large number of pileup interactions.

This analysis is focused on high- p_T^{miss} values where the impact of pileup interactions on the p_T^{miss} values is not significant. The JEC propagated to the \vec{p}_T^{miss} together with

a minimum jet p_T threshold of 20 GeV reduces these contributions allowing a good accuracy in the estimation of \vec{p}_T^{miss} .

The p_T^{miss} performance (scale and resolution) in 2016 data is studied in events with an identified Z boson or an isolated photon [93]. The electron, muon and photon are reconstructed with resolutions lower than 6%, while the jets are reconstructed with a resolution between 5% and 15%. As a consequence of the lower precision of jet reconstruction, the p_T^{miss} resolution is dominated by the hadronic activity in the event. The p_T^{miss} performance studying the events with $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ is shown in Figure 5.13, and the performance on events with γ + jets is shown in Figure 5.14. There is an overall agreement between data and simulation within the range of uncertainty.

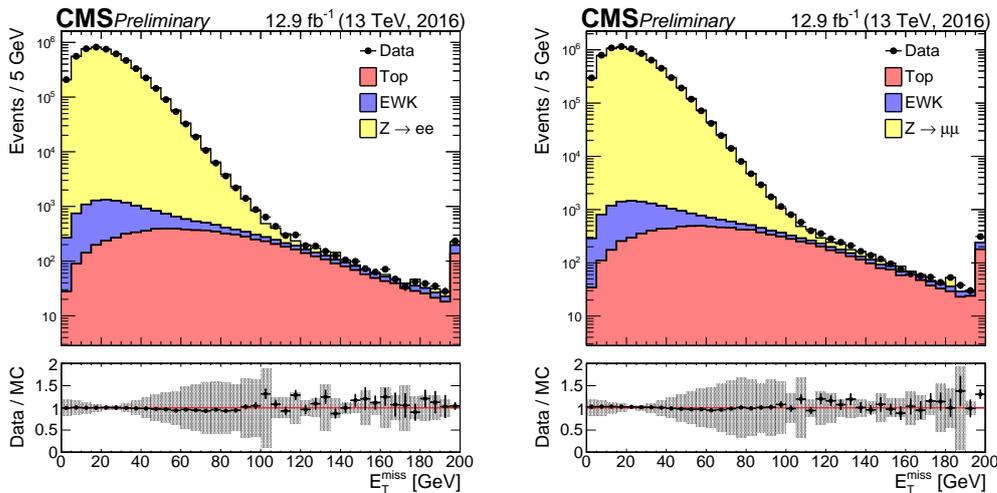


FIGURE 5.13: Distribution of p_T^{miss} in $Z \rightarrow e^+e^-$ (left) and $Z \rightarrow \mu^+\mu^-$ events. The points in the lower panel of each plot show the ratio between the data (black dots) and Monte Carlo simulation (solid colors), including the statistical uncertainties of both data and simulation. The systematic uncertainties due to the jet energy corrections and the unclustered energy are exhibit with the statistical uncertainty on the ratio bottom plot. The last bin includes all the overflow content [93].

5.2 Data samples

The data samples analyzed for this work correspond to an integrated luminosity of $35.9 \pm 0.9 \text{ fb}^{-1}$ [100] of pp collisions at $\sqrt{s} = 13 \text{ TeV}$ collected with the CMS detector during the 2016 data taking period at the LHC. The dataset is saved according to the run period in the course of 2016. Each run period is composed in turn by the collision runs of stable proton's beams. It is not unusual that several technical problems arise during that runs making necessary the commissioning and management of the quality of recorded data. Eventually, only the certified runs as good run are selected for physics analyses.

The collected data events used in this analysis are determined by the presence of one or two leptons and only certified runs have been used. A complete list of the data

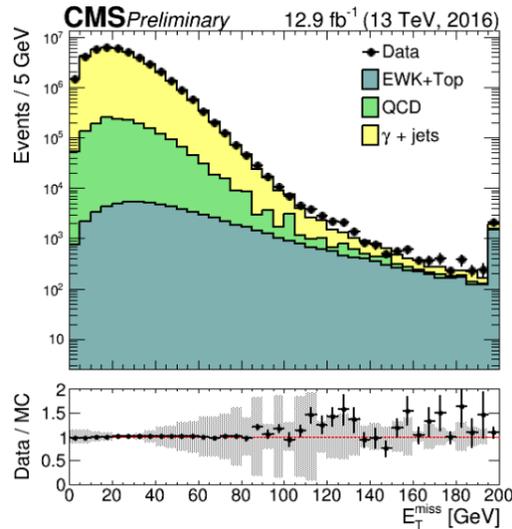


FIGURE 5.14: Distribution of p_T^{miss} in γ + jets events. The EWK contribution corresponds to the diboson, $Z\gamma$ and $W\gamma$ production processes. Same conditions as Figure 5.13 [93].

samples is gathered in Table A.1. This analysis has followed a blinding procedure in which signal regions have been defined and fixed before having a look at the data.

5.2.1 Triggers

At a crossing frequency of 25 ns, the average crossing rate of the two proton beams in LHC is over 30 millions of times per second, which means millions of pp collisions per second (see Section 3.2.1). This huge amount of data far exceeds the input rate of the on-line computer farm. In order to select the most interesting events produced in each collision, in the design of CMS has been included a highly customized trigger system (Section 3.3.7).

The events of interest for this analysis are selected using triggers that requires the presence of two well identified and isolated leptons (ee , $\mu\mu$, $e\mu$). The threshold on the transverse momentum of the leading lepton is 23 GeV for the ee and $e\mu$ triggers, and 17 GeV for the $\mu\mu$ triggers. The threshold for the trailing lepton is 8 (12) GeV for muons (electrons). To increase the efficiency of the trigger selection, the events accepted by triggers requiring at least one electron (muon) with $p_T > 25$ (24) GeV are also included if that leptons pass tighter identification criteria than the ones applied in the double-lepton triggers. Additional selection on the lepton longitudinal impact parameter (d_z) was added to some trigger paths in order to reduce the large rate of runs with high instantaneous luminosity, mostly at the end of the data taking period, avoiding then the prescale of the data collections.

A summary of the used HLT trigger paths can be found in Table 5.4. For each lepton flavour pair all the paths are combined with a logical OR. The trigger performances are measured with leptons from the $Z \rightarrow l^+l^-$ decays using the tag-and-probe technique. The combined efficiency of the dilepton and single-lepton triggers for signal events is found to range between 90% and 99%, depending on the p_T and η of the leptons.

Lepton pair	Primary dataset	HLT paths
$e\mu$	MuonEG	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v* HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v* HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v* HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*
	SingleMuon	HLT_IsoTkMu24_v* HLT_IsoMu24_v*
	SingleElectron	HLT_Ele27_WPTight_Gsf_v* HLT_Ele25_eta2p1_WPTight_Gsf_v*
ee	DoubleEG	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
	SingleElectron	HLT_Ele27_WPTight_Gsf_v* HLT_Ele25_eta2p1_WPTight_Gsf_v*
$\mu\mu$	DoubleMuon	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v* HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v* HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v* HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*
	SingleMuon	HLT_IsoTkMu24_v* HLT_IsoMu24_v*

TABLE 5.4: HLT triggers paths used in this analysis. For each lepton flavour pair the paths are combined using a logical OR. The leptons are required to satisfy a basic identification and isolation criteria. The phrase "v*" refers to all the available versions have been included.

5.3 Monte Carlo simulated samples

The samples of MC simulated events are used to study the contribution of SM processes to the selected dataset and the expected acceptance of the different signal models. A summary of all MC samples used in this analysis is listed in Table A.2 of Appendix A.

Events from top quark-antiquark pair ($t\bar{t}$) production are generated with POWHEG v2 [118, 138, 139] and normalized to the expected cross section calculated at NNLO in perturbative quantum chromodynamics (QCD), including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [140].

Events with a single top quark produced in association with a W boson (tW) are generated with POWHEG v1 [141] and normalized to an approximate NNLO cross section calculation [142]. Diboson production (WW , WZ , and ZZ) via quark-antiquark annihilation is simulated at NLO using POWHEG v2 [143, 144]. The yields of events from WW production are scaled to the NNLO cross section [145]. Events from $q\bar{q} \rightarrow ZZ$ production are reweighted via NNLO/NLO K factors as functions of the generated ZZ system mass [146]. Two additional sets of K factors, as functions of the generated ZZ system p_T and of the azimuthal separation ($\Delta\phi$, ϕ in radians) between the Z bosons, are used to evaluate the uncertainty in the kinematic properties of ZZ production. Diboson production via gluon fusion is simulated using MCFM v7 [147], and LO cross sections obtained from the generator are corrected with the NNLO/LO K factors [146, 148].

Drell-Yan events are generated with MADGRAPH5_aMC@NLOv2.2.2 [119] at LO, and event yields are scaled to the NNLO cross section [149]. Events from $t\bar{t}W$, $t\bar{t}Z$, triboson, and $H \rightarrow WW$ production are generated at NLO [150, 151] with the MADGRAPH5_aMC@NLO generator.

Chargino pair production and top squark pair production events are generated using MADGRAPH5_aMC@NLO at LO with up to two extra partons in the matrix element calculations, and are normalized to the respective cross sections computed at NLO plus next-to-leading logarithmic (NLL) precision [47, 152–159], with all the other sparticles assumed to be heavy and decoupled. In the case of chargino pair production, calculations are performed in a limit of mass-degenerate wino $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^\pm$, and light bino $\tilde{\chi}_1^0$.

All processes are generated using the NNPDF3.0 [160] parton distribution function (PDF) set. The parton showering, hadronization, and the underlying event are modeled using PYTHIA 8.212 [123] with the CUETP8M1 [161] underlying event tune for all the processes, except in the generation of $t\bar{t}$ events, where the first emission is done at the matrix element level with POWHEG v2 and the CUETP8M2T4 [162] tune is used. Weights for the estimation of theoretical systematic uncertainties, including those related to the choice of PDFs, and renormalization and factorization scales, are included in simulated events [163].

The detector response to the generated events is simulated using a realistic model of the CMS detector based on GEANT4 [126] for SM processes, while for signal events a fast simulation (FastSim) [164] of the detector based on a parametrization of the average response to particles is used. Simulated events are subsequently reconstructed using the same algorithms as applied to data.

In order to model the effect of multiple interactions per bunch crossing or pileup, the SM simulated events are mixed with minimum bias events simulated with PYTHIA, and are reweighted in order to match the rate of multiple interactions observed in data. A variation of $\pm 5\%$ on the minimum bias cross section is used to estimate the uncertainties due to pileup modeling. The pileup profile in data has been centrally provided for the corresponding runs by the physics validation group.

The modeling and normalization of the main background processes are studied in data, as discussed in Chapter 6. The modeling of $t\bar{t}$, tW , and WW production is studied in data control regions (CRs), and their normalization is determined via a maximum likelihood (ML) fit to data. The normalization of the yields of events from $t\bar{t}Z$, WZ , ZZ , and Drell-Yan production is corrected by the event rates measured in dedicated CRs.

To improve the modeling of jets from initial-state radiation (ISR) in simulated signal events, reweighting factors are applied, which make the distribution of observables for related SM processes in simulation to agree with control samples in data. For chargino pair production, mediated by the electroweak interaction, the reweighting procedure is based on studies of p_T balance in inclusive Z boson production events [165]. Events are then reweighted according to the total transverse momentum (p_T^{ISR}) of the system of supersymmetric particles. The reweighting factors range between 1.18 at $p_T^{ISR} \approx 125$ GeV and 0.78 for $p_T^{ISR} > 600$ GeV. A global reweighting is further applied in order not to alter the signal production cross section.

As the top squark pair production occurs via strong interactions, a different set of reweighting factors is derived as a function of the multiplicity of ISR jets (N_{jet}^{ISR}) in a sample of $t\bar{t}$ events selected by requiring an OC electron-muon pair and two jets identified as coming from bottom quark hadronization. The measured reweighting factors vary between 0.92 and 0.51 for N_{jet}^{ISR} between 1 and 6, with an additional scale factor applied to leave invariant the total yields of produced events.

Differences have been observed in the modeling of \vec{p}_T^{miss} resolution in events simulated with FastSim and with the full detector simulation. To account for this effect, the acceptance for signal events is computed both using \vec{p}_T^{miss} at generator level and after the event reconstruction. The average value of the two acceptances in each analysis bin is taken as the central value for the acceptance.

Simulated events are reweighted to account for differences with respect to data in the efficiencies of the lepton reconstruction, identification, and isolation requirements, and in the performance of b-jet identification. The values of the data-to-simulation scale factors differ from unity by less than 10% with typical efficiency corrections of 2-3 (5)% for the identification of leptons (b jets) with $p_T > 20$ GeV and $|\eta| < 2.4$. The scale factor for leptons requirements and b-jet identification have been officially provided by the particle analysis group of supersymmetry and the particle object group of b-tagging in the CMS Collaboration, respectively.

Chapter 6

Event selection and background estimation

The search strategy is developed for two signal hypotheses: the chargino pair and top squark pair productions. The first signal hypothesis is studied along the whole $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0})$ mass plane, while for the second one the analysis is optimized on the compressed scenario, where the mass difference of the top squark and the lightest neutralino lies between the top quark and W boson masses. Both searches involve the same techniques for the background estimation and signal extraction, while they differ slightly in the signal region (SR) selection in order to improve their respective sensitivities.

This chapter describes the common event selection in Section 6.1. It identifies the main backgrounds in the two lepton channel in Section 6.2. The dedicated signal regions (SRs) are defined in Section 6.3. The treatment of the SM processes contributing to the background in SRs is explained in Section 6.4. Finally the systematic uncertainties that affect the MC predictions are studied in Section 6.5.

6.1 Event selection

The signal models for top squark and chargino pair production are characterized by a common final state with two OC leptons and two lightest neutralinos contributing to the large p_T^{miss} . Based on this, a high-acceptance baseline selection is defined for both searches. Two OC isolated leptons satisfying the selection criteria related in Section 5.1, with $p_T > 25$ (20) GeV for the leading (trailing) lepton are required. Events with τ leptons decaying into electrons or muons that satisfy the selection requirements are taken into account in this analysis. To reduce the contributions from low mass resonances as those produced by Drell-Yan process, $Z \rightarrow \tau\tau$ production, and nonprompt leptons from hadronic jets, the invariant mass of the lepton pair, $m_{\ell\ell}$, is required to be greater than 20 GeV, and if both leptons have the same flavour (SF), $m_{\ell\ell}$ is further required to satisfy $|m_{\ell\ell} - m_Z| > 15$ GeV, where m_Z is the mass of the Z boson. High $p_T^{miss} (\geq 140$ GeV) is required. Events are further rejected if they contain a third lepton with $p_T > 15$ GeV and $|\eta| < 2.4$, satisfying the veto lepton selection as detailed in Section 5.1. This condition rejects those backgrounds with at least three leptons in the final state. A summary of this baseline selection is found in Table 6.1.

The expected contribution from SM backgrounds to this region comes from those processes decaying into final states with two well defined high- p_T lepton candidates

Variable	Selection
Lepton flavour	$e^+e^-, \mu^+\mu^-, e^\pm\mu^\mp$
Leading lepton	$p_T > 25 \text{ GeV}, \eta < 2.4, I_1^{rel} < 0.12$
Trailing lepton	$p_T > 20 \text{ GeV}, \eta < 2.4, I_1^{rel} < 0.12$
Third lepton veto	$p_T > 15 \text{ GeV}, \eta < 2.4, I_1^{rel} < 0.4$
$m_{\ell\ell}$	$\geq 20 \text{ GeV}$
$ m_{\ell\ell} - m_Z $	$> 15 \text{ GeV}$ only for ee and $\mu\mu$ events
p_T^{miss}	$\geq 140 \text{ GeV}$

TABLE 6.1: Definition of the baseline selection used in the searches for chargino and top squark pair production in dilepton final states.

and high p_T^{miss} . Further signal region definitions as described in Section 6.3 will exploit the kinematic differences between signal and background to optimize a phase space with a significant discriminatory power.

Selection	$\tilde{\chi}_1^\pm \rightarrow \tilde{l}\nu(\tilde{\nu}l) \rightarrow l\nu\tilde{\chi}_1^0$ ($m_{\tilde{\chi}_1^\pm}=500 \text{ GeV}, m_{\tilde{\chi}_1^0}=200 \text{ GeV}$)	$\tilde{t}_1 \rightarrow t\tilde{\chi}_1^0$ ($m_{\tilde{t}_1}=350 \text{ GeV}, m_{\tilde{\chi}_1^0}=225 \text{ GeV}$)
$\sigma\mathcal{B}(\tilde{\chi}_1^\pm\tilde{\chi}_1^\mp / \tilde{t}_1\tilde{t}_1 \rightarrow \ell\ell X)\mathcal{L}$	794	14270
Two OC leptons	257	2373
Third lepton veto	256	2301
$m_{\ell\ell} > 20\text{GeV}$	255	2237
$ m_{\ell\ell} - m_Z > 15 \text{ GeV}$ (SF events)	246	1938
$p_T^{miss} > 140 \text{ GeV}$	143	466

TABLE 6.2: Expected signal yields at different stages of the baseline event selection for two representative signal points, one for chargino pair production and one for top squark pair production. The yields are normalized to an integrated luminosity of 35.9 fb^{-1} .

Considering two representative signal points, Table 6.2 shows the stream of signal's yield through the different cuts of the baseline selection.

6.2 Main backgrounds in the dilepton channel

The set of SM processes with signature similar to events from top squark and chargino pair production presents an irreducible background to the data in the signal search regions of this analysis. These processes are individually studied and their expected contributions are estimated with dedicated methods. In this section, the considered SM backgrounds are classified in different categories.

Top backgrounds

- **Top-antitop quarks pair production ($t\bar{t}$):** The production of this process in pp collisions at LHC is possible via gluon-gluon fusion or $q\bar{q}$ annihilation at leading order as illustrated in Figure 6.1. Due to their large mass, top quarks (t) decay before hadronization. Almost to 100%, the top quark decays into a W^\pm boson and a b quark through the electroweak interaction. In turn the W^\pm boson decays leptonically or hadronically determining thus the final states of $t\bar{t}$ decays.

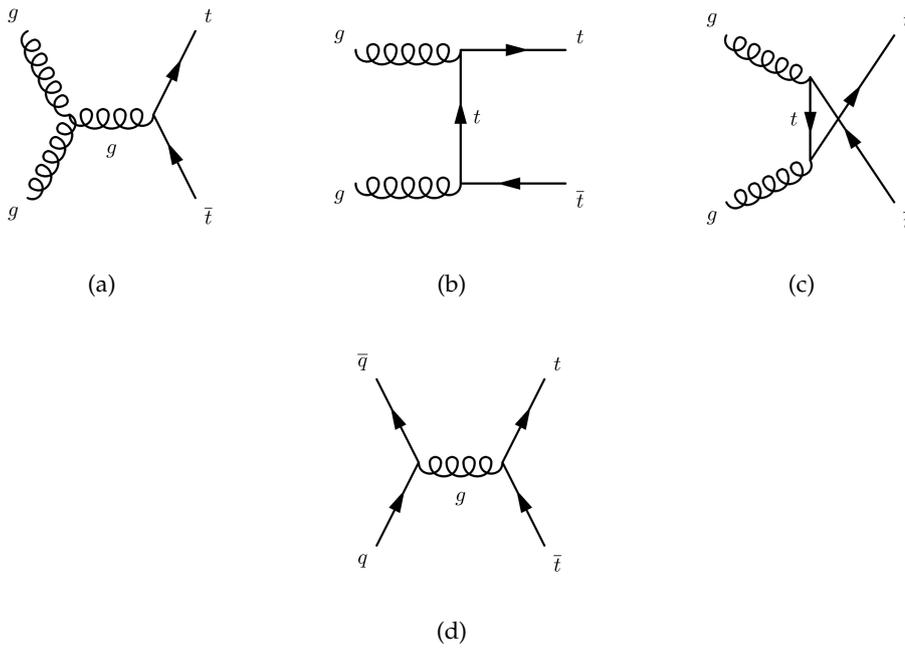


FIGURE 6.1: Leading-order Feynman diagrams for $t\bar{t}$ production in proton-proton collisions. The four diagrams represent contributions from gluon fusion through the s -channel (a), t -channel (b) or the u -channel (c), and quark-antiquark annihilation (d).

The full decay chain is shown in Figure 6.2. When the W^\pm decays leptonically, namely into a lepton and a neutrino, the $t\bar{t}$ reproduces the $\tilde{t}_1\tilde{t}_1$ signature and its contribution represents a predominant background in the signal regions. In the case of $\tilde{\chi}_1^\pm\tilde{\chi}_1^\mp$ search, the presence of the W^\pm leptonic decays gives a major contribution to the total backgrounds in the corresponding signal regions.

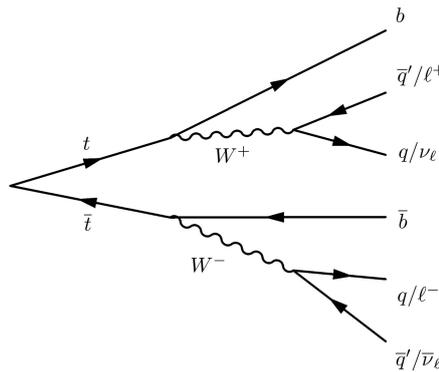


FIGURE 6.2: Feynman diagrams for the decay chain of a $t\bar{t}$ pair into a $b\bar{b}$ pair and W^\pm bosons at tree level, including the subsequent hadronic and leptonic W decays with $\ell = e, \mu, \tau$.

- **Single top quark production in association with a W boson (tW):** The top quarks are mainly produced in $q\bar{q}$ pairs through the strong interaction at the LHC, but they can also be produced individually via a charged-current electroweak interaction. At leading order, the single top production proceeds through three separate sub-processes as it is displayed in Figure 6.3. The single

top quark production in association with a W boson (Figure 6.3(c)–(d)) is one of the main backgrounds in this analysis when the W bosons decay leptonically.

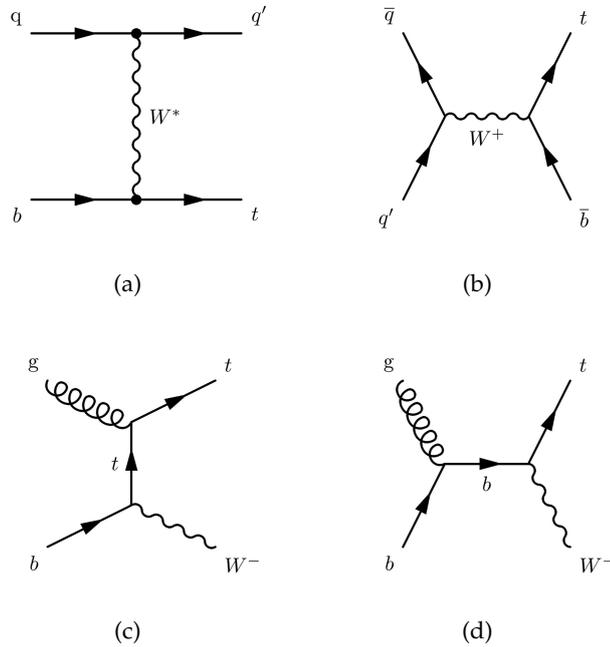


FIGURE 6.3: Leading-order Feynman diagram for single top quark production in the t -channel (a), the leading-order diagram in the s -channel (b), and the tW production channel (c)–(d).

- **Top-antitop pair production in association with a boson ($t\bar{t}Z$, $t\bar{t}W$):** The production of $t\bar{t}$ associated with a W^\pm or a Z boson can also have an expected contribution in the signal regions. The dominant production modes at leading order for $t\bar{t}Z$ and $t\bar{t}W$ in pp collisions at LHC are shown in Figure 6.4. These processes can enter as backgrounds in our signal regions when decaying into at least two leptons and satisfy the p_T^{miss} requirements.

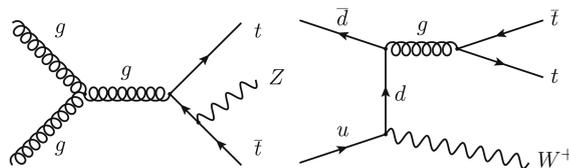


FIGURE 6.4: Primary Feynman diagrams for $t\bar{t}Z$ (left) and $t\bar{t}W^+$ (right) production at LHC. The charge conjugate process of $t\bar{t}W^+$ produces $t\bar{t}W^-$.

Z boson backgrounds

In pp collisions at LHC, the Z boson production is dominated by the Drell-Yan process via $q\bar{q}$ annihilation as it is shown in Figure 6.5(a)–(b). Although in a minor way the Z boson can be produced by other processes as Figure 6.5(c)–(d) illustrates. The Figure 6.6 shows the fermionic decay chain of a Z boson.

- **Drell-Yan process:** This process occurs when a quark from one hadron annihilates with an antiquark from another hadron and is converted into a virtual photon (γ^*) or a Z boson, which decays into two leptons of same flavour and opposite charge. Final states with different flavour leptons are mediated through the $\tau\tau$ decays ($Z \rightarrow \tau\tau \rightarrow \mu\nu_\mu\nu_\tau + e\nu_e\nu_\tau$), where the $\mathcal{BR}(\tau \rightarrow e\mu) \sim 17\%$.

Due to the interference between the Z boson and γ^* production, the individual contribution of each particle cannot be easily disentangled. However, in the case of Z exchange, the invariant mass value of the produced lepton pair will mostly fall into a window of width 30 GeV centered at the mass of the Z boson (91.2 GeV). While the γ^* will produce low mass resonances typically below the 50 GeV.

- **Z+jets production:** It is usual the production of neighboring jets at leading order (Figure 6.5(b), (d)). These events can pass our selection criteria contributing to the signal region. The invariant mass of the decay products will also reproduce the resonance peak around the Z boson mass.

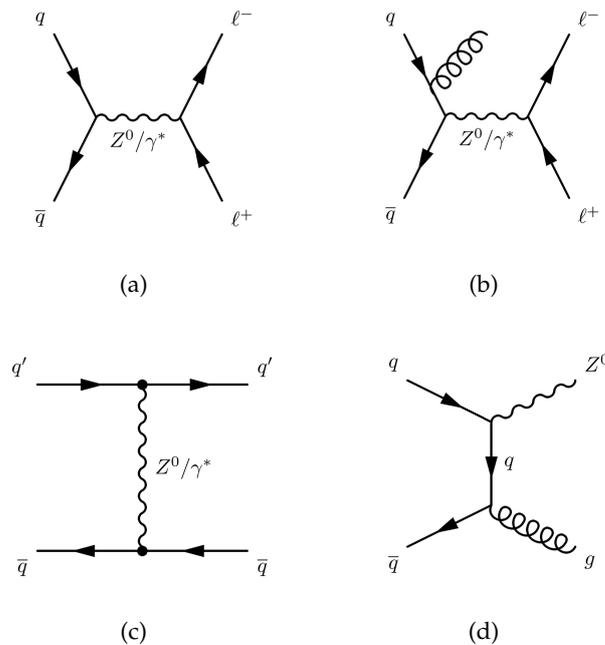


FIGURE 6.5: Leading order Feynman diagrams for Z boson production at LHC through Drell-Yan process (a)–(b), the $q\bar{q}$ scattering in the t -channel (c) and the $q\bar{q}$ annihilation producing a gluon and a Z^0 boson (d).

Typically, in events with two leptons, the Drell-Yan and Z + jets processes populate the lower regions of p_T^{miss} values. But, because of the detector resolution effects, misreconstruction or mismeasurement issues, this background can be found at high p_T^{miss} selection with a non-negligible expected contribution.

Diboson backgrounds

- **W^\pm boson pair production (WW):** The leading order production of WW at LHC is dominated by quark-antiquark annihilation through the triple gauge

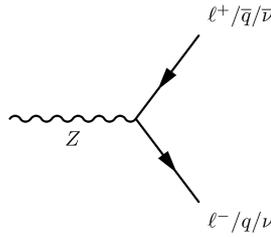


FIGURE 6.6: Feynman diagram for the Z boson fermionic decays.

coupling (TGC) in the s -channel or the quark-antiquark scattering in the t -channel, as shown in Figure 6.7(a) and (b) respectively. The gluon-gluon fusion is also possible via a quark loop similar to that illustrated in Figure 6.7(c). When both W^\pm bosons decay leptonically this process becomes the second greater contributor to the background in the signal regions. The final states with two leptons and p_T^{miss} exactly reproduce the $\tilde{\chi}_1^\pm \tilde{\chi}_1^\mp$ signature being essential in the $\tilde{t}_1 \tilde{t}_1$ search.

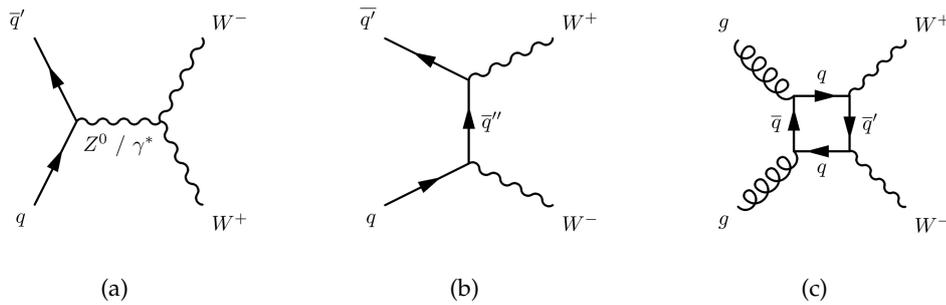


FIGURE 6.7: Representative leading order Feynman diagrams of WW production at LHC through a triple gauge vertex via $q\bar{q}$ annihilation (a), or by the $q\bar{q}$ scattering in the t -channel (b) and through a quark loop initiated by gluon-gluon interaction (c).

- **Z boson pair production (ZZ):** At LHC, this process is also predominantly produced through quark-antiquark interaction followed by gluon fusion diagrams similar to those for the WW process. Figure 6.8 shows two important ZZ production mechanisms at leading order in pp collisions. The ZZ decay chain into two leptons and two neutrinos generates events typically with high p_T^{miss} that populate the tails of our distributions.
- **W^\pm production in association with a Z boson ($W^\pm Z$):** The $W^\pm Z$ diboson production at leading order arises predominantly in the LHC from quark-antiquark initial states. Figure 6.9 shows the LO Feynman diagrams for the $W^\pm Z$ production from $q\bar{q}'$ annihilation. Events from WZ production mainly enter in our SRs when both bosons decay leptonically and the third lepton fails the veto requirement.

Triboson backgrounds

The triboson processes usually involve triple and quartic interaction vertices between charged bosons or charged and neutral bosons ($W^+W^-W^+W^-$, W^+W^-Z , $W^+W^- \gamma$, ...). In this analysis only three of these final states are considered: WWW ,

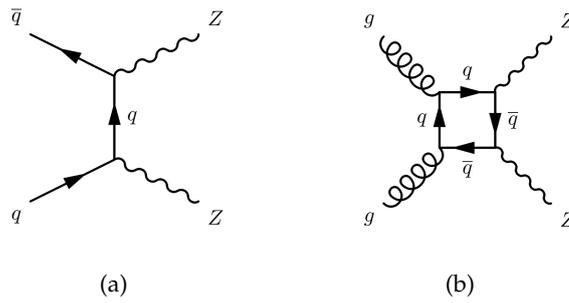


FIGURE 6.8: Example of Feynman diagrams for ZZ production initiated by quarks (a) and gluon fusion (b).

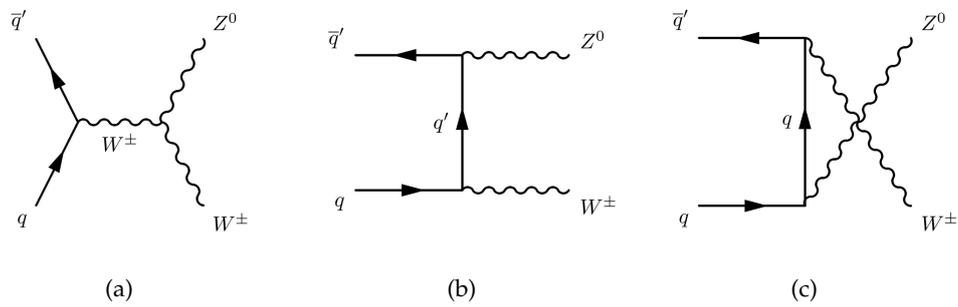


FIGURE 6.9: Leading-order Feynman diagrams for WZ production in proton-proton collisions. The three diagrams represent the contributions from the s -channel through a triple gauge vertex (a), t -channel (b), and u -channel (a).

WWZ, and WZZ. The leading order structures of Feynman diagrams contributing to those processes cross section at LHC are shown in Figure 6.10. The $q\bar{q} \rightarrow VVV$ production is dominant with respect to the gluon production.

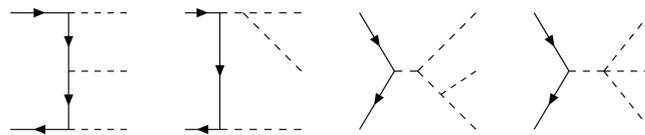


FIGURE 6.10: Tree-level structures of Feynman diagrams contributing to the $q\bar{q} \rightarrow VVV$ cross section, with $V = Z, W^\pm$ and the dashed lines represent W^\pm, Z , goldstone bosons or photons.

- **triboson production (WWW, WWZ, WZZ, ZZZ):** An example of leading order Feynman diagrams for the triboson production at LHC is shown in Figure 6.11. These processes can contribute to the total background in our signal regions when decaying into at least two leptons and they satisfy the p_T^{miss} requirements.

Higgs backgrounds

- **Higgs production:** In pp collisions, the Higgs boson can be produced in different ways. The main mechanisms of Higgs production at leading order in the

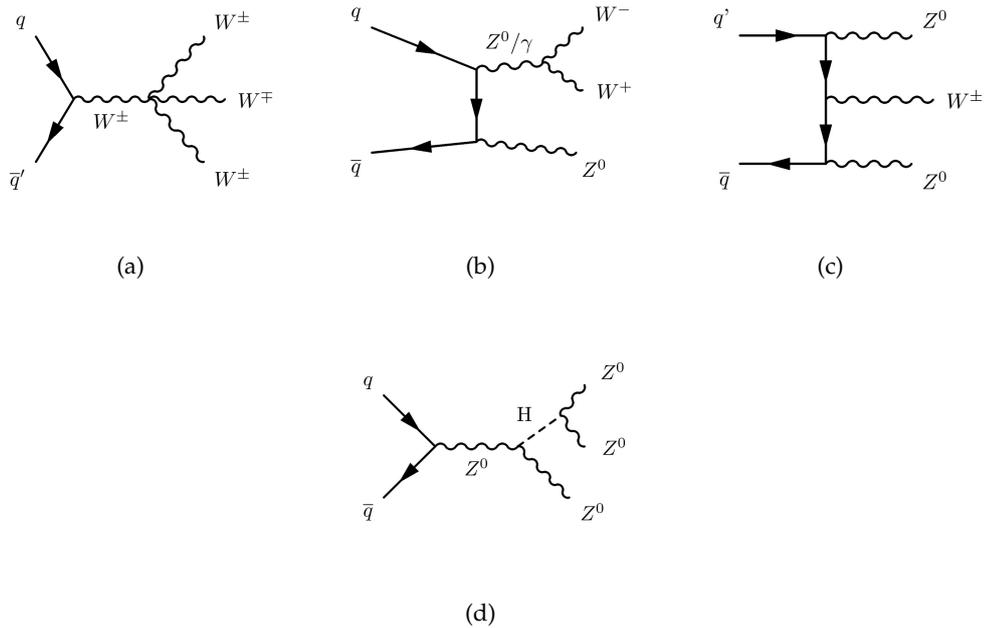


FIGURE 6.11: Representative tree-level diagrams for triboson processes: WW (a), WWZ (b), WZZ (c) and ZZZ (c) production.

LHC are shown in Figure 6.12. Two decays channels have been considered for this analysis: Higgs decays into WW and $\tau\bar{\tau}$.

Nonprompt lepton backgrounds

In this analysis, the processes under this category are characterized by events with two leptons where one of them is a nonprompt or fake lepton mistaken as prompt lepton. While a prompt lepton comes from the interaction vertex, the fake leptons are produced by other type of process such as meson decays in jets, cosmic rays or jets misidentified as leptons, for instance. When this fake particle passes all the selection criteria enters in the SRs without suspicion. These kind of backgrounds are mostly suppressed by the identification, isolation and selection requirements imposed to the electrons and muons, and the remaining contribution after the signal regions selection can be negligible. Those processes are:

- **W^\pm boson production in association with jets ($W + jets$):** Two simple LO Feynman diagrams for the $W + jets$ production are shown in Figure 6.13. This background consists of a real W boson produced in association with quarks or gluons. The gluon can split into a pair of heavy flavour quarks. As a result, in the W leptonic final states there will be jets, one lepton and p_T^{miss} . If a fake lepton arises from the misidentified jet, or is taken from a neighbouring decay, these events have the same final state of the signal (two leptons and missing transverse energy).
- **$t\bar{t}$ semileptonic production:** This process can contribute to the background in the dilepton channel only if there is one fake lepton. An example of semileptonic $t\bar{t}$ decay is shown in Figure 6.14.

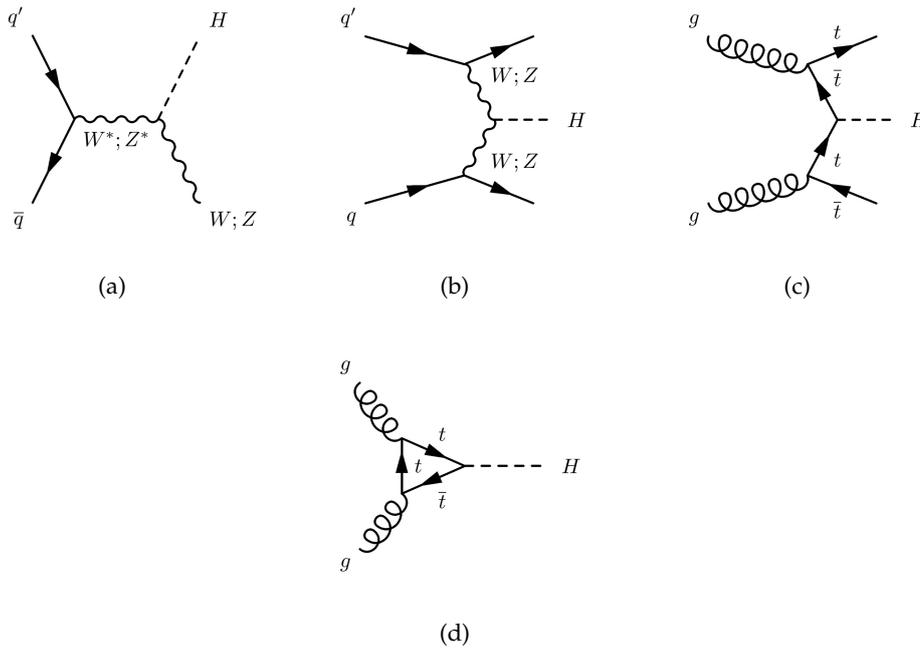


FIGURE 6.12: The main channels for Higgs boson production at LO in pp collisions: the Higgs strahlung (a), through vector boson fusion (VBF) (b), in association with a top quark production (c) and through the gluon fusion (d).

6.3 Analysis strategy and signal region definition

The SM processes that contribute most after the baseline selection are $t\bar{t}$, tW , and WW production. For all these backgrounds, the lepton pair and the \vec{p}_T^{miss} come from a W boson pair. The transverse mass (m_T) of the system formed by one visible and one massless invisible particles, in this case one lepton and one neutrino, can be written as follows,

$$m_T = \sqrt{2E_T^{lep} p_T^{miss} \cdot (1 - \cos(\Delta\phi))}, \quad (6.1)$$

where the mass of the lepton is considered negligible and the p_T^{miss} corresponds to the transverse momentum of the neutrino. The angular difference $\Delta\phi$ is the angular

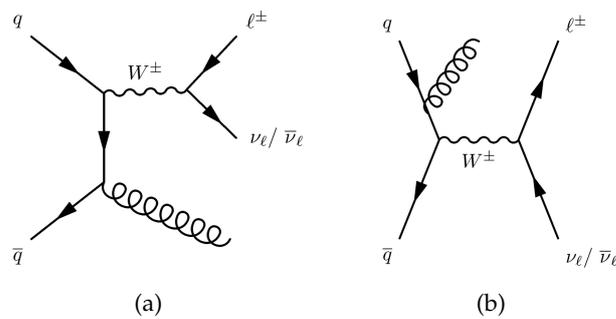


FIGURE 6.13: LO diagrams for W +jets production in the LHC.

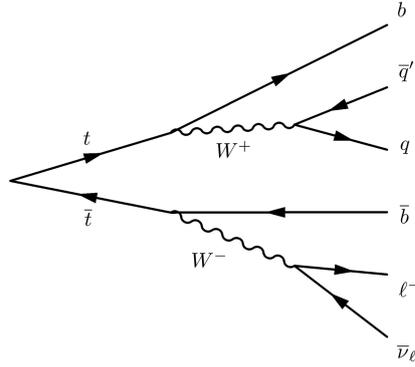


FIGURE 6.14: Feynman diagrams for the semileptonic decay chain of a $t\bar{t}$ pair into a $b\bar{b}$ pair and W^\pm bosons at tree level, with the subsequent leptonic and hadronic W decays ($\ell = e, \mu, \tau$).

separation between the neutrino and the lepton in the transverse plane. The magnitude m_T has the property that,

$$m_T \leq m_W, \quad (6.2)$$

being the equality possible for events where the lepton and the neutrino are produced with the same rapidity.

The generalization of this variable in a system of two particles with the same mass decaying semi-invisibly cannot use Equation 6.1 anymore because of the p_T of the individual missing particles is unknown. Instead of that, a new variable whose value is based on minimizing the set of transverse masses formed by all possible values of p_T^{miss1} and p_T^{miss2} which sum gives the measured p_T^{miss} is developed. This observable denoted by m_{T2} [166] is by definition an event-by-event quantity constructed as follows:

$$m_{T2}(ll) = \min_{\vec{p}_T^{miss1} + \vec{p}_T^{miss2} = \vec{p}_T^{miss}} (\max[m_T(\vec{p}_T^{\ell ep1}, \vec{p}_T^{miss1}), (m_T(\vec{p}_T^{\ell ep2}, \vec{p}_T^{miss2}))]), \quad (6.3)$$

where \vec{p}_T^{miss1} , and \vec{p}_T^{miss2} are the transverse momentum vectors of any possible neutrino pair in the phase space of the event and $\vec{p}_T^{\ell ep1}$, and $\vec{p}_T^{\ell ep2}$ are the measured transverse momentum vectors of the leptons from the WW decay. Inheriting the m_T property that,

$$m_{T2}(ll) \leq m_W. \quad (6.4)$$

That can be equal, under special kinematic conditions, when both visible leptons and invisible particles are produced with the same rapidity.

This observable reaches a kinematic endpoint at the m_W for the considered backgrounds. Signal events, instead, present the $m_{T2}(ll)$ spectra without such an endpoint because of the additional contribution to the \vec{p}_T^{miss} give by the neutralinos. The sensitivity of the analysis is further enhanced by dividing the baseline region in three bins of p_T^{miss} : $[140, 200)$, $[200, 300)$, ≥ 300 GeV. This allows the analysis not only to exploit the larger tails in the p_T^{miss} distribution of the signal events, but also to optimize the sensitivity to signals with different mass separation between the produced supersymmetric particles and LSP. Each p_T^{miss} bin is in turn divided into events with

SF and different flavour (DF) lepton pairs, which allows to exploit the smaller contamination from WZ , ZZ , and Drell-Yan production in the latter channel.

Each p_T^{miss} bin constitutes an exclusive SR that is further subdivided according to the specific characteristic of each signal model. A veto on b-tagged jets is applied to reject $t\bar{t}$, tW , and $t\bar{t}Z$ events in the chargino search. Selected events in the p_T^{miss} bins below 300 GeV are then split into two different subregions depending on the presence of a jet with $p_T \geq 20$ GeV and $|\eta| < 2.4$. This exploits the larger fraction of signal events with no jets respect to top background ($t\bar{t}$ and tW), which still contaminates the SRs after applying the b-tagged jet veto. Events with b-tagged jets are kept as a CR for the normalization of the background from $t\bar{t}$ and tW production.

On the other hand, the final states produced in the top squark decays are characterized by the presence of two bottom quarks. When the difference in the mass of the top squark and the neutralino is close to the m_W edge of the compressed region satisfying $\Delta m \gtrsim m_W$, the bottom quarks are soft and give rise to jets with relatively low momentum that have lower probability to be tagged. In this case, the kinematics of the top squark is similar to those for chargino pair production, and requiring a veto on b-tagged jets is again an effective strategy to define SRs with reduced contamination from $t\bar{t}$, tW and $t\bar{t}Z$ backgrounds. The main irreducible backgrounds are the $t\bar{t}$, tW and WW processes. For signal scenarios with larger Δm , instead, the b jets have higher momentum and the final final states kinematics is more comparable to $t\bar{t}$ kinematics leading a similar phase space. Consequently, sensitivity to top squark production is enhanced by requiring a b-tagged jet to reduce the background from diboson production, especially WW production and Drell-Yan events. For these events the main irreducible backgrounds are $t\bar{t}$ and tW production.

The p_T^{miss} and $m_{T2}(ll)$ distributions for events with and without b-tagged jets after the baseline selection are shown in Figures 6.15 and 6.16 for different lepton flavour pairs in final states. The cut on $p_T^{miss} > 140$ GeV has not been applied on Figures 6.15. The contributions and the shape behaviour of each background highlight the main features of the SM processes in our signal regions. An overall good agreement is observed in both variables. Note that in the following, the name VVV stands for triboson production and unless the opposite is said, the name VZ stands for $WZ/ZZ \rightarrow 2\ell 2q$ decay modes. They are considered as minor backgrounds.

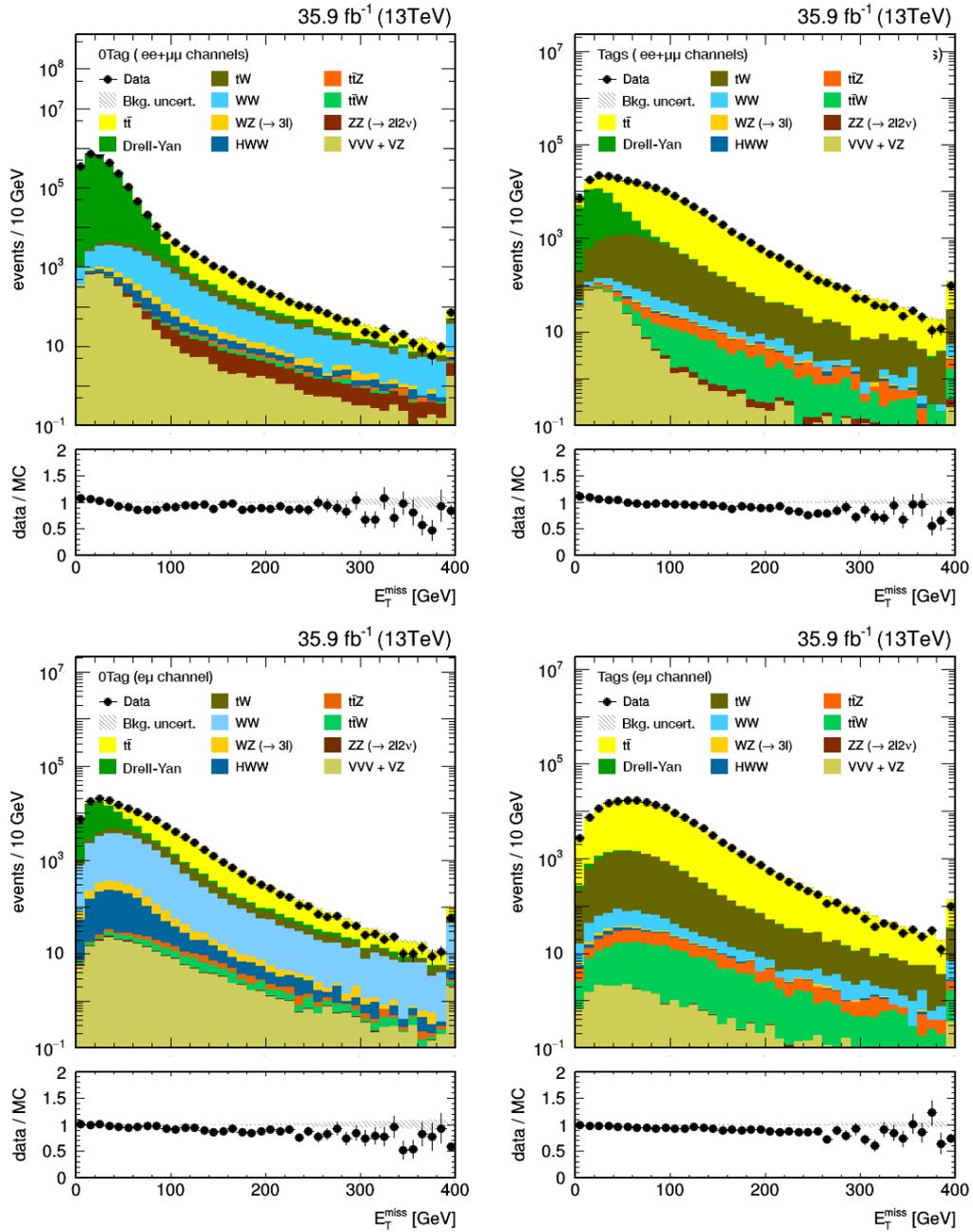


FIGURE 6.15: Observed and expected distributions of p_T^{miss} in events with two opposite sign leptons passing the Z veto selection. Left plots show events without b-tagged jets, while right plots show events with at least one b-tagged jet. Top plots, final states with same lepton flavour pairs ($ee+\mu\mu$); bottom plots, different lepton flavour pair ($e\mu$) final states. The lower panels show the ratio data to total SM contribution in each bin. Only statistic uncertainties are considered.

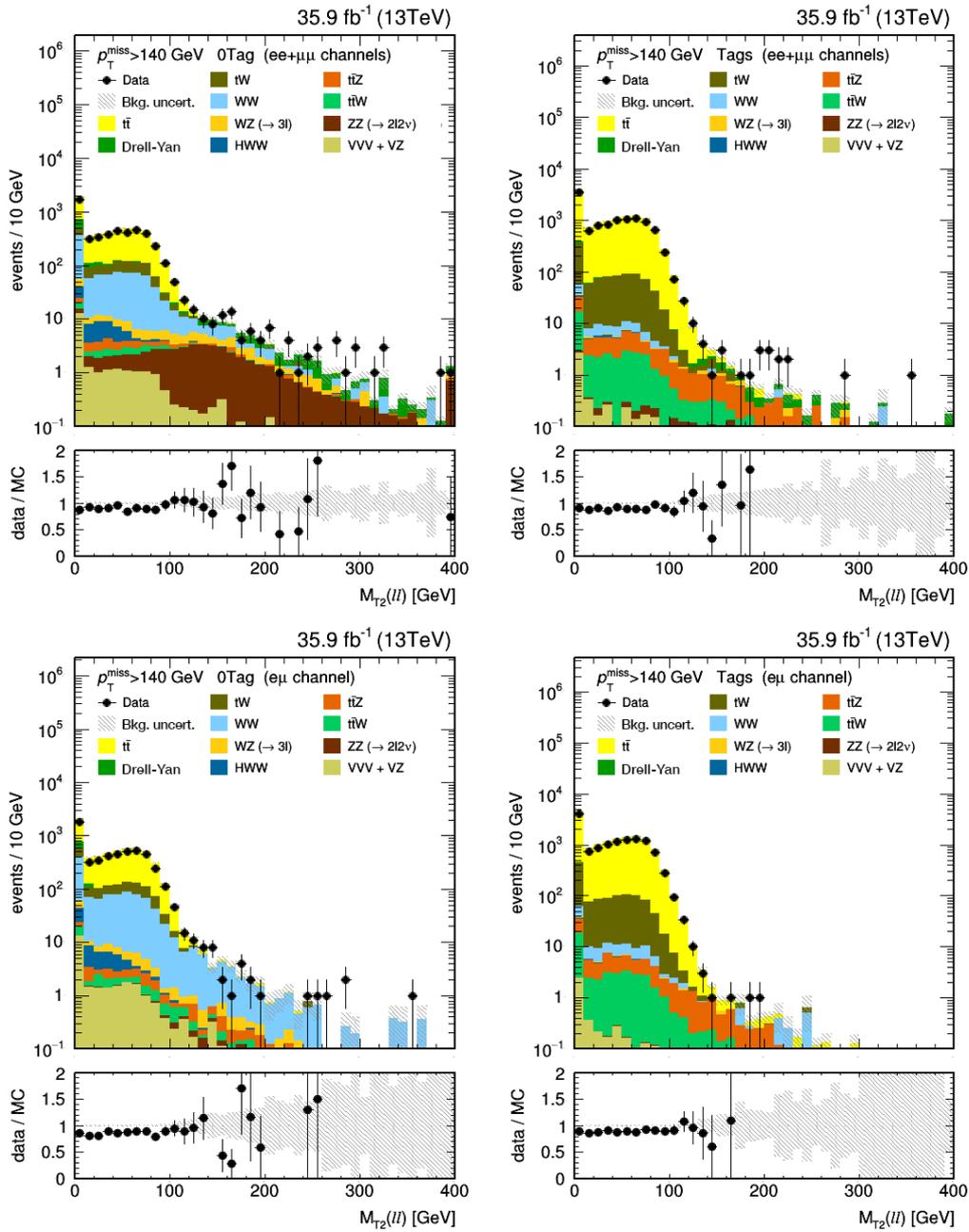


FIGURE 6.16: Observed and expected distributions of $m_{T2}(ll)$ in events with two opposite sign leptons passing the Z veto selection and with $p_T^{miss} > 140$ GeV. Left plots show events without b -tagged jets, while right plots show events with at least one b -tagged jet. Top plots, final states with same lepton flavour pairs ($ee + \mu\mu$); bottom plots, different lepton flavour pair ($e\mu$) final states. The lower panels show the ratio data to total SM contribution in each bin. Only statistic uncertainties are considered.

The correlation between these two variables in events with two leptons can be observed for $t\bar{t}$ (left) and top squark (right) MC samples in Figure 6.17. A strong correlation is exhibited between the two observables up to $m_{T2}(ll)$ values of the order of m_W , becoming largely uncorrelated at high p_T^{miss} values.

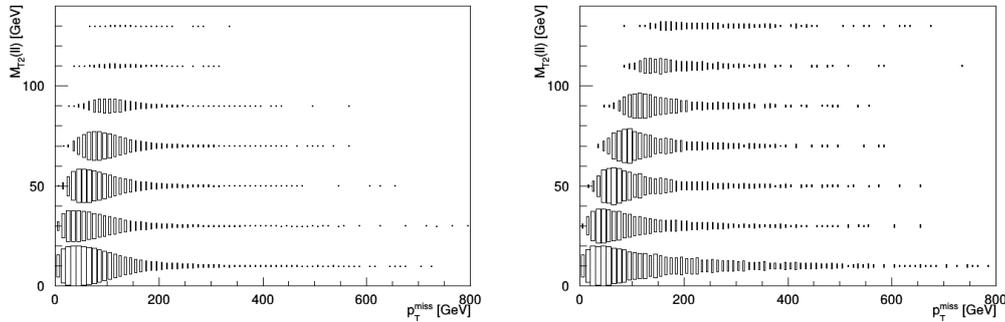


FIGURE 6.17: Distribution of $m_{T2}(ll)$ and p_T^{miss} variables for events with two leptons in the $t\bar{t}$ sample (left) and the top squark sample for the mass point $(m_{\tilde{t}_1}, m_{\tilde{\chi}_1^0}) = (350, 225)$ GeV (right).

Another useful means to discriminate top squark production from SM processes is given by the presence of high- p_T jets from ISR in the events (see Section 4.3.1). The invisible particles (neutrinos and neutralinos) produced in the decay chain of the top squark in the compressed scenario are expected to be soft; events with harder neutralinos, however, can arise when the top squark pair system recoils against a high- p_T ISR jet. In this hard ISR regime, background is still constrained by the kinematic m_W endpoint in $m_{T2}(ll)$ and can be effectively separated from the signal. Hard ISR events are selected by requiring that the leading jet satisfies $p_T > 150$ GeV and is not b tagged. In order to favor the topology in which the jet recoils against the rest of the system, the $\Delta\phi$ between the jet and the \vec{p}_T^{miss} is required to be larger than 2.5 rad. This requirement is found to be effective in discriminating top squark production from background events at high p_T^{miss} and is therefore applied only for events with $p_T^{miss} > 300$ GeV.

However, since the chargino pair production is studied in the whole mass plane, in the phase space of this signal events can appear larger mass differences between chargino and neutralino giving naturally signatures with higher p_T^{miss} value. The improvement of the ISR jet requirement in the chargino signal regions has been found negligible and is therefore not considered in this search.

A summary of the SRs for the chargino and top squark searches is given in Tables 6.3 and 6.4, respectively, indicating the p_T^{miss} range, the selection on the multiplicity of jets (N_{jets}) and b jets ($N_{b jets}$) in the event, and the ISR jet requirement. The observed distributions of main observables used to define the SRs after the baseline selection are compared to SM expectations in Figure 9.1. An excellent agreement is observed.

Each of the SRs defined in Tables 6.3 and 6.4 is further divided into seven $m_{T2}(ll)$ bins of 20 GeV width, starting from 0 GeV and with the last bin collecting all events with $m_{T2}(ll) > 120$ GeV. A simultaneous binned ML fit of the $m_{T2}(ll)$ distribution in all the SRs is then performed to extract the signal strength and set the upper limits on the production cross section for each signal point. Since the first $m_{T2}(ll)$ bins have a low signal contribution, we exploit them to constrain the contributions of the dominant backgrounds in the SRs with one b-tagged jet (dominated by $t\bar{t}$ and tW production)

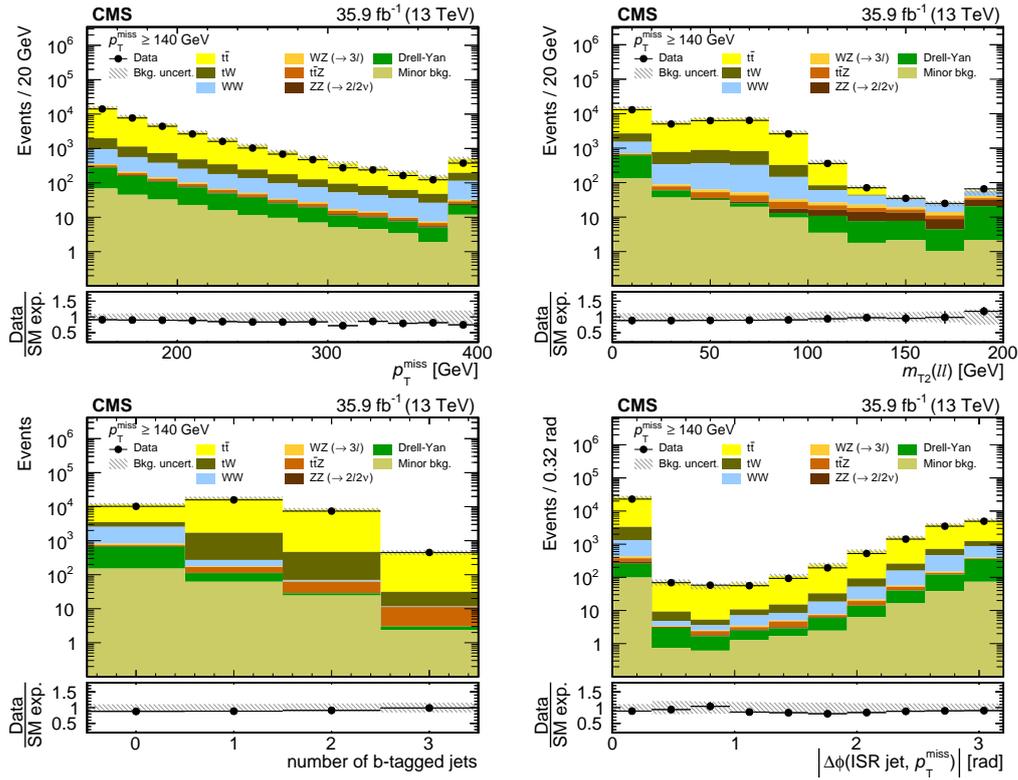


FIGURE 6.18: Observed and SM expected distributions of some observables used to define the SRs for events with two OC isolated leptons and $p_T^{\text{miss}} \geq 140$ GeV. Clockwise from top left: p_T^{miss} , $m_{T2}(ll)$, $\Delta\phi$ between the \vec{p}_T^{miss} and the leading jet (required not to be b-tagged and with $p_T > 150$ GeV, events missing this requirements are shown in the first bin), and multiplicity of b-tagged jets in the event. The last bin includes the overflow entries. The contributions of minor backgrounds such as $t\bar{t}W$, $H \rightarrow WW$, and triboson production are grouped together. In the bottom panel, the ratio of observed and expected yields is shown. The hatched band represents the total uncertainty (statistical and systematic) in the background expectation, as described in Section 6.5 [19].

	$SR1_{0tag}^{0jet}$	$SR1_{0tag}^{jets}$	$CR1_{tags}$	$SR2_{0tag}^{0jet}$	$SR2_{0tag}^{jets}$	$CR2_{tags}$	$SR3_{0tag}$	$CR3_{tags}$
p_T^{miss} [GeV]	140–200	140–200	140–200	200–300	200–300	200–300	≥ 300	≥ 300
$N_{b\ jets}$	0	0	≥ 1	0	0	≥ 1	0	≥ 1
N_{jets}	0	≥ 1	≥ 1	0	≥ 1	≥ 1	≥ 0	≥ 1
Channels	SF, DF	SF, DF	SF, DF	SF, DF	SF, DF	SF, DF	SF, DF	SF, DF
$m_{T2}(ll)$	0–20, 20–40, 40–60, 60–80, 80–100, 100–120, ≥ 120 GeV							

TABLE 6.3: Definition of the SRs for the chargino search as a function of the p_T^{miss} value, the b-jet multiplicity and jet multiplicity. Also shown are the CRs with b-tagged jets used for the normalization of the $t\bar{t}$ and tW backgrounds. Each of the regions is further divided in seven $m_{T2}(ll)$ bins as described in the last row.

	$SR1_{0tag}$	$SR1_{tags}$	$SR2_{0tag}$	$SR2_{tags}$	$SR3_{0tag}^{ISR}$	$SR3_{tag}^{ISR}$
p_T^{miss} [GeV]	140–200	140–200	200–300	200–300	≥ 300	≥ 300
$N_{b\ jets}$	0	≥ 1	0	≥ 1	0	≥ 1
N_{jets}	≥ 0	≥ 1	≥ 0	≥ 1	≥ 1	≥ 2
ISR jets	≥ 0	≥ 0	≥ 0	≥ 0	≥ 1	≥ 1
Channels	SF, DF	SF, DF	SF, DF	SF, DF	SF, DF	SF, DF
$m_{T2}(ll)$	0–20, 20–40, 40–60, 60–80, 80–100, 100–120, ≥ 120 GeV					

TABLE 6.4: Definition of the SRs for top squark search as a function of the p_T^{miss} value, the b-jet multiplicity and the ISR jet requirement. Each of the regions is further divided in seven $m_{T2}(ll)$ bins as described in the last row.

and without b-tagged jets (where WW production becomes relevant) through the fit.

6.4 Background estimation

The main contributions from SM processes to the SRs come from $t\bar{t}$, tW , and WW . The normalization of these backgrounds is determined by the ML fit, as it is described in Chapter 7. The shape of $m_{T2}(ll)$ distribution for $t\bar{t}$, tW , and WW processes has a natural endpoint at the W boson mass by construction. Likewise, the events with $m_{T2}(ll) > 80$ GeV that enter in the relevant regions for signal extraction are mainly due to detector resolution effects, whose contribution is not easy to model in simulation. For this reason, we study the modeling of the $m_{T2}(ll)$ distribution for these processes in dedicated CRs in data and examine the contribution of non-prompt leptons in the events, as described in Section 6.4.1. The contributions of the subleading $t\bar{t}Z$, WZ , ZZ and Drell-Yan backgrounds are also tested in CRs, where correction factors for their normalization are extracted using the equation:

$$SF_{norm} = \frac{Data - \sum_{i \neq X} Bkg^i}{X}. \quad (6.5)$$

The studied process is denoted by X and the sum runs over the yields of all residual backgrounds (Bkg^i) contributing to the CR. This sum is subtracted from data to

estimate the observed X contribution. The ratio between the X contribution in data and the expected contribution in simulation is the scale factor applied to simulation in order to correct the cross section normalization by the most accurate event rate, as discussed in Section 6.4.2 and the following.

Remaining minor backgrounds from $t\bar{t}W$, $H \rightarrow WW$, and triboson production (WWW , WWZ , WZZ) give small contributions in the SRs, and the estimates for these processes are taken directly from simulation. Background contributions from the rest of the processes are found to be negligible. The contribution of signal to any of the CRs used in the analysis is found to be negligible compared to SM processes.

6.4.1 Modeling of $m_{T2}(ll)$ in $t\bar{t}$, tW , and WW events

In this section, the modeling of $m_{T2}(ll)$ tails in simulation is studied by comparison with the observed distribution in data. The simulated $m_{T2}(ll)$ distributions for $t\bar{t}$, tW , and WW backgrounds are validated in two CRs. To design the first one, the baseline selection is modified by requiring $100 < p_T^{miss} < 140$ GeV. The events in this CR are further separated according to their b-jet multiplicity defining two sub-regions with different content in top quark ($t\bar{t}$ and tW) and WW backgrounds. In order to reject events from Drell-Yan production, only DF events are considered. Figure 6.19 shows the $m_{T2}(ll)$ distribution in $e^\pm\mu^\mp$ events with at least one b-tagged jet, which are dominated by top production. An overall good agreement between data and MC simulation is observed. The bin-by-bin comparison can be seen in the bottom plot, where the maximum discrepancy between bins is less than 5% and no trend is found.

On the other hand, Figure 6.20 shows the $m_{T2}(ll)$ distribution for events with no b-tagged jets, that enhance the contribution of WW production. In this figure two cases have been considered, on the left DF events with no b-tagged jets to further reduce the Drell-Yan contribution, while on the right, DF and SF events with 0 jets have been selected. When only b-tagged jets are vetoed, the comparison between the observed and expected tails of $m_{T2}(ll)$ distribution, does not present a significant trend under the range of the statistical uncertainty. The maximum deviation between bins is of about 5%. However, this deviation is increased to 30% on the last bins of $m_{T2}(ll)$, when the requirement of 0-jets is applied. A slight trend in the shape, above the statistical uncertainties, is suggested.

The second CR aims at validating the modeling of the $m_{T2}(ll)$ distribution in events with $p_T^{miss} > 140$ GeV. This region overlaps with the SRs of the analysis and a signal-free space is not easily accessible. Therefore, events from $WZ \rightarrow 3\ell 1\nu$ production are selected to emulate the $m_{T2}(ll)$ shape of WW background. The \vec{p}_T of the lepton from the Z boson decay that has the same charge as the lepton from the W boson is vectorially added to the \vec{p}_T^{miss} of the event treating it effectively like a neutrino. If the event satisfies the baseline selection, the observable $m_{T2}(ll)$ is then recomputed with the remaining leptons and the new p_T^{miss} value.

Events are selected by requiring three leptons and vetoing the presence of a fourth lepton passing the veto lepton requirements. A veto is applied to events with b-tagged jets to remove residual $t\bar{t}$ events. Among the three leptons, a pair of OC same flavour leptons with an invariant mass within 10 GeV of the Z boson mass is required to identify the Z boson. If no pair is found the event is rejected. The result

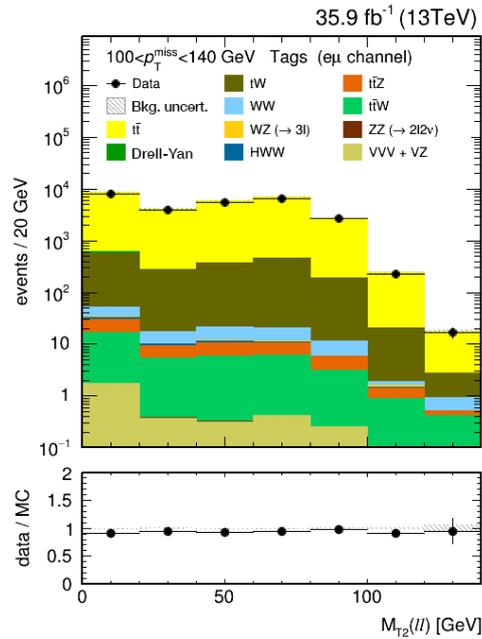


FIGURE 6.19: Observed and expected $m_{T2}(ll)$ distribution for $e^{\pm}\mu^{\mp}$ events with a b-tagged jet in the validation region $100 < p_T^{miss} < 140$ GeV. The lower panel shows the ratio data to total SM contribution in each bin. Only statistic uncertainties are considered.

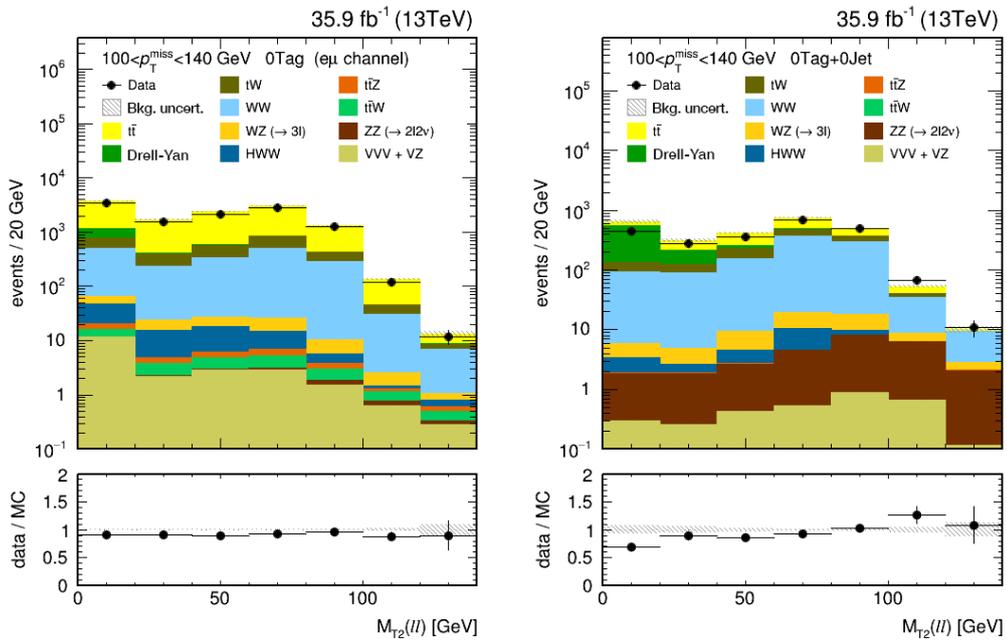


FIGURE 6.20: Observed and simulated $m_{T2}(ll)$ distribution for $e^{\pm}\mu^{\mp}$ events without b-tagged jet (left) and all lepton flavour pairs with no jets (right) in the validation region $100 < p_T^{miss} < 140$ GeV. The lower panels show the ratio data to total SM contribution in each bin. Only statistic uncertainties are considered.

of this test is shown in Figure 6.21. Good agreement between data and simulation is observed within the statistical uncertainty.

The difference in mass between the W and Z bosons is considered a negligible effect in the modeling of $m_{T2}(ll)$ tails. In Figure 6.22, the $m_{T2}(ll)$ shape for WZ (in its WW mimic) and WW nominal process are compared. There is no relevant discrepancy between distributions apart from the bin with $m_{T2}(ll)$ from 80 to 100 GeV where the difference in W and Z bosons mass may be appreciable. For WW events, an endpoint of the $m_{T2}(ll)$ shape is expected at the W boson mass, while for WZ events, due to the larger Z boson mass, the endpoint is expected to be shifted. This difference does not represent a bias in the $m_{T2}(ll)$ tails validation.

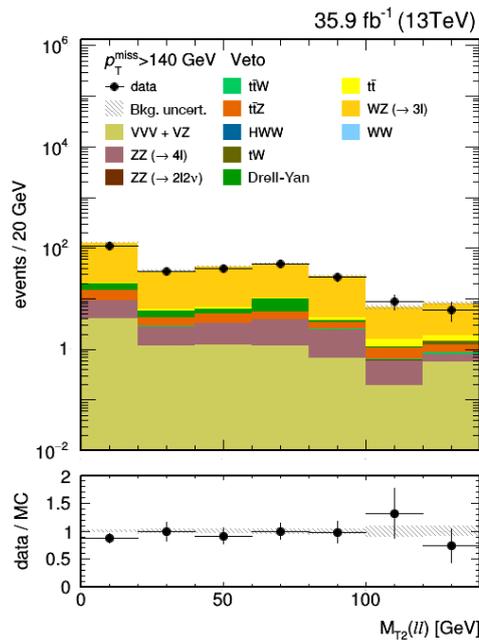


FIGURE 6.21: Distribution of $m_{T2}(ll)$ in events with three leptons, b-tagged jet veto and $p_T^{miss} > 140$ GeV. The WW background is mimicked by the $WZ \rightarrow 3l1\nu$ process. The \vec{p}_T of a lepton from the Z boson decay with the same charge as the lepton from W boson decay is vectorially added to the total \vec{p}_T^{miss} . The new events must satisfy the baseline selection. The lower panel shows the ratio between data and total SM contribution in each bin. Only statistic uncertainties are considered.

Eventually, the simulation is found to describe the data well in the CRs. Based on the statistical precision of these CRs, a conservative uncertainty of 5, 10, 20 and 30% is taken for the bins $60 \leq m_{T2}(ll) < 80$ GeV, $80 \leq m_{T2}(ll) < 100$ GeV, $100 \leq m_{T2}(ll) < 120$ GeV, and $m_{T2}(ll) \geq 120$ GeV, respectively. These uncertainties are applied to top quark and WW production, and treated as uncorrelated between the two types of backgrounds.

Nonprompt leptons $m_{T2}(ll)$ modeling

Another potential source of mismodeling in the tails of $m_{T2}(ll)$ distributions arises from nonprompt leptons originating, for instance, from semileptonic decays of B hadrons in b jets or from hadronic jets accidentally passing the lepton selection. The value of $m_{T2}(ll)$ in $t\bar{t}$, tW , and WW events with one nonprompt lepton replacing a

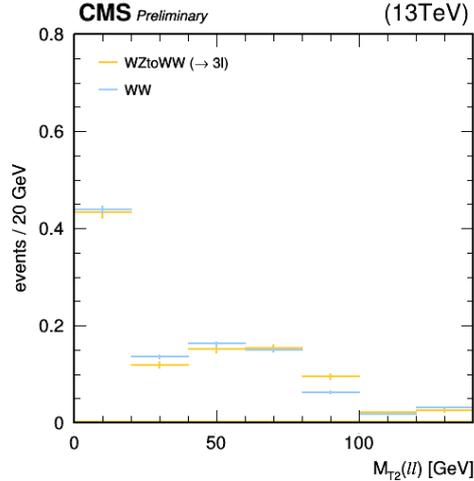


FIGURE 6.22: Comparison of $m_{T2}(ll)$ shape for nominal WW (blue) and mimicked WW (yellow) processes. Only statistical uncertainties have been considered.

prompt one failing the selection requirements will not be bound by the m_W endpoint. The contribution of these events is found to be less than 1% of the expected background across the different SRs. It becomes more relevant only at large values of $m_{T2}(ll)$ and p_T^{miss} , where it constitutes up to 20% of the $t\bar{t}$ background.

Since the nonprompt leptons rate is usually not very accurately reproduced by MC simulation, the modelling of this source of background is also studied. The shape of $m_{T2}(ll)$ is studied in events selected by removing the third lepton veto from the baseline selection. Since the two tight leptons have been already reconstructed, the third looser lepton is expected to be a good representative of a fake lepton candidate. The $m_{T2}(ll)$ value is then recomputed with a new lepton pair by randomly promoting the third lepton to the rank of one of the two tight leptons, emulating thus the $m_{T2}(ll)$ of events with fake leptons in the SRs. The \vec{p}_T of the excluded tight lepton is not added to the \vec{p}_T^{miss} since it is very likely that such loose lepton, failing the identification and isolation requirements, could have a reconstructed track and associated energy deposits in the calorimeter.

The results are shown in Figure 6.23 for events satisfying $100 \leq p_T^{miss} < 140$ GeV (left) and $p_T^{miss} > 140$ GeV (right). A b-tagged jet is also required to suppress the $WZ \rightarrow 3\ell 3\nu$ production. The $m_{T2}(ll)$ shape for events with the emulated fake leptons look correctly modeled in the simulations within the statistical uncertainties.

The modelling of the fake lepton rate is also studied for events with two leptons of same charge (SC) in the regions with $100 < p_T^{miss} < 140$ GeV and $p_T^{miss} > 140$ GeV. At least one b-tagged jet is also required to select a phase space similar to $t\bar{t}$ process. Figure 6.24 shows the result. The dominant contribution to this sample comes from $t\bar{t}$ dileptonic and semileptonic decays. An overall disagreement over the statistical uncertainties is observed between data and SM expectation in the $100 < p_T^{miss} < 140$ GeV region, while the disagreement is focused only on one $m_{T2}(ll)$ bin for events in the $p_T^{miss} > 140$ GeV region. No trend can be recognized in the bottom panels.

Finally, the difference between events with positive or negative SC lepton pairs has been also considered in the region of the analysis with $p_T^{miss} > 140$ GeV. Figure 6.25

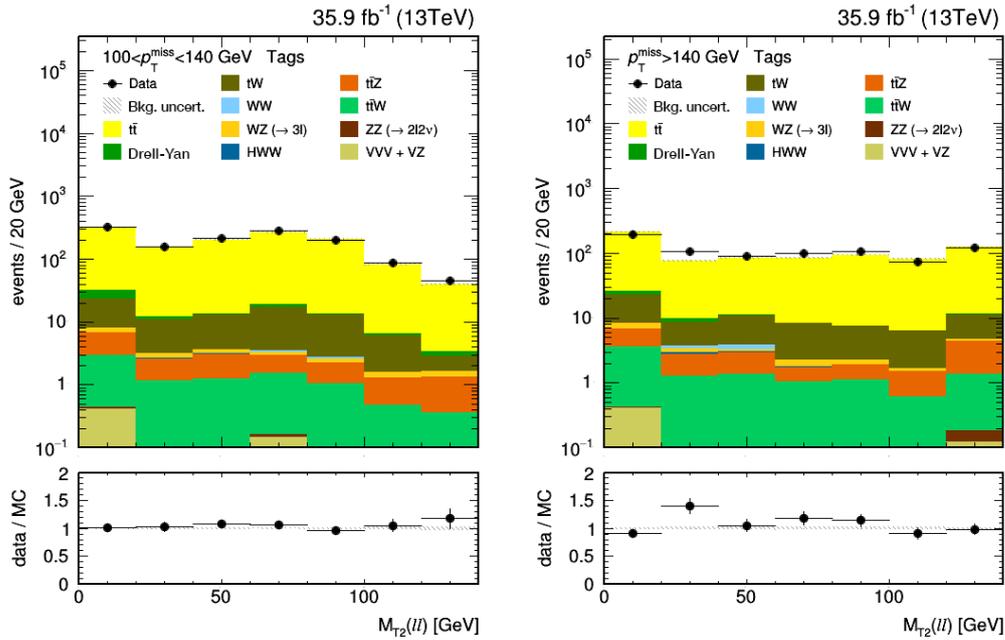


FIGURE 6.23: Distribution of $m_{T2}(ll)$ in events with two tight leptons and a third lepton passing the veto requirements but failing the tight definition. The $m_{T2}(ll)$ variable has been built by randomly substituting one of the first two tight leptons with the third looser one. A b-tagged jet is also required to suppress WZ production. Distributions are shown for data and simulated events with $100 \leq p_T^{miss} < 140$ GeV on the left, and with $p_T^{miss} \geq 140$ GeV on the right one. Only statistic uncertainties are considered.

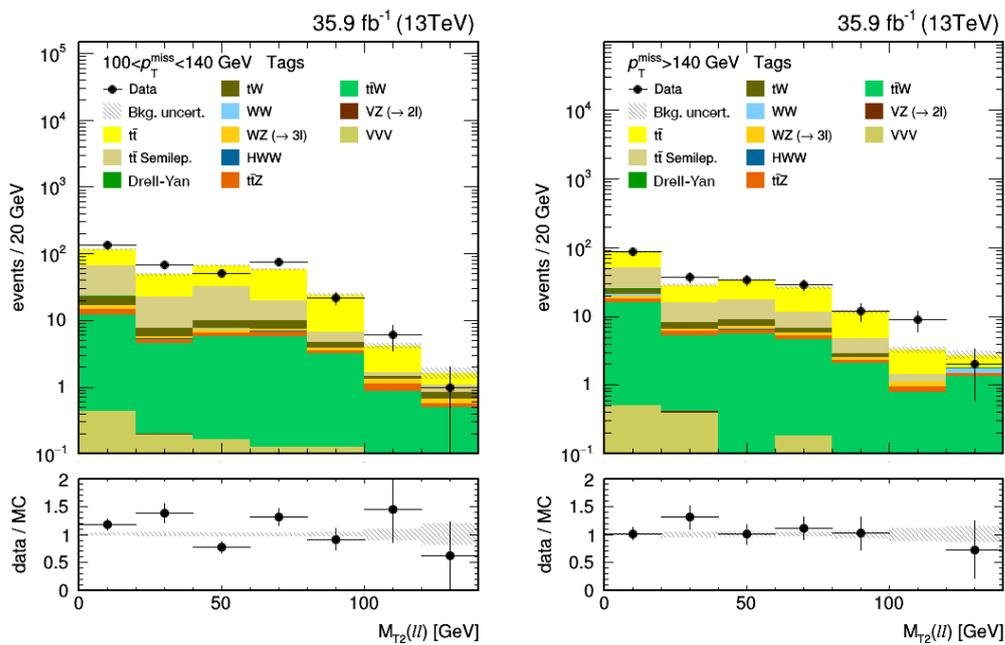


FIGURE 6.24: Distribution of $m_{T2}(ll)$ in events with same-charged lepton pairs and at least one b-tagged jet. Distributions are shown for data and simulated events with the same p_T^{miss} requirement as the validation region ($100 \leq p_T^{miss} < 140$ GeV) on the left plot, and with the same p_T^{miss} requirement as the signal regions ($p_T^{miss} \geq 140$ GeV) on the right one. Only statistic uncertainties are considered.

shows the $m_{T2}(ll)$ distribution for negative charged lepton pairs on the left and for the positive charged lepton pairs on the right.

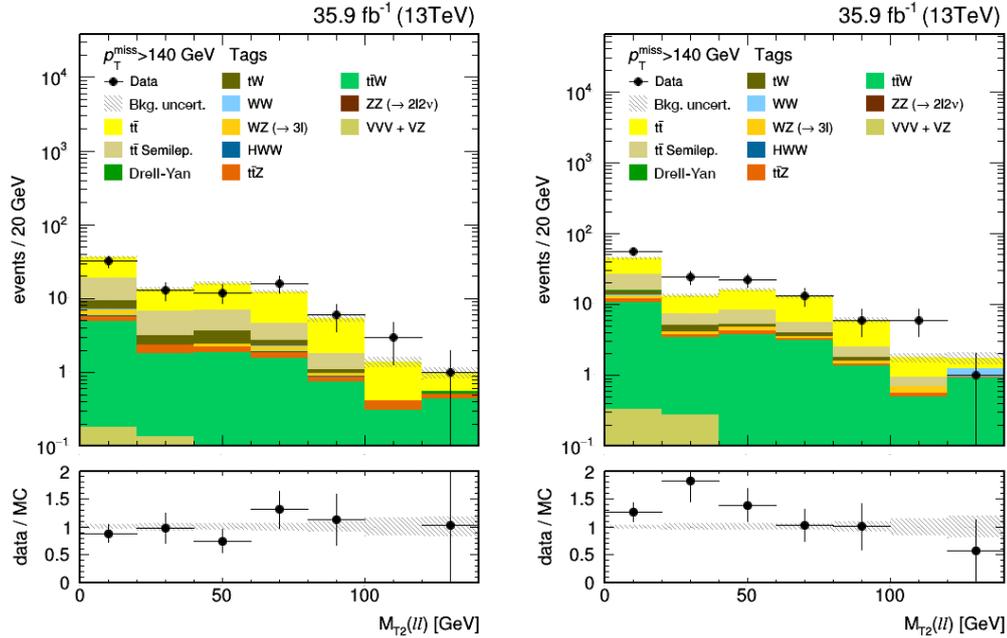


FIGURE 6.25: Distribution of $m_{T2}(ll)$ in events with same-sign lepton pairs with negative (left) or positive (right) charge, at least one b-tagged jet and $p_T^{miss} > 140$ GeV. Only statistic uncertainties are considered.

Based on the observed agreement between data and simulation, a correction factor of 1.08 ± 0.21 is derived for the nonprompt lepton rate in simulation. The uncertainty is set to cover the difference between events with positive and negative lepton pairs.

6.4.2 Estimation of $t\bar{t}Z$ background

The production of $t\bar{t}Z$ events where two W bosons decay leptonically and the Z boson decays into neutrinos leads to final states with the same experimental signature as the signal events and with no natural endpoint for reconstructed $m_{T2}(ll)$ distribution, due to the additional contribution of the neutrinos from the Z boson decay to the \vec{p}_T^{miss} . After the event selection, in SRs with at least one b tagged jet and OC same lepton flavour pair, this process results dominant on the $m_{T2}(ll)$ tails for values of $m_{T2}(ll) > 140$ GeV, see Figure 6.16.

The normalization of this background is validated in events with three tight leptons, $p_T^{miss} > 140$ GeV, and at least two jets with $p_T > 20$ GeV and $|\eta| < 2.4$, of which at least one is tagged as b jet. At least one pair of OC same flavour leptons with an invariant mass not further than 10 GeV from the Z boson mass is also required. Table 6.5 summarizes the selection. This control region increases the acceptance of $t\bar{t}Z$ events with the Z boson decaying into two leptons and the $t\bar{t}$ process decaying semileptonically, where one of the produced W bosons decays hadronically and the other leptonically. The observed and expected yields after applying the selection requirements are shown in Figure 6.26.

Object	Selection
leptons (e, μ) flavour and sign	Three tight leptons with $p_T > 25, 20$ and 20 GeV at least one OC SF pair
$ m_{\ell\ell} - m_Z $	< 10 GeV
p_T^{miss}	≥ 140 GeV
N_{jets}	≥ 2
$N_{b\ jets}$	≥ 1

TABLE 6.5: Control region definition for the validation of $t\bar{t}Z$ background normalization.

The main contribution in this region comes from $t\bar{t}Z$ events, followed by the WZ contribution. The rest of yields due to other SM process are negligible. A normalization scale factor is measured comparing the observed and predicted numbers in Figure 6.26 following the Equation 6.5.

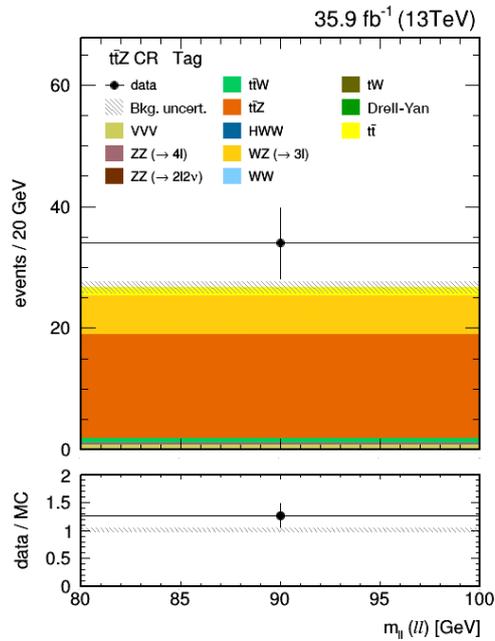


FIGURE 6.26: Observed and expected yields of $m_{\ell\ell}$ distribution in $t\bar{t}Z$ control region. The lower panel shows the ratio between data and total SM contribution in each bin. Only statistic uncertainties are shown.

The estimation provides a normalization scale factor of 1.44 ± 0.36 for $t\bar{t}Z$ production. The error includes the statistical uncertainty both for data and SM simulated events and the systematic uncertainties on the number of expected events from the residual processes in the CRs.

6.4.3 Estimation of WZ background

Events from the WZ production can pass the signal region selection when both bosons decay leptonically and one of the three leptons fails the veto lepton requirement. Therefore, if the other two leptons pass the tight requirement and baseline

selection the event is accepted. The modeling of this source of background is tested again in a three tight lepton CR, requiring $p_T^{miss} \geq 140$ GeV, but no b-tagged jets present in the event, see Table 6.6. The result is shown through the $m_{T2}(ll)$ distribution in Figure 6.27. This region is highly dominated by $WZ \rightarrow 3\ell 1\nu$ events, other backgrounds such as $ZZ \rightarrow 4\ell$ or $t\bar{t}V \rightarrow 3\ell + X$ with $V \sim Z, W$ give a negligible contribution.

An overall agreement is observed between data and simulation, and a normalization scale factor can be derived from this distribution. Following the Equation 6.5 with the corresponding processes for this estimation, a normalization scale factor of 0.97 ± 0.09 is derived for the WZ background, where the uncertainty includes the statistical uncertainty both for data and SM simulated events and the systematic uncertainties on the number of expected events from the residual processes in the CRs.

Object	Selection
leptons (e, μ) flavour and sign $ m_{\ell\ell} - m_Z $ p_T^{miss} $N_{b \text{ jets}}$	Three tight leptons with $p_T > 25, 20$ and 20 GeV at least one OC SF pair < 10 GeV ≥ 140 GeV 0

TABLE 6.6: Control region definition for the validation of WZ background normalization.

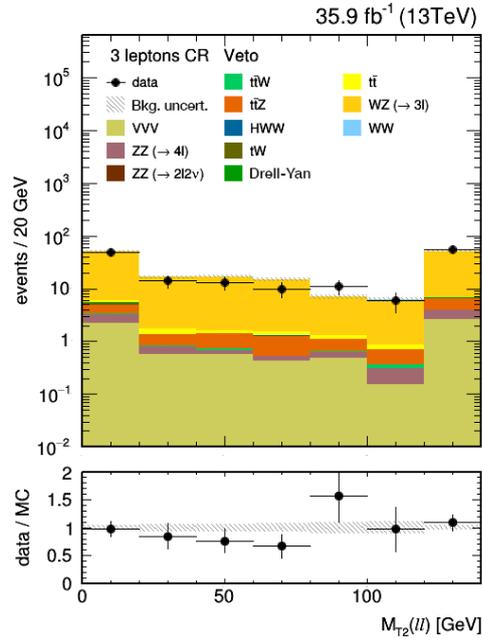


FIGURE 6.27: Observed and simulated $m_{T2}(ll)$ distribution in data and simulations for events in the WZ control region. The lower panel shows the ratio between data and total SM contribution in each bin. Only statistic uncertainties are shown.

6.4.4 Estimation of ZZ background

The remaining backgrounds from diboson production contributing to the SRs selection are the $WZ \rightarrow 2\ell 2q$ and ZZ processes where a Z boson decays into two leptons ($Z \rightarrow 2\ell$) and the other decays into two quarks ($Z \rightarrow q\bar{q}$) or neutrinos ($Z \rightarrow \nu\nu$). After the requirement of $p_T^{miss} > 140$ GeV, the $ZZ \rightarrow 2\ell 2\nu$ is the dominant background process.

The contribution of this process is significant in the tail of $m_{T2}(l\bar{l})$ distribution for SF lepton pair in the first two bins of p_T^{miss} when a veto on the presence of b-tagged jets has been required, being larger in 0-jet SRs.

Due to the similarities with the signal signature (two OC leptons and high p_T^{miss}), finding a control region orthogonal to the SRs for estimating this background is not easy. Therefore, this contribution is studied by mimicking the $ZZ \rightarrow 2\ell 2\nu$ production via $ZZ \rightarrow 4\ell$ events, where the p_T of one of the reconstructed Z bosons (randomly chosen between the pairs of leptons satisfying $|m_{\ell\ell} - m_Z| < 15$ GeV) is added to the total \vec{p}_T^{miss} of the event. The $m_{T2}(l\bar{l})$ variable is recomputed using the remaining lepton pair if it satisfies the η and p_T baseline requirements and the new \vec{p}_T^{miss} .

The events are selected by requiring four leptons (e, μ), with one lepton allowed to pass the looser selection for veto lepton requirement in order to increase the acceptance for ZZ production while maintaining very high purity in the control region. The events are retained if the four leptons can be arranged into two pairs of OC SF lepton, both with an invariant mass within 30 GeV of Z boson mass, and at least one within 15 GeV.

In addition to the MC samples of ZZ simulated events listed in Table A.2 (Appendix A) this study uses a set of simulated samples of $ZZ \rightarrow 4\ell$ events from different production processes, as summarized in Table 6.7.

Sample	σ [pb]	Events/M
/ZZTo4L_13TeV_powheg_pythia8	1.256	6.7
/GluGluToContinToZZTo4e_13TeV_MCFM701_pythia8	0.001586	1.0
/GluGluToContinToZZTo4mu_13TeV_MCFM701_pythia8	0.001586	1.0
/GluGluToContinToZZTo4tau_13TeV_MCFM701_pythia8	0.001586	0.5
/GluGluToContinToZZTo2e2tau_13TeV_MCFM701_pythia8	0.003194	0.5
/GluGluToContinToZZTo2e2mu_13TeV_MCFM701_pythia8	0.003194	1.5
/GluGluToContinToZZTo2mu2tau_13TeV_MCFM701_pythia8	0.003194	0.5
/GluGluHToZZTo4L_M125_13TeV_powheg2_JHUgenV6_pythia8	0.012	1.0
/VBF_HToZZTo4L_M125_13TeV_powheg2_JHUgenV6_pythia8	0.001	0.5

TABLE 6.7: Simulated samples of $ZZ \rightarrow 4\ell$ production events used in the analysis for $ZZ \rightarrow 2\ell 2\nu$ estimation. The cross section assigned to each process and the number of processed events are also shown.

The cross section for the $q\bar{q} \rightarrow ZZ \rightarrow 4\ell$ and $gg/VBF \rightarrow H \rightarrow ZZ \rightarrow 4\ell'$ processes have been computed at NNLO, while the $gg \rightarrow ZZ \rightarrow 4\ell/2\ell 2\ell'$ production at LO. The nominal samples used in the analysis: $gg \rightarrow ZZ \rightarrow 2e2\nu/2\mu 2\nu$ and $q\bar{q} \rightarrow ZZ \rightarrow 2\ell 2\nu/2\ell 2q$ are generated at LO and NLO respectively. The difference in the QCD perturbative precision may lead to differences in the modelling that can be propagated to the analysis. In order to make compatible the predictions a set of K factors has been used to correct the samples (as mentioned in Section 5.3). The $q\bar{q} \rightarrow ZZ \rightarrow 2\ell 2\nu$ production (at NLO) is reweighted by NNLO/NLO K factors,

which are available in three different sets as a function of total mass of the generated ZZ system, the total transverse momentum and the azimuthal angle between each generated Z boson. The $gg \rightarrow ZZ \rightarrow 4\ell/2\ell 2\ell'$ and $gg \rightarrow ZZ \rightarrow 2e2\nu/2\mu 2\nu$ samples (both at LO) are reweighted by NNLO/LO K factors which are available as a function of the total ZZ system mass.

In the case of NNLO/NLO K factors, the effect of each set in the modelling of the $q\bar{q} \rightarrow ZZ \rightarrow 2\ell 2\nu$ process has been compared. The general acceptance of ZZ process after $p_T^{miss} > 140$ GeV and no b-tagged jet requirements is shown in Table 6.8. The number of expected events after reweighting the sample by $m(ZZ)$ and $\Delta\phi(Z, Z)$ K factors dependent is very similar, while an excess in the simulation prediction with respect to data is observed when the processes are reweighted by $p_T(ZZ)$ K factors dependent.

	ZZ-mass k-factors	ZZ- p_T k-factors	ZZ- $\Delta\phi$ k-factors
Data	80	80	80
ZZ	72.2	88.5	71.3
Other backgrounds	4.0	4.0	4.0

TABLE 6.8: Observed and expected events of ZZ process in the 4ℓ CR where the one of Z boson reconstructed p_T has been added to the total \vec{p}_T^{miss} . The requirement of $p_T^{miss} > 140$ GeV and no b-tagged jets is then applied.

The $m_{T2}(ll)$ shape normalized to unity is compared in Figure 6.28 for $ZZ \rightarrow 2\ell 2\nu$ events requiring $p_T^{miss} > 140$ GeV and no b-tagged jets. The four distributions correspond to the nominal shape and the resulting shapes after applying independently each of the three different sets of K factors. The ratio between every reweighted distribution and the nominal one is shown in the bottom plot. In the case of $m(ZZ)$ and $p_T(ZZ)$ dependent sets, small differences can be appreciated mostly in first bins of $m_{T2}(ll)$ can be appreciated. They become bigger for $\Delta\phi(Z, Z)$ sets, even in last bins.

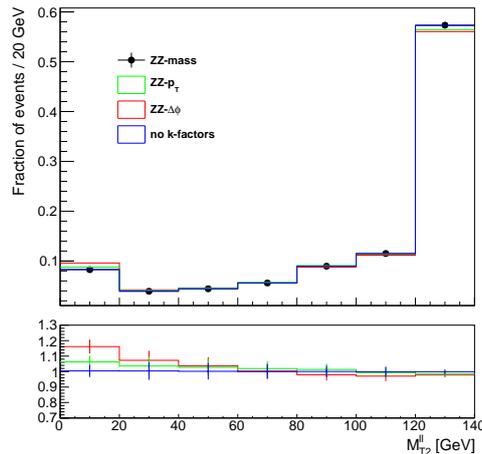


FIGURE 6.28: Normalized shape of the $m_{T2}(ll)$ distribution in $ZZ \rightarrow 2\ell 2\nu$ events with $p_T^{miss} > 140$ GeV obtained with the different K factor sets and with no K factors applied. On the bottom plot, the ratio between the nominal and each of the reweighted distributions.

Consistently to these results, the normalization of ZZ background in the signal regions is carried out using the set of K factors as a function of the ZZ mass to reweigh the $q\bar{q} \rightarrow ZZ \rightarrow 2\ell 2\nu$ sample. A general normalization scale factor of 1.05 ± 0.12 is obtained where the uncertainty is dominated by the data statistics.

To take into account the bias that introduces the choice of one particular set, a shape uncertainty has been defined as the difference between the normalized shape of $m_{T2}(ll)$ before and after applying $\Delta\phi(Z, Z)$ K factors. Since this specific set presents the largest variation with respect to the nominal shape, the uncertainty covers the three options. The relative variations range from 16% for $m_{T2}(ll) < 20$ GeV to about 2% for $m_{T2}(ll) > 120$ GeV.

Finally, since the chargino search uses separate SRs for events with and without jets, the corrected ZZ modelling has been studied as a function of the number of jets for the mimicked events, those obtained after adding the p_T of one of the reconstructed Z bosons into the total \vec{p}_T^{miss} and applying $p_T^{miss} > 140$ GeV and no b-tagged jets selection. As result, Figure 6.29 shows the observed and expected jet multiplicity distribution. Since the data to simulation agreement suggests a significant trend, it makes unable to employ the previous estimations in dedicated chargino regions. Hence, two different normalization scale factors have been measured depending on the presence of 0 jet or at least one jet in the event: 0.74 ± 0.19 and 1.21 ± 0.17 respectively.

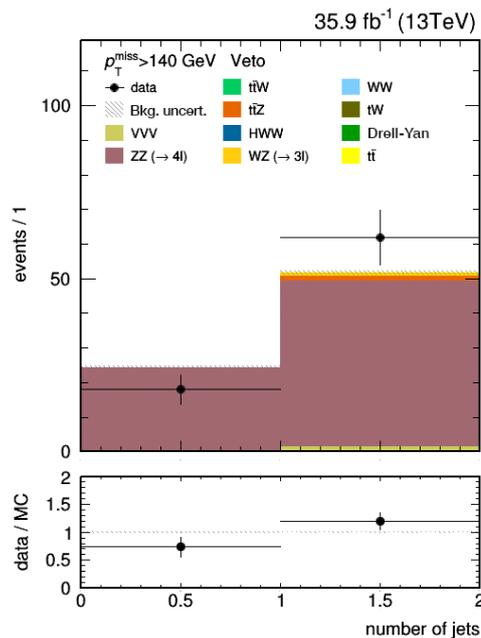


FIGURE 6.29: Observed and expected jet multiplicity distribution for ZZ process corrected by $m(ZZ)$ K factors set. The p_T of one of the reconstructed Z boson has been added to the total \vec{p}_T^{miss} . Events with 0 or at least 1 no b-tagged jet have been selected after the $p_T^{miss} > 140$ GeV requirement. The lower panel shows the ratio between data and total SM contribution in each bin. Only statistic uncertainties are shown.

A summary of the scale factors derived in Sections 6.4.2, 6.4.3 and 6.4.4 is given in Table 6.9. For all the quoted scale factor values, uncertainties include the statistical

uncertainties on data and simulated events and the systematic uncertainties on the number of expected events from the residual processes in the CRs.

Process	Scale factors		
	$N_{jets} = 0$ (a)	$N_{jets} > 0$ (a)	$N_{jets} \geq 0$ (b)
$t\bar{t}Z$	1.44 ± 0.36	1.44 ± 0.36	1.44 ± 0.36
WZ	0.97 ± 0.09	0.97 ± 0.09	0.97 ± 0.09
ZZ	0.74 ± 0.19	1.21 ± 0.17	1.05 ± 0.12

TABLE 6.9: Summary of the normalization scale factors for $t\bar{t}Z$, WZ , and ZZ backgrounds in the SRs used for the chargino (a) and top squark (b) searches. Uncertainties include the statistical uncertainties of data and simulated event samples, and the systematic uncertainties on the number of expected events from the residual processes in the CRs.

6.4.5 Estimation of Z boson background

The Z boson production at LHC is dominated by the Drell-Yan process which is usually accompanied by jets. In the dileptonic phase space of this analysis this background is mainly concentrated at low p_T^{miss} values. After applying the complete baseline selection, the events from Drell-Yan are highly suppressed by the $p_T^{miss} > 140$ GeV requirement. However, its contribution has been found non-negligible in signal regions without b-tagged jets, same flavour and low p_T^{miss} bins, mainly concentrated on first and last bins of the $m_{T2}(ll)$ distribution.

The event rate and the modelling of the $m_{T2}(ll)$ shape is studied for this process in a Z boson enriched region defined by events with two OC SF leptons that satisfy $|m_{\ell\ell} - m_Z| < 15$ GeV, no additional leptons, and no b-tagged jets are required. The last condition rejects as much as possible the top background.

Since even in this control region after $p_T^{miss} > 140$ GeV requirement the Drell-Yan events are not dominant in the background, a lower p_T^{miss} range ($100 < p_T^{miss} < 140$ GeV), as close as possible to signal region selection where Drell-Yan events are still dominant SM background is also considered. Figure 6.30 shows the $m_{T2}(ll)$ distribution in these two ranges of p_T^{miss} . At high p_T^{miss} selection (right), where the Drell-Yan relative contribution is about 13%, the observed distribution is well modelled by the MC simulation, however, after lower p_T^{miss} requirement is applied (left), the contribution of Drell-Yan increases to about 77% presenting a trend in the $m_{T2}(ll)$ simulation modelling.

In order to study first the modelling of $m_{T2}(ll)$ distribution for Drell-Yan simulation, a normalization scale factor is first measured in events with $100 < p_T^{miss} < 140$ GeV following Equation 6.5. The obtained value ($SF_{norm} = 0.68$) is then applied on Drell-Yan events to match the data. In Figure 6.31 left, the observed and simulated $m_{T2}(ll)$ distribution after renormalization of Drell-Yan events is shown. The ratio between data and simulation is used to derive bin-by-bin weights successively employed for the $m_{T2}(ll)$ shape correction in Drell-Yan events with $p_T^{miss} > 140$ GeV. The correction ranges from a few percent at low $m_{T2}(ll)$ to about 50% for $m_{T2}(ll) > 100$ GeV. Figure 6.31 right shows the region at high p_T^{miss} after applying the shape correction weights. An improvement on data to simulation agreement is observed.

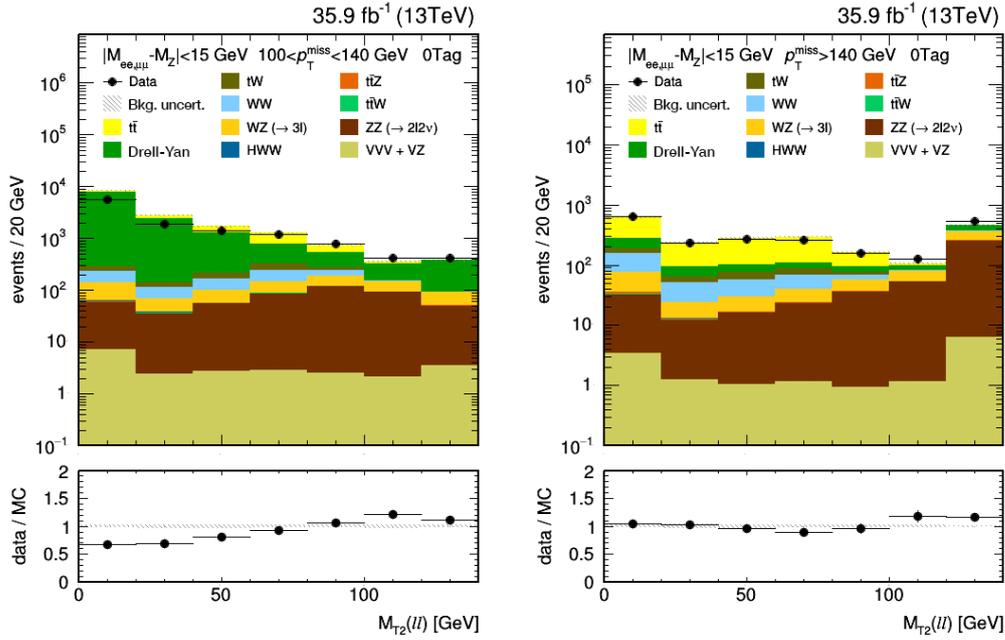


FIGURE 6.30: Observed and expected distribution of $m_{T2}(ll)$ observed in data and simulated events with two OC SF leptons satisfying $|m_{\ell\ell} - m_Z| < 15$ GeV and $100 < p_T^{miss} < 140$ GeV (left) $p_T^{miss} > 140$ GeV (right). The lower panel shows the ratio between data and total SM contribution in each bin. Only statistic uncertainties are shown.

The estimation of Drell-Yan contribution in the SRs is taken from simulation in the control region at high p_T^{miss} after applying the shape correction weights. The difference between the derived weight and the unity in each bin is taken as shape uncertainty. An overall normalization uncertainty of 32% is established according to the observed disagreement between data and simulated number of events with $100 < p_T^{miss} < 140$ GeV. The total uncertainty on the yield estimation bin by bin is the root square of the quadratic sum of the two uncertainties.

Finally, the predictions for Drell-Yan events with no jets are tested in the same Z boson enriched region for the two p_T^{miss} ranges. This process is subdominant background in the SRs with no jets and populates the first bins of $m_{T2}(ll)$ distribution ($m_{T2}(ll) < 20$ GeV).

The number of simulated events with $p_T^{miss} > 140$ GeV has been found largely underestimated with respect to the observed yields. On the other hand, events with $100 < p_T^{miss} < 140$ GeV exhibit a rather different trend not compatible with the former excess. The disagreement exhibited between data and simulation suggests a high dependence on the p_T^{miss} value, avoiding any inference estimation between regions.

Therefore, due to the difficulties in the modeling of MC simulation in this phase space, together with the large statistical uncertainties of the predictions, a conservative normalization uncertainty of 100% is applied on its contribution in those SRs with no jets. Because Drell-Yan production is a subdominant background in the SRs with no jets, this uncertainty has a negligible impact on the expected sensitivity for signal production.

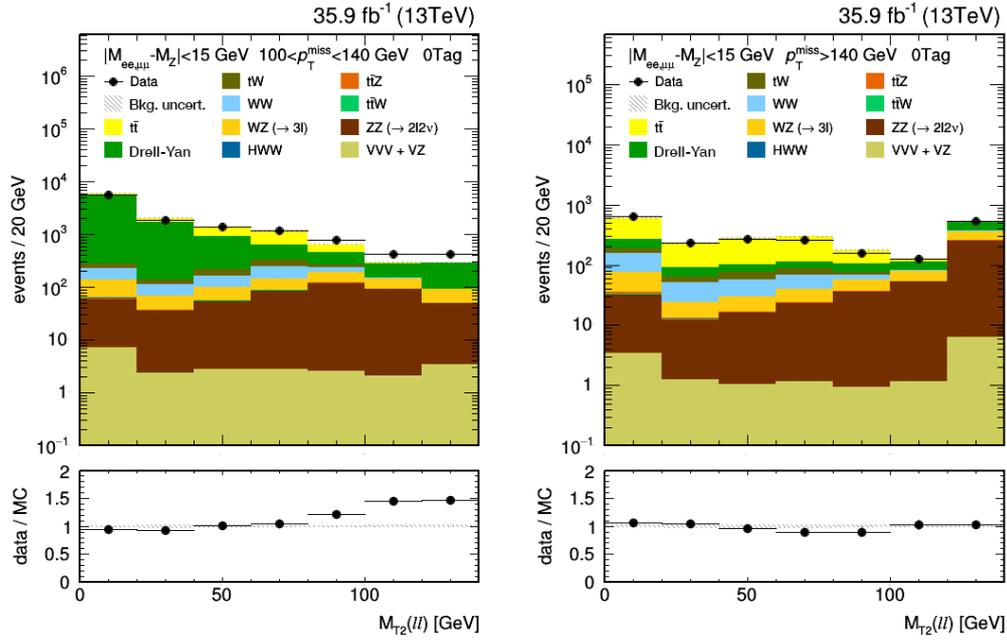


FIGURE 6.31: Observed and expected $m_{T2}(ll)$ distribution for events with $|m_{\ell\ell} - m_Z| < 15 \text{ GeV}$ and $100 < p_T^{\text{miss}} < 140 \text{ GeV}$ after applying the Drell-Yan event renormalization (left) and events with $|m_{\ell\ell} - m_Z| < 15 \text{ GeV}$ and $p_T^{\text{miss}} > 140 \text{ GeV}$ after applying the derived shape corrections of Drell-Yan modeling (right). The lower panel shows the ratio between data and total SM contribution in each bin. Only statistic uncertainties are shown.

6.5 Systematic uncertainties

The uncertainty on a measurement is associated to the nature of the measured system and the apparatus of measurement. In high energy physics, when we analyze data from pp collisions, two types of uncertainty sources are considered: the statistical uncertainty and the systematic uncertainty.

The statistical uncertainty is related to the inherent variability of sampling a random variable with a finite number of events. The precision on the measurement is limited by its uncertainty that in this case depends on the size of the sample, decreasing just as the sample size increases. Typical sources of statistical uncertainty are random fluctuations from noise in the detector, tight selections or a finite number of generated events.

The systematic uncertainty is related to the error introduced by the methods of measurement. In particular, data analyses in physics are subject to multiple sources of systematic uncertainty. These must be evaluated, computed and finally quoted after the statistical error. However, there is not a clear recipe for the detection and estimation of systematic uncertainties. Even so, it can be taken a pragmatic definition as a guide for their determination:

"Systematic uncertainties are measurement errors which are not due to statistical fluctuations in real or simulated data samples." [167]

With this definition, the general sources of systematic errors and biases that can be contemplated in this analysis are:

1. Detector resolution and acceptance or trigger efficiency.
2. Erroneous detector calibrations.
3. Reconstruction and identification of physics objects.
4. Uncertainties in the underlying theoretical aspects of the simulation.
5. Uncertainties on input parameters as luminosity or cross section.
6. Uncertainty on the estimation of poorly knowledge of backgrounds.
7. Uncertainties on the correction factors applied to take into account observed disagreement between data and simulation.

The following lines explain the sources of uncertainty considered and their estimate.

Statistical uncertainty

The number of events in each SR and in each bin of the $m_{T2}(ll)$ distribution is used to extract the signal strength through the ML fit. The effect of the simulated data samples sizes on the modelling of the $m_{T2}(ll)$ distributions is taken into account by treating the statistical uncertainty in each bin for each process as an additional uncorrelated uncertainty.

Systematic uncertainties

Several sources of systematic uncertainties that affect both the normalization and the $m_{T2}(ll)$ shape of the SM background and signal events are considered in this analysis.

- The overall uncertainty in the integrated luminosity is estimated to be 2.5% [100].
- The uncertainty on the measured trigger efficiency is 2% .
- The uncertainty on the pileup modeling of SM samples is obtained by reweighting by the pileup distribution in data after varying the cross section of inelastic proton-proton collision within its theoretical uncertainty ($\pm 5\%$). The change in the acceptance in simulated events after the full selection is taken as the total uncertainty.
- Lepton identification and isolation efficiencies are corrected by data-to-simulation scale factors measured in $Z \rightarrow \ell\ell$ events. The corresponding uncertainties are typically smaller than 3% per lepton.
- The jet energy scale is varied by its uncertainty [129], and the changes are propagated to all the related observables in the event.
- The energy scale of the low- p_T particles that are not clustered in jets, or unclustered energy, is varied by its uncertainty, and the changes are propagated to the p_T^{miss} .
- The efficiencies and misidentification rates of the b-jet identification algorithms are also corrected by data-to-simulation scale factors measured in inclusive jet and $t\bar{t}$ events [132]. The respective uncertainties range between 1 and 6%, depending on the p_T and η of the jets.

The estimates of the SM backgrounds are also affected by specific uncertainties in the modeling of different processes.

- A **background normalization uncertainty** is applied for each background separately. The normalization of the $t\bar{t}$, tW , and WW processes is determined by the ML fit, as described in Section 7.1. A common normalization parameter is assigned for $t\bar{t}$ and tW as top background and another for WW production. No explicit normalization uncertainty is defined for $t\bar{t}$ and WW , while a 10% uncertainty is set for the tW process to take into account its relative normalization with respect to $t\bar{t}$ production as well as any interference effect between them. The uncertainties applied to $t\bar{t}Z$ (25%), WZ (9%) and ZZ (26% in the SRs with 0 jets, 14% in the SRs with at least 1 jet, and 11% in the rest of the SRs) correspond to the scale factors obtained in Sections 6.4.2, 6.4.3 and 6.4.4, respectively. Minor backgrounds ($t\bar{t}W$, HWW , triboson production) are assigned a conservative uncertainty of 50%. Finally, Drell-Yan events have a 100% of normalization uncertainty in the SR with no jets and 32% in all other SRs, as explained in Section 6.4.5.
- The modeling of the yields of events with no jets has been explicitly studied for ZZ and Drell-Yan production in Sections 6.4.4 and 6.4.5 respectively. For other SM processes, a related uncertainty has been introduced by adding two free parameters in the ML fit, scaling respectively the rate of events with no jets for diboson and b-enriched ($t\bar{t}$, tW , $t\bar{t}Z$, and $t\bar{t}W$) backgrounds. The total number of expected events without b-tagged jets is constrained to remain invariant, so that only a migration of events between the SRs with and without jets is allowed. Details are given in Appendix B.
- The modeling of $m_{T2}(ll)$ shapes in events of backgrounds with an endpoint at the m_W ($t\bar{t}$, tW , and WW) has been studied in Section 6.4.1: an uncertainty of 5, 10, 20 and 30% is assigned for the last four $m_{T2}(ll)$ bins. These are treated as uncorrelated between signal regions and between the top ($t\bar{t}$, tW) and WW backgrounds.
- The choice of the set of NNLO/NLO K factors applied to the $q\bar{q} \rightarrow ZZ$ events affects the modelling of the $m_{T2}(ll)$ shape for the ZZ background (as described in Section 6.4.4). Relative variations range from 16% for $m_{T2}(ll) < 20$ GeV to about 2% for $m_{T2}(ll) > 120$ GeV and are taken as uncertainties.
- The $m_{T2}(ll)$ distribution in Drell-Yan events has been corrected by scale factors derived in bins of $m_{T2}(ll)$ in the validation region $100 < p_T^{miss} < 140$ GeV, as discussed in Section 6.4.5. The correction ranges from few percent at low $m_{T2}(ll)$ to about the 50% for $m_{T2}(ll) > 100$ GeV. The full size of the correction in each bin is taken as uncertainty.
- The weight of nonprompt lepton events in simulated samples is varied by the 19% uncertainty in the correction factor derived in events with two same-charge leptons, as described in Section 6.4.1.
- The spectrum of top quark p_T in $t\bar{t}$ events has been observed to be softer in data than in simulated events [168–170]. An uncertainty is derived from the observed variations when reweighting the $t\bar{t}$ events to the p_T distribution observed in data.
- Uncertainties in the renormalization and factorization scales, and PDFs are propagated by taking the largest changes in the acceptance when independently doubling and halving the renormalization and factorization scales, and when varying the choice of PDFs between the NNPDF3.0 replicas. The PDF

uncertainties are not considered for signal models as they are found to be redundant, once the uncertainty in the ISR modeling is included.

Finally, additional uncertainties in the modelling of signal events are taken to take into account, mostly related to the performance of the event reconstruction in FastSim.

- The lepton identification efficiency in events simulated with FastSim is corrected by a scale factor relative to the lepton identification efficiency in generated events using full detector simulation. The additional uncertainty is estimated to be 2%.
- The analogous uncertainty of correction factors in the b-tagging efficiency in FastSim samples is applied. The uncertainty ranges between 0.2-0.5%.
- The p_T^{miss} modeling in FastSim events is studied by comparing the acceptances computed using the p_T^{miss} at generator level and after the event reconstruction. Since the average of two is taken as central value for the acceptance, half of their difference is taken as an uncertainty, fully correlated among bins.
- An uncertainty in the modeling of pileup events in FastSim signal samples is derived by studying the dependence of the acceptance on the multiplicity of primary vertices reconstructed in the event. This uncertainty varies from 0 to 10% across the SRs and the $m_{T2}(ll)$ bins.
- Simulated signal events are reweighted to improve the modeling of the ISR, as described in Section 5.3. Uncertainties on the reweighting procedure are derived from closure tests. For chargino models, the deviation from unity is taken as the systematic uncertainty in the p_T^{ISR} reweighting factors. For top squark models, half of the deviation from unity in the N_{jet}^{ISR} factors is taken.

Tables 6.10 and 6.11 summarize the effect of the systematic uncertainties on the predicted yields for SM processes and for two reference signal points, respectively. The range of the systematic uncertainties across the different signal regions is given in Appendix C.

Source of uncertainty	SM processes	
	Change in yields	Change in $m_{T2}(ll)$ shape
Integrated luminosity	2.5%	—
Trigger	2%	—
Lepton ident./isolation	4–5%	<1%
Jet energy scale	1–6%	3–15%
Unclustered energy	1–2%	2–16%
b tagging	<3%	<2%
Renorm./fact. scales	1–10%	1–6%
PDFs	1–5%	2–8%
$t\bar{t}Z$ normalization	<1%	<9%
WZ normalization	<1%	<1%
ZZ normalization	<1%	<5%
Drell-Yan normalization	<4%	1–11%
$m_{T2}(ll)$ shape (top quark)	—	4–18%
$m_{T2}(ll)$ shape (WW)	—	1–15%
ZZ K factors	—	<3%
$m_{T2}(ll)$ shape (Drell-Yan)	—	1–13%
Nonprompt leptons	<1%	<4%
$t\bar{t}$ p_T reweighting	1–4%	1–8%

TABLE 6.10: Sizes of systematic uncertainties in the predicted yields for SM processes. The first column shows the range of the uncertainties in the global background normalization across the different SRs. The second column quantifies the effect on the $m_{T2}(ll)$ shape. This is computed by taking the maximum variation across the $m_{T2}(ll)$ bins (after renormalizing for the global change of all the distribution) in each SR. The range of this variation across the SRs is given.

Source of uncertainty	$\tilde{\chi}_1^\pm \rightarrow \tilde{l}\nu$ ($\tilde{\nu}l$) $\rightarrow l\nu\tilde{\chi}_1^0$		$\tilde{t}_1 \rightarrow t\tilde{\chi}_1^0$	
	$m_{\tilde{\chi}_1^\pm}=500$ GeV, $m_{\tilde{\chi}_1^0}=200$ GeV	Yields	$m_{\tilde{t}_1}=350$ GeV, $m_{\tilde{\chi}_1^0}=225$ GeV	Yields
Integrated luminosity	2.5%	—	2.5%	—
Trigger	2%	—	2%	—
Lepton ident./isolation	4–5%	<1%	4–5%	<1%
Jet energy scale	1–3%	3–11%	1–4%	2–14%
Unclustered energy	1–2%	8–13%	1–2%	2–7%
b tagging	<1%	<1%	1–3%	<1%
Renorm./fact. scales	1–3%	1–3%	1–3%	1–3%
Lept. id./iso. (FastSim)	4%	<1%	4%	<1%
b tagging (FastSim)	<1%	<1%	<1%	<1%
\vec{p}_T^{miss} (FastSim)	1–4%	7–28%	1–6%	6–20%
Pileup (FastSim)	1–6%	4–9%	2–4%	2–14%
ISR reweighting	1–2%	1–6%	2–8%	1–6%

TABLE 6.11: Same as in Table 6.10 for a representative signal point of chargino and top squark pair production.

Chapter 7

Results and interpretation

The previous chapters have set the basis for this search emphasizing the elements and procedures used for signal extraction. In this chapter a review on the performed maximum likelihood fit to data is done in Section 7.1, the outcomes on the simulated $m_{T2}(ll)$ distributions are presented in Section 7.2 and Section 7.3 ends with the further interpretation of the results in terms of the considered SMS.

7.1 Signal extraction

The supersymmetric signal is extracted using a simultaneous binned ML fit to the $m_{T2}(ll)$ distribution in all SRs. Assuming that the number of events in each bin follows a Poisson distribution, the total likelihood function $L(\text{data} | \mu, \theta)$ is expressed in terms of a product of Poisson probabilities over all the bins, channels and SRs:

$$L(\text{data} | \mu, \theta) = \prod \text{Poisson}(\text{data} | \mu \cdot s(\theta) + b(\theta)) \cdot p(\theta | \tilde{\theta}). \quad (7.1)$$

The parameter μ is known as the signal strength modifier and its value is taken as the change factor of the signal cross section given the observed data. It can be expressed as the ratio between the estimated number of signal events and the number of signal events predicted by the model (σ / σ_{nom}). The parameter θ represents a collection of nuisance parameters with a probability density function $p(\theta)$ associated to a default value $\tilde{\theta}$, that reflect our best estimate of the nuisance value. The function $\text{Poisson}(\text{data} | \mu s + b)$ stands for the product of Poisson probabilities to observe n_i events in N bins i when the mean expected number in each bin is $\nu = \mu s_i + b_i$,

$$\prod_i^N \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)}. \quad (7.2)$$

Prior to fit the data, both the signal yield (s) and background yield (b) are subject to different sources of uncertainty. Those sources are handle by introducing the nuisance parameters θ , that makes the signal and background expectations become functions of these nuisance parameters ($s(\theta)$, $b(\theta)$). The differences between SM and beyond the SM contribution on the predicted $m_{T2}(ll)$ shape will be then exploited to set an upper limit on the signal cross section for each model.

In addition to the parameters of interest which are free in the fit, such as the rate (cross section) of the signal process, the nuisance parameters, whose values are not

taken as known a priori, must also be fitted from data. The choice of the probability distributions $p(\theta|\bar{\theta})$, that models our prior knowledge about them, includes additional constraints in the likelihood functions. This method to parametrize the systematic effects introduces an additional flexibility that will result, as it should be, in a certain loss in sensitivity.

In this analysis, the uncertainties due to signal and background normalization, luminosity and trigger efficiency are included through nuisance parameters with log-normal prior distributions. The uncertainties in the shape of the $m_{T2}(ll)$ distributions are included with Gaussian prior distributions. For more details see Section 6.5. The normalization of the main backgrounds from top quark ($t\bar{t}$, tW) and WW production are left to be determined in the fit via the constraint provided by the low $m_{T2}(ll)$ region with and without b-tagged jets. An individual multiplicative factor to each predicted rate is introduced as free parameter in the simultaneous ML fit. Since each normalization parameter is correlated among decay channels and bins, the overall normalization of the backgrounds are constrained by the dominant parts of their contribution when fitting in the signal regions. To take into account the relative rate of tW with respect to the $t\bar{t}$ process, a normalization uncertainty on the tW estimate is introduced as a nuisance parameter with a log-normal prior. This uncertainty covers the relative difference between the cross section as well as any interference effect between the two processes (see Section 6.5).

As explained in Section 6.3, in the chargino search the signal regions with no b-tagged jets are split in two sub-regions according to whether any jet with $p_T > 20$ GeV is found in the event or not. Due to the rough modeling of jet multiplicity distribution at high p_T^{miss} , two additional normalization parameters are added to the ML fit in order to estimate the rate of events with no jets for diboson (WW , WZ) and jet enriched ($t\bar{t}$, tW , $t\bar{t}Z$, $t\bar{t}W$) backgrounds respectively. These two rate parameters are allowed to float within the normalization uncertainties derived in control regions. The variations are treated independently in the p_T^{miss} bins. The total number of yields expected for each process in the two regions with and without jets is constrained to remain invariant, so that only a migration of events from the two signal regions is allowed. For more details see Appendix B. These constrained rate parameters can be taken as a nuisance parameter expressed with a flat prior.

In general, the systematic uncertainties used in this analysis are deemed to be correlated, or partially correlated, across the SRs and the various processes. The fitting procedure takes these correlations into account for the final result.

7.2 Results

The results of the fit in the SRs for the chargino search are shown in Figures 7.1 and 7.2 for DF and SF events, respectively. The results for the top squark search are shown in Figures 7.3 and 7.4. Each figure compares the number of observed events in the SRs with the expected yields from SM processes after a background-only fit. As a comparison, the expected yields for a representative signal point are given. No excess over SM prediction is observed in data. The total expected SM contributions before the fit and after a background+signal fit are also shown. Detailed information on the observed and expected yields after the background-only fit are given in Tables 7.1-7.2 for the DF and SF dilepton final states and all SRs. The expected yields of

two representative signal points used in the chargino search and top squark search are presented respectively in Tables 7.3 and 7.4.

$m_{T2}(ll)$ [GeV]		0–20	20–40	40–60	60–80	80–100	100–120	≥ 120
DF events								
SR1 _{0tag} ^{jets}	Predicted	1493 ± 32	558 ± 12	719 ± 16	730 ± 16	316 ± 10	45.1 ± 3.1	13.7 ± 2.8
	Observed	1484	532	732	725	298	47	13
SR1 _{0tag} ^{0jet}	Predicted	41.9 ± 5	27.4 ± 3.8	34.1 ± 4.8	42 ± 5.5	21.1 ± 3.4	6 ± 1.3	7.9 ± 2.1
	Observed	39	24	33	44	13	6	9
SR2 _{0tag} ^{jets}	Predicted	534 ± 15	158.6 ± 5.9	167.9 ± 6.1	157.9 ± 6.5	42.4 ± 2.9	5.9 ± 1	9 ± 1.7
	Observed	511	162	156	176	43	5	9
SR2 _{0tag} ^{0jet}	Predicted	10.3 ± 1.7	7 ± 1.5	6.5 ± 1.3	6.9 ± 1.3	2.19 ± 0.69	1.59 ± 0.7	7.8 ± 1.8
	Observed	10	4	4	6	2	2	7
SR3 _{0tag}	Predicted	127.9 ± 7.2	28.3 ± 2	30.2 ± 2.4	23.1 ± 2	4.96 ± 0.73	1.12 ± 0.38	4.5 ± 1.2
	Observed	116	35	29	21	3	1	5
SF events								
SR1 _{0tag} ^{jets}	Predicted	1310 ± 29	499 ± 12	623 ± 14	634 ± 15	271.7 ± 8.9	51.6 ± 3.5	48.6 ± 5.5
	Observed	1324	499	609	659	284	57	47
SR1 _{0tag} ^{0jet}	Predicted	44.1 ± 7.5	28.5 ± 4.1	33.5 ± 4.4	33.5 ± 4.5	18.6 ± 2.6	7.7 ± 1.6	12.5 ± 2.5
	Observed	43	40	39	33	17	6	12
SR2 _{0tag} ^{jets}	Predicted	474 ± 14	134.8 ± 5.1	155.1 ± 5.5	128.5 ± 5.5	37.1 ± 2.5	7.29 ± 0.91	23.9 ± 2.4
	Observed	493	123	166	118	33	7	25
SR2 _{0tag} ^{0jet}	Predicted	10.9 ± 1.9	7.8 ± 1.8	7.3 ± 1.4	7.9 ± 1.3	1.9 ± 0.52	1.28 ± 0.58	7.1 ± 1.4
	Observed	8	12	11	10	3	2	7
SR3 _{0tag}	Predicted	112.8 ± 6.3	27.9 ± 2.2	24.2 ± 1.8	22.5 ± 1.8	5.2 ± 1	1.36 ± 0.36	10.6 ± 1.2
	Observed	110	35	26	26	2	1	14

TABLE 7.1: Observed and expected yields of DF (the upper half of Table) and SF (the lower half) events in the SRs for the chargino search after the background-only fit. The quoted uncertainties in the background predictions include statistical and systematic contributions.

A good agreement between data and total SM predictions within the interval of uncertainty is observed in all signal regions. However some fluctuations may be highlighted in the regions with little statistics, the $m_{T2}(ll)$ tails and higher p_T^{miss} bins, where the statistical uncertainty is dominant. There is a general improvement in the ratio data to simulation after the fit. The SM predictions have been scaled according to the observed data taking into account both the systematic and statistical uncertainties and the normalization parameters. This effect is more appreciable in SR2_{0tag}^{0jet} and SR3_{0tag} regions where the freedom in the normalization of main backgrounds together with the constrained parameter introduced for the rate of events with 0 jets have a higher impact on the predictions. A study of how the nuisance parameters behave during the fit is done in Section 7.2.1.

The asymptotic approximation of the CLs criterion [171–173] has been used to set upper limits at 95% confidence level (CL) on the production cross sections for the different signal models considered. A description of the test statistic used for deriving the exclusion limits and the interpretation plots can be found in Section 7.3.

7.2.1 Pull of nuisance parameters

The presence of nuisance parameters broadens the likelihood function and it is reflected as loss of information about the estimation of the parameter of interest. With the aim of verifying how the nuisance parameters are modified during the fitting

$m_{T2}(ll)$ [GeV]		0–20	20–40	40–60	60–80	80–100	100–120	≥ 120
DF events								
SR1 _{tags}	Predicted	3525 ± 80	1505 ± 31	1958 ± 42	2049 ± 46	897 ± 22	108.4 ± 7.3	13.4 ± 2.2
	Observed	3534	1494	1938	2068	879	111	15
SR1 _{0tags}	Predicted	1542 ± 33	588 ± 13	756 ± 15	771 ± 19	338.3 ± 9.3	50.6 ± 3.8	21 ± 3.8
	Observed	1523	556	765	769	311	53	22
SR2 _{tags}	Predicted	1036 ± 37	363 ± 13	415 ± 14	377 ± 14	105.1 ± 6.5	12.3 ± 2	5.02 ± 0.82
	Observed	1045	357	412	389	111	11	1
SR2 _{0tags}	Predicted	545 ± 18	164.3 ± 7.3	173.2 ± 6.2	165.1 ± 6.8	44.8 ± 3.1	7.1 ± 1.4	15.5 ± 3
	Observed	521	166	160	182	45	7	16
SR3 ^{ISR} _{tags}	Predicted	152.1 ± 9.9	35.5 ± 2.7	32.3 ± 2.3	25 ± 2.2	4.67 ± 0.77	0.41 ± 0.38	0.41 ± 0.26
	Observed	133	44	36	26	2	1	0
SR3 ^{ISR} _{0tags}	Predicted	103.9 ± 6.8	21.3 ± 1.9	22.2 ± 2.1	15.4 ± 1.6	3.51 ± 0.6	0.53 ± 0.21	0.53 ± 0.34
	Observed	100	27	22	12	3	0	1
SF events								
SR1 _{tags}	Predicted	2979 ± 68	1277 ± 30	1644 ± 35	1712 ± 37	762 ± 19	91.9 ± 6.1	18.1 ± 2.1
	Observed	3003	1266	1674	1671	798	85	16
SR1 _{0tags}	Predicted	1350 ± 33	526 ± 13	656 ± 15	670 ± 17	289.2 ± 7.6	57.9 ± 4.2	61.8 ± 5.8
	Observed	1367	539	648	692	301	63	59
SR2 _{tags}	Predicted	888 ± 30	319 ± 12	363 ± 14	323 ± 13	90.5 ± 5.5	10.8 ± 1.5	7.43 ± 0.98
	Observed	900	315	343	325	86	13	11
SR2 _{0tags}	Predicted	487 ± 16	140.7 ± 5.5	161.9 ± 5.9	134.5 ± 6.2	39.6 ± 2.7	8.1 ± 1.1	30.6 ± 3
	Observed	501	135	177	128	36	9	32
SR3 ^{ISR} _{tags}	Predicted	129.6 ± 8.9	29.6 ± 2.1	27.8 ± 2.1	22.2 ± 1.9	3.71 ± 0.57	0.47 ± 0.42	0.71 ± 0.38
	Observed	123	27	28	38	4	1	1
SR3 ^{ISR} _{0tags}	Predicted	91.5 ± 6.1	20.1 ± 1.8	16.5 ± 1.4	13.7 ± 1.4	3.14 ± 0.58	0.78 ± 0.36	1.63 ± 0.42
	Observed	92	26	17	12	1	1	2

TABLE 7.2: Observed and expected yields of DF (the upper half of Table) and SF (the lower half) events in the SRs for the top squark search after the background-only fit. The quoted uncertainties in the background predictions include statistical and systematic contributions.

process and quantifying their correlation with the signal strength estimation, the pull distribution has been studied for each parameter using simulated pseudo-data, here referred to as the “Asimov” dataset [171]. Since the MC simulation and pseudo-data shapes and yields are equal, when fitting to Asimov dataset one is expected to be able to just constrain the nuisance parameters.

After the ML fit, the initial values of the nuisance parameters and their respective uncertainties could have changed as a result of a better description of the fitted dataset. The shifted value of the initial estimation divided by the initial uncertainty should be encountered very close to zero. In the ideal situation when our original assumptions describing the data coincide exactly with the fitted estimation, the shifted value is equal to zero. Any deviation from zero indicates how the fit has moved our original guess. Similarly, the nuisance uncertainty can be constrained during the fitting process. Ideally, the change is small and the ratio between the associated uncertainty after and before the fit is close to one. Likewise, it is also useful to know which parameters are strongly correlated with the signal strength since these can have the largest impact on the observed and expected limits.

In the Appendix D, the pulls of the nuisance parameters after the fit for both hypothesis, only background and signal + background, with two reference mass points are reported. No special pathologies have been found.

$m_{T2}(ll)$ [GeV]	0–20	20–40	40–60	60–80	80–100	100–120	≥ 120
DF events							
SR1 _{0tag} ^{jets}	1.93 ± 0.37	0.37 ± 0.16	0.93 ± 0.26	1.08 ± 0.28	1.59 ± 0.33	2.42 ± 0.41	10.54 ± 0.86
SR1 _{0tag} ^{0jet}	0.16 ± 0.1	0.36 ± 0.15	0.31 ± 0.14	0.21 ± 0.12	0.83 ± 0.24	1.23 ± 0.29	9 ± 0.78
SR2 _{0tag} ^{jets}	1.03 ± 0.26	0.33 ± 0.15	0.65 ± 0.21	0.7 ± 0.22	0.66 ± 0.21	0.49 ± 0.19	13.87 ± 0.99
SR2 _{0tag} ^{0jet}	0.16 ± 0.1	0.1 ± 0.06	0.26 ± 0.13	0.2 ± 0.11	0.13 ± 0.09	0.38 ± 0.16	10.59 ± 0.84
SR3 _{0tag}	1.74 ± 0.32	0.23 ± 0.12	0.12 ± 0.09	0.39 ± 0.16	0.61 ± 0.2	0.27 ± 0.13	6.13 ± 0.64
SF events							
SR1 _{0tag} ^{jets}	1.47 ± 0.33	0.81 ± 0.24	1.1 ± 0.28	1.35 ± 0.31	1.56 ± 0.33	2.38 ± 0.41	9.34 ± 0.81
SR1 _{0tag} ^{0jet}	0.2 ± 0.11	0.23 ± 0.12	0.43 ± 0.17	0.74 ± 0.22	0.84 ± 0.24	1.28 ± 0.29	10.38 ± 0.84
SR2 _{0tag} ^{jets}	1.38 ± 0.31	0.3 ± 0.14	0.31 ± 0.14	0.58 ± 0.2	0.81 ± 0.24	0.97 ± 0.26	12.33 ± 0.93
SR2 _{0tag} ^{0jet}	0.07 ± 0.07	0.14 ± 0.1	0.23 ± 0.12	0.03 ± 0.03	0.1 ± 0.08	0.17 ± 0.1	8.58 ± 0.76
SR3 _{0tag}	1.1 ± 0.25	0.48 ± 0.17	0.34 ± 0.14	0.2 ± 0.11	0.37 ± 0.15	0.43 ± 0.16	7.43 ± 0.72

TABLE 7.3: Expected yields of DF (the upper half of Table) and SF (the lower half) events in the signal regions used in the chargino search, for a reference $\tilde{\chi}_1^\pm \rightarrow \tilde{l}\nu \rightarrow l\nu\tilde{\chi}_1^0$ signal with $m_{\tilde{\chi}_1^\pm} = 500$ GeV and $m_{\tilde{\chi}_1^0} = 200$ GeV. The yields are normalized to an integrated luminosity of 35.9 fb^{-1} . Quoted uncertainties are statistical only.

7.3 Interpretation of results

The procedure for setting the observed and expected exclusion limits is based on the modified frequentist method CL_s for the asymptotic approximation of the LHC test-statistic distributions, which is build on a profile likelihood ratio [171]:

$$t_\mu = -2 \ln \frac{\mathcal{L}(\text{data}|\mu, \hat{\Theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\Theta})}, \quad (7.3)$$

where μ is the signal strength parameter established by the model and $\hat{\Theta}_\mu$ refers to the conditional maximum likelihood estimator of Θ given a value of μ and the data. The pair of parameter estimators $\hat{\mu}$, and $\hat{\Theta}_\mu$ correspond to the global (unconditional) maximum likelihood given the same data. The maximum likelihood estimation of the parameter $\hat{\mu}$ is the value of the signal strength μ that better describes the data.

For each model the value of the signal strength parameter is scanned and the corresponding hypothesis is tested until a value which results in $\text{CL}_s = 0.05$ is found. The upper limit on the cross section for each model corresponds to such value of $\hat{\mu}$ for which a $\text{CL}_s=0.05$ is obtained and the model can be excluded when the cross section ratio $\sigma/\sigma_{nom} < 1$ is given.

In this analysis, the 95% CL upper limits on chargino pair production cross sections with the chargino decaying into sleptons are shown in Figure 9.2 (left). The $\tilde{\chi}_1^\pm \rightarrow \tilde{l}\nu \rightarrow l\nu\tilde{\chi}_1^0$ and $\tilde{\chi}_1^\pm \rightarrow \tilde{\nu}l \rightarrow l\nu\tilde{\chi}_1^0$ decay chains are given a \mathcal{B} of 50% each, and the sleptons are assumed to be degenerate, with a mass equal to the average of the chargino and neutralino masses. By comparing the upper limits with $pp \rightarrow \tilde{\chi}_1^\pm \tilde{\chi}_1^\mp$ production cross sections, observed and expected exclusion regions in the $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0})$ plane are also determined. Masses are excluded up to values of about 800 and 320 GeV for the chargino and the neutralino, respectively. Limited sensitivity is found when the chargino is assumed to decay into a W boson and the lightest neutralino, due to the relatively small \mathcal{B} for the leptonic decay of the W boson. For this scenario, we

$m_{T2}(ll)$ [GeV]	0–20	20–40	40–60	60–80	80–100	100–120	≥ 120
DF events							
SR1 _{tags}	16.5 ± 1.5	6.2 ± 0.88	6.01 ± 0.82	8.9 ± 1.1	9.4 ± 1	4.95 ± 0.8	1.78 ± 0.48
SR1 _{0tag}	21.6 ± 1.7	7.9 ± 1	10.6 ± 1.2	13.7 ± 1.4	14.5 ± 1.5	12.2 ± 1.3	4.32 ± 0.81
SR2 _{tags}	13 ± 1.2	4.12 ± 0.69	3.9 ± 0.65	4 ± 0.69	3.01 ± 0.64	3.07 ± 0.63	1.65 ± 0.46
SR2 _{0tag}	18.1 ± 1.5	4.63 ± 0.77	5.22 ± 0.85	5.32 ± 0.85	6.9 ± 0.97	4.78 ± 0.81	3 ± 0.64
SR3 _{tags} ^{SR}	8.55 ± 0.95	1.53 ± 0.39	1.15 ± 0.34	0.73 ± 0.28	0.72 ± 0.28	0.73 ± 0.29	0.82 ± 0.29
SR3 _{0tag} ^{SR}	9.8 ± 1.1	2.3 ± 0.53	2.02 ± 0.51	1.97 ± 0.5	1.48 ± 0.44	1.13 ± 0.38	1.09 ± 0.38
SF events							
SR1 _{tags}	13.9 ± 1.3	4.36 ± 0.72	4.75 ± 0.75	7.53 ± 0.93	7.66 ± 0.95	4.64 ± 0.74	1.63 ± 0.48
SR1 _{0tag}	18 ± 1.6	6.88 ± 0.99	9 ± 1.1	9.2 ± 1.2	12.4 ± 1.3	10.4 ± 1.2	3.41 ± 0.72
SR2 _{tags}	9.9 ± 1.1	3.06 ± 0.66	2.4 ± 0.59	3.28 ± 0.61	3.61 ± 0.63	1.62 ± 0.42	1.07 ± 0.33
SR2 _{0tag}	12.4 ± 1.3	2.86 ± 0.64	4.14 ± 0.75	3.26 ± 0.65	5.31 ± 0.87	5.32 ± 0.85	2.75 ± 0.63
SR3 _{tags} ^{SR}	4.98 ± 0.71	1.31 ± 0.4	0.88 ± 0.3	1.1 ± 0.35	0.64 ± 0.25	0.63 ± 0.24	0.57 ± 0.25
SR3 _{0tag} ^{SR}	6.74 ± 0.91	1.12 ± 0.38	1.27 ± 0.4	1.19 ± 0.38	0.94 ± 0.35	0.8 ± 0.32	1.22 ± 0.39

TABLE 7.4: Expected yields of DF (the upper half of Table) and SF (the lower half) events in the signal regions used in the top squark search, for a reference $\tilde{t}_1 \rightarrow t\tilde{\chi}_1^0$ signal with $m_{\tilde{t}_1} = 350$ GeV and $m_{\tilde{\chi}_1^0} = 225$ GeV. The yields are normalized to an integrated luminosity of 35.9 fb^{-1} . Quoted uncertainties are statistical only.

derive upper limits on chargino pair production cross section assuming a lightest neutralino mass of 1 GeV. Observed and expected upper limits as a function of the chargino mass are compared to theoretical cross sections in Figure 9.2 (right).

Figure 9.3 shows the observed and expected 95% CL upper limits on top squark production cross section for the two SMS considered. While the search strategy has been optimized for a compressed scenario, the results are presented on the whole $(m_{\tilde{t}_1}, m_{\tilde{\chi}_1^0})$ plane for completeness. Also shown are the expected and observed exclusion regions when assuming NLO+NLL top squark pair production cross sections. When assuming the top squark to decay into a top quark and a neutralino, top squark (neutralino) masses are excluded up to about 420 (360) GeV in the compressed mass region where Δm lies between the top quark and W boson masses. For the $\tilde{t}_1 \rightarrow b\tilde{\chi}_1^\pm \rightarrow bW\tilde{\chi}_1^0$ decay mode, a lower bound $\Delta m \approx 2m_W$ is set by the assumption that $m_{\tilde{\chi}_1^\pm} = (m_{\tilde{t}_1} + m_{\tilde{\chi}_1^0})/2$. For $\Delta m \approx 2m_W$, top squark masses are excluded in the range 225–325 GeV. The uncovered region around a top squark mass of 200 GeV in Figure 9.3 (right) corresponds to a signal phase space similar to that of $t\bar{t}$, with little contribution from the neutralinos to p_T^{miss} . In this situation, the uncertainty in the modeling of p_T^{miss} FastSim events becomes too large to provide any signal sensitivity.

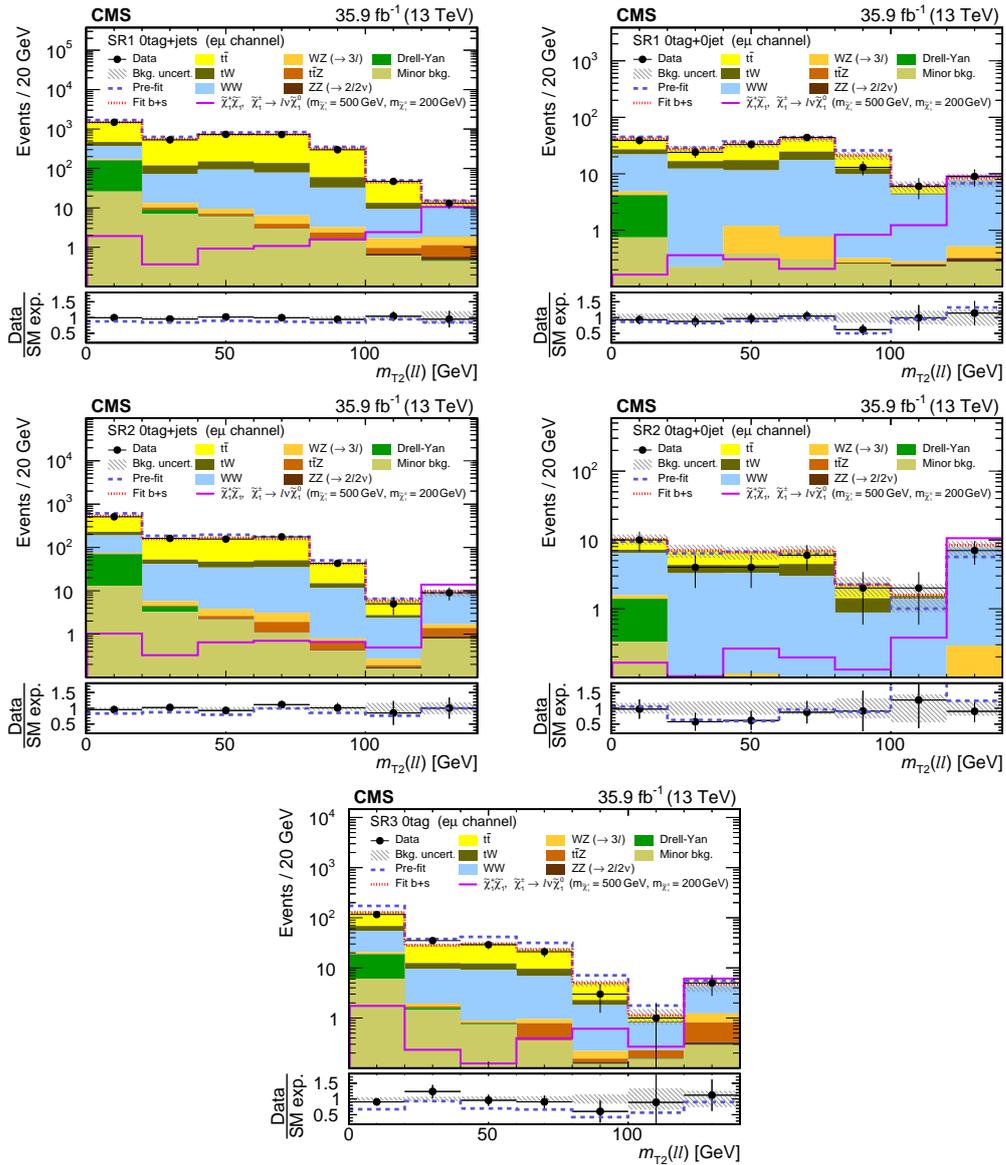


FIGURE 7.1: Distributions of $m_{T2}(ll)$ after the fit to data in the chargino SRs with $140 < p_T^{miss} < 200$ GeV (upper plots), $200 < p_T^{miss} < 300$ GeV (middle), and $p_T^{miss} > 300$ GeV (lower), for DF events without b-tagged jets and at least one jet (left plots) and no jets (right plots). The lower plot for the SR with $p_T^{miss} > 300$ GeV shows all the events without b-tagged jets regardless of their jet multiplicity. The solid magenta histogram shows the expected $m_{T2}(ll)$ distribution for chargino pair production with $m_{\tilde{\chi}_1^\pm} = 500$ GeV and $m_{\tilde{\chi}_1^0} = 200$ GeV. Expected total SM contributions before the fit (dark blue dashed line) and after a background+signal fit (dark red dotted line) are also shown. The last bin includes the overflow entries. In the bottom panel, the ratio of data and SM expectations is shown for the expected total SM contribution after the fit using the background-only hypothesis (black dots) and before any fit (dark blue dashed line). The hatched band represents the total uncertainty after the fit [19].

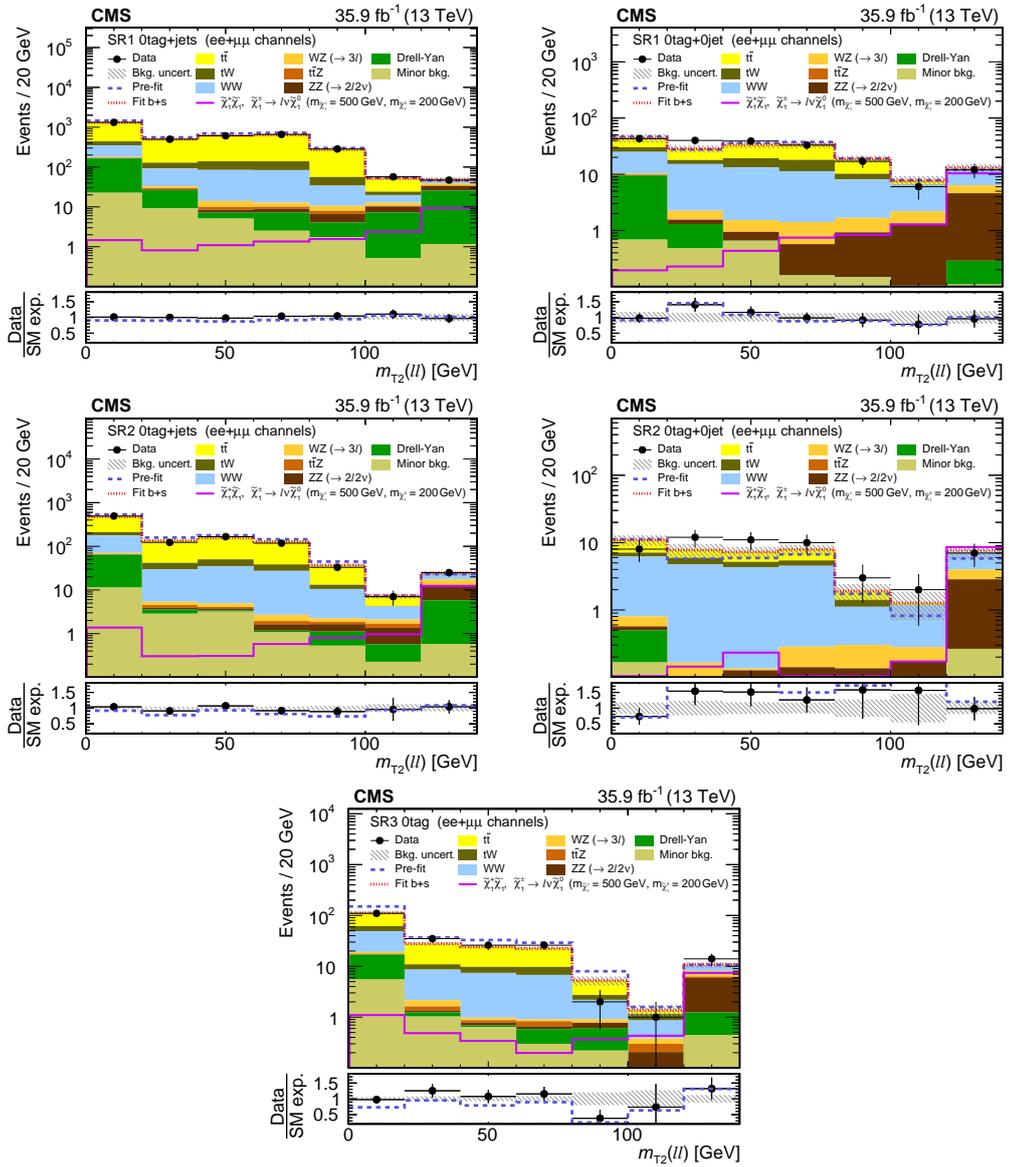


FIGURE 7.2: The same distributions of $m_{T2}(ll)$ as Fig. 7.1, but for SF events [19].

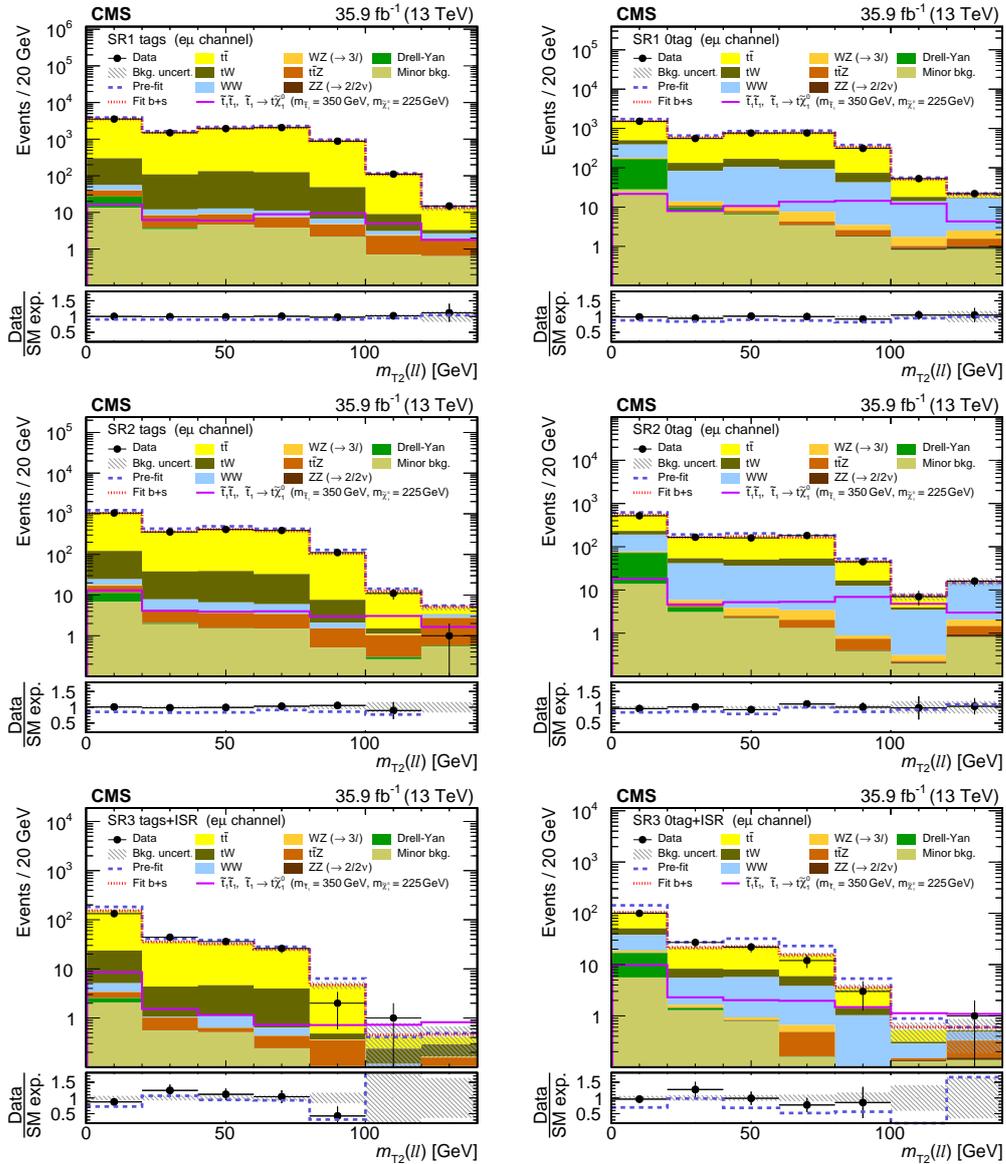


FIGURE 7.3: Distributions of $m_{T2}(ll)$ after the fit to data in the top squark SRs with $140 < p_T^{miss} < 200$ GeV (upper plots), $200 < p_T^{miss} < 300$ GeV (middle), or $p_T^{miss} > 300$ GeV (lower), for DF events with b-tagged jets (left plots) and without b-tagged jets (right plots). The solid magenta histogram shows the expected $m_{T2}(ll)$ distribution for top squark pair production with $m_{\tilde{t}_1} = 350$ GeV and $m_{\tilde{\chi}_1^0} = 225$ GeV.

Expected total SM contributions before the fit (dark blue dashed line) and after a background+signal fit (dark red dotted line) are also shown. The last bin includes the overflow entries. In the bottom panel, the ratio of data and SM expectations is shown for the expected total SM contribution after the fit using the background-only hypothesis (black dots) and before any fit (dark blue dashed line). The hatched band represents the total uncertainty after the fit [19].

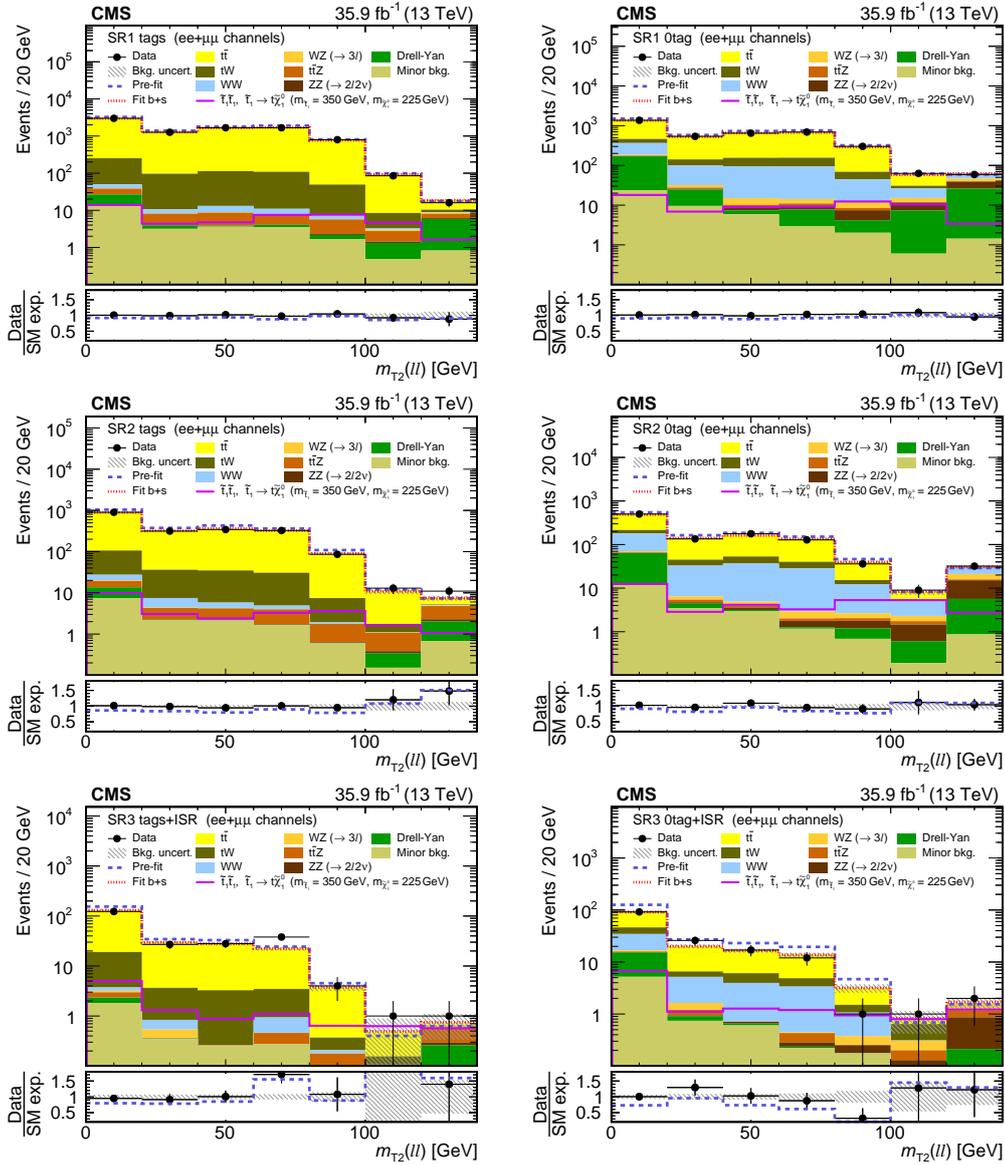


FIGURE 7.4: The same distributions of $m_{T2}(ll)$ as Fig. 7.3, but for SF events [19].

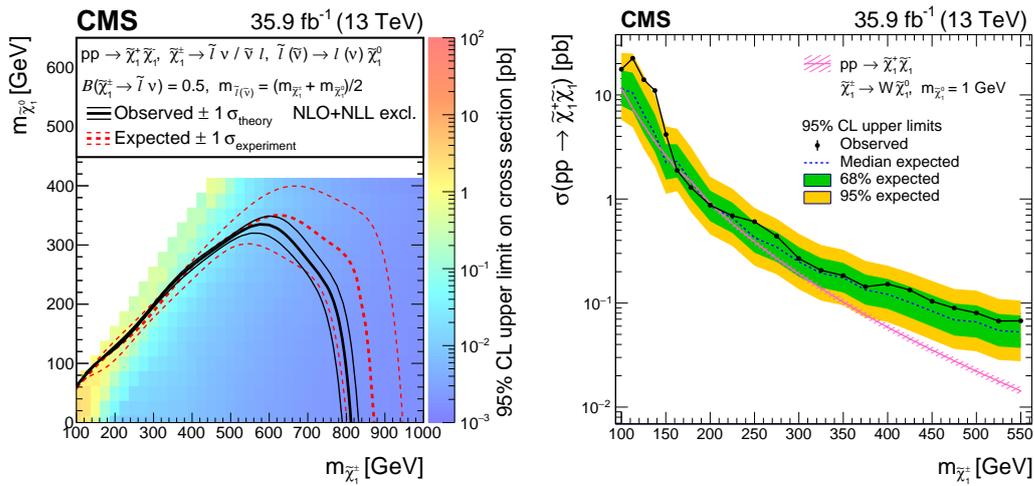


FIGURE 7.5: Left: upper limits at 95% CL on chargino pair production cross section as a function of the chargino and neutralino masses, when the chargino undergoes a cascade decay $\tilde{\chi}_1^\pm \rightarrow \tilde{l} \nu$ ($\tilde{l} \nu \rightarrow l \nu \tilde{\chi}_1^0$). Exclusion regions in the plane $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0})$ are determined by comparing the upper limits with the NLO+NLL production cross sections. The thick dashed red line shows the expected exclusion region. The thin dashed red lines show the variation of the exclusion regions due to the experimental uncertainties. The thick black line shows the observed exclusion region, while the thin black lines show the variation of the exclusion regions due to the theoretical uncertainties in the production cross section. Right: observed and expected upper limits at 95% CL as a function of the chargino mass for a neutralino mass of 1 GeV, assuming chargino decays into a neutralino and a W boson ($\tilde{\chi}_1^\pm \rightarrow W \tilde{\chi}_1^0$) [19].

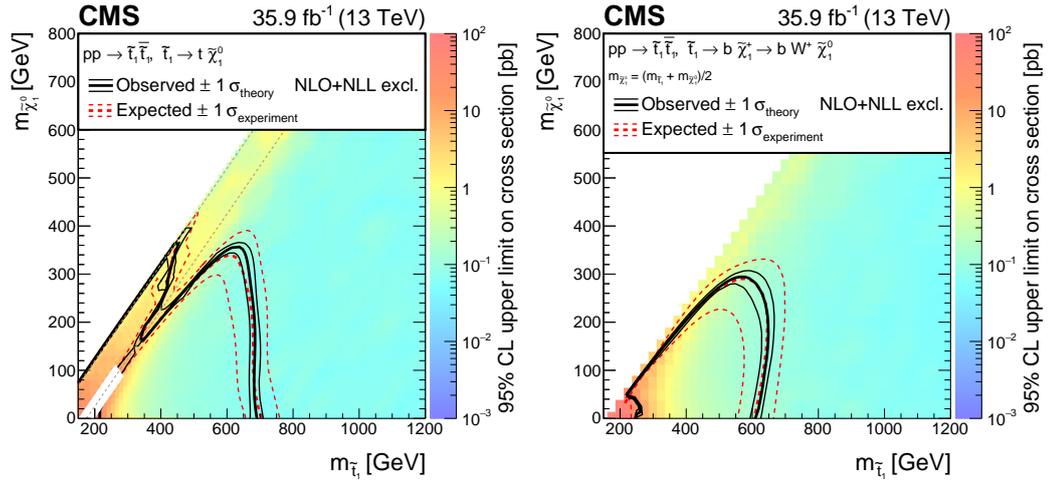


FIGURE 7.6: Upper limits at 95% CL on top squark production cross section as a function of the top squark and neutralino masses. The plot on the left shows the results when top squark decays into a top quark and a neutralino are assumed. The two diagonal gray dashed lines enclose the compressed region where $m_W < m_{\tilde{t}_1} - m_{\tilde{\chi}_1^0} \lesssim m_t$. The plot on the right gives the limits for top squarks decaying into a bottom quark and a chargino, with the latter successively decaying into a W boson and a neutralino. The mass of the chargino is assumed to be equal to the average of the top squark and neutralino masses. Exclusion regions in the plane $(m_{\tilde{t}_1}, m_{\tilde{\chi}_1^0})$ are determined by comparing the upper limits with the NLO+NLL production cross sections. The thick dashed red line shows the expected exclusion region. The thin dashed red lines show the variation of the exclusion regions due to the experimental uncertainties. The thick black line shows the observed exclusion region, while the thin black lines show the variation of the exclusion regions due to the theoretical uncertainties in the production cross section [19].

Chapter 8

Conclusions

A search for new physics in events with two oppositely charged isolated leptons and missing transverse momentum has been presented. The data used consist of a sample of proton-proton collisions collected with the CMS detector during the 2016 LHC run at a centre-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 35.9 fb^{-1} . No evidence for a deviation with respect to Standard Model predictions was observed in data. The results have been interpreted as upper limits on the cross sections of supersymmetric particle production for several simplified model spectra.

Chargino pair production has been investigated in two possible decay modes. If the chargino is assumed to undergo a cascade decay through sleptons, an exclusion region in the $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0})$ plane can be derived, extending to chargino masses of 800 GeV and neutralino masses of 320 GeV. These are the most stringent limits on this model to date. For chargino decays into a neutralino and a W boson, limits on the production cross section have been derived assuming a neutralino mass of 1 GeV, and chargino masses in the range 170–200 GeV have been excluded.

Top squark pair production was also tested, with a focus on compressed decay modes. A model with the top squark decaying into a top quark and a neutralino was considered. In the region where $m_W < m_{\tilde{t}_1} - m_{\tilde{\chi}_1^0} \lesssim m_t$, limits extend up to 420 and 360 GeV for the top squark and neutralino masses, respectively. An alternative model has also been considered, where the top squark decays into a chargino and a bottom quark, with the chargino subsequently decaying into a W boson and the lightest neutralino. The mass of the chargino is assumed to be average between the top squark and neutralino masses, which gives a lower bound to the mass difference (Δm) between the top squark and the neutralino of $\Delta m \approx 2m_W$. This search reduces by about 50 GeV the minimum Δm excluded in the previous result with two leptons in the final state [59] from the CMS Collaboration, excluding top squark masses in the range 225–325 GeV for $\Delta m \approx 2m_W$.

In summary, by exploiting the full data set collected by the CMS experiment in 2016, this search extends the existing exclusion limits on the pair production of charginos decaying via sleptons [58], improving of about 70 GeV the limit on the chargino mass for a massless neutralino. Exclusion limits on the top squark pair production extend the results obtained by the CMS Collaboration in final states with two oppositely charged leptons [59] to the compressed region, where they are competitive with the results obtained by the ATLAS Collaboration in the same decay channel [65].

The amount of data collected by CMS from proton-proton collisions during 2017 and 2018 has substantially increased the available statistics. The combined analysis

of the data collected by the CMS experiment during 2016, 2017 and 2018 is expected to improve the current limits if the possibility of discovery is dismissed.

Different strategies of optimization can be performed to increase the sensitivity of the searches. In the top squark search, the development of more accurate requirements on jets from the initial state radiation and the optimization of the $m_{T2}(ll)$ distribution binning can improve the discrimination between signal and background in the compress region at high missing transverse momentum. In the chargino search, extending the analysis with finer bins in the high $m_{T2}(ll)$ region can also provide more sensitivity rejecting most of the dominant backgrounds.

Furthermore, the signal regions definition can be improved exploiting variables such as the transverse momentum of the tagged jet or the scalar sum of all visible object (jets and leptons) instead of the jet multiplicity. Finally, the introduction of new techniques for estimating the contribution of main background could reduce the value of systematic uncertainties related to that source.

Chapter 9

Resumen

El Modelo Estándar o en inglés *Standard Model* (SM) [20–24], es una teoría cuántica de campos efectiva compatible con la relatividad especial que describe tres de las cuatro interacciones fundamentales de la naturaleza, el electromagnetismo, la interacción débil y la interacción fuerte, así como su influencia sobre las partículas elementales que constituyen la materia. Establece la existencia de 12 partículas con espín 1/2 (fermiones) clasificadas como leptones ($e, \mu, \tau, \nu_e, \nu_\mu, \nu_\tau$) y quarks (u, d, c, s, t, b). A éstas se le añaden sus 12 antipartículas con igual masa pero números cuánticos opuestos. En función de la masa, los fermiones se agrupan en tres generaciones: (e, ν_e, u, d), (μ, ν_μ, c, s), (τ, ν_τ, t, b); siendo la primera generación la que compone toda la materia estable observada en el Universo. Tanto leptones como quarks son susceptibles a la interacción débil, sin embargo, sólo aquellos eléctricamente cargados serán sensibles también a la interacción electromagnética. Los quarks portadores de carga de color, son a su vez susceptibles a la acción de la fuerza fuerte.

En la descripción del Modelo Estándar, cada interacción está determinada por un grupo de simetría gauge. Las partículas encargadas de mediar la acción de dicha interacción se denominan bosones gauge, poseen espín 1 y son los generadores del correspondiente grupo de simetría. En particular, la fuerza electromagnética, descrita por la electrodinámica cuántica es mediada por un bosón eléctricamente neutro, el fotón (γ), generador del grupo de simetría $U(1)$. Puesto que el fotón no tiene masa, el rango de interacción es infinito. La interacción débil descrita por el grupo de simetría $SU(2)$ es mediada por dos bosones cargados (W^\pm) y uno neutro (Z^0). Debido a que los mediadores tienen masas mayores ($m_W = 80.4 \text{ GeV}$ y $m_Z = 91.2 \text{ GeV}$), el rango de acción $\sim 10^{-3} \text{ fm}$ es menor. La interacción fuerte, descrita por la cromodinámica cuántica, es representada por el grupo de simetría gauge $SU(3)$. Está mediada por un octete de color de bosones eléctricamente neutros llamados gluones (g), con un rango de acción de $\sim 1 \text{ fm}$. Debido a la naturaleza de esta interacción, ni los quarks ni los gluones han sido nunca observados como partículas libres, sino en los estados confinados llamados hadrones.

Las simetrías gauge, garantizan un Lagrangiano del Modelo Estándar bien definido y renormalizable, sin embargo, no pueden dotar de masa a las partículas elementales. En 1964, Robert Brout y Francois Englert, Peter Higgs, e independientemente Gerald Guralnik, C. R. Hagen, y Tom Kibble [32–34], publicaron casi simultáneamente un mecanismo para dotar de masa a las partículas manteniendo la invariancia gauge del Lagrangiano. Normalmente conocido como el mecanismo de Higgs, este método propone la ruptura espontánea de la simetría electrodébil ($SU(2) \otimes U(1)$) dando lugar a una nueva partícula neutra de masa no nula y espín 0 conocida como

el bosón de Higgs (H). Los fermiones y bosones W^\pm y Z^0 adquirirían su masa a través del acoplamiento con esta nueva partícula.

Tras el descubrimiento en 2012 del bosón de Higgs [35, 36], el Modelo Estándar se ha establecido firmemente como una teoría efectiva, válida hasta energías del orden de la escala electrodébil (\sim TeV). Las últimas evidencias experimentales demuestran que su actual formulación es incompleta y existen numerosas cuestiones aún por resolver. Entre ellas se encuentran la existencia de la materia oscura, la asimetría entre materia y antimateria o el problema de la jerarquía entre la masa del bosón de Higgs y la masa de Planck. Parece entonces plausible pensar en modelos de nueva física, que describan la naturaleza a escalas más altas de energía en un rango fenomenológico más amplio. Una de las alternativas más prometedoras al Modelo Estándar se encuentra en la supersimetría.

El principio de supersimetría asigna a cada partícula del Modelo Estándar un compañero supersimétrico a través de una transformación de simetría espacio-temporal que cambia en $1/2$ el momento angular de espín de la partícula original. La extensión supersimétrica más simple del Modelo Estándar se denomina, en inglés, *the Minimal Supersymmetric Standard Model* (MSSM) [6–14] y constituye la base para la mayoría de los desarrollos teóricos y experimentales en física de partículas hoy en día. Este modelo puede resolver algunos de los problemas más importantes del actual Modelo Estándar como proveer de un candidato a materia oscura o solucionar el problema de la jerarquía [1, 2].

En el MSSM, los compañeros supersimétricos de los fermiones (leptones y quarks) de espín $1/2$ son partículas de espín 0 denominadas sfermions (sleptones y squarks), mientras que los compañeros supersimétricos de los bosones de espín 1 son partículas de espín $1/2$ conocidas como gauginos. El caso del bosón de Higgs con espín 0 es algo más complejo, teniendo como compañero supersimétrico el higgsino. Aparecen además nuevas partículas de espín $1/2$ los neutralinos y los charginos. Para asegurar la conservación del número leptónico y bariónico, se introduce una simetría extra, denominada R -parity, cuyo número cuántico asociado se define para cada partícula. Si R -parity es una simetría conservada, aparecen interesantes consecuencias a nivel fenomenológico. En concreto, si la partícula supersimétrica más ligera, en inglés *lightest supersymmetric particle* (LSP), del modelo es estable y neutra, ésta adquiere las características esenciales de una partícula candidata a materia oscura.

Esta tesis doctoral presenta la búsqueda de dos partículas supersimétricas producidas en colisiones proton-proton en el LHC con una energía de centro de masas de $\sqrt{s} = 13$ TeV. La muestra de datos analizada ha sido recogida por el detector CMS en 2016 y corresponde a una luminosidad integrada de 35.9 fb^{-1} . El análisis se centra en la producción del chargino y del top squark más ligeros ($\tilde{\chi}_1^\pm, \tilde{t}_1$) en pares de partícula-antipartícula decayendo en estados finales con dos leptones y momento transversal faltante. Los resultados son interpretados en términos de varios modelos de señal simplificados asumiendo conservación de la simetría R -parity y asumiendo como LSP el neutralino más ligero ($\tilde{\chi}_1^0$), llamado a partir de aquí simplemente neutralino.

La búsqueda de parejas de chargino se realiza en todo el plano de masas, considerando como principal referencia un modelo donde el chargino decae en un leptón (ℓ), un neutrino (ν) y el neutralino a través de la desintegración de un sleptón ($\tilde{\chi}_1^\pm \rightarrow \tilde{l}\nu \rightarrow l\nu\tilde{\chi}_1^0$), o de un sneutrino ($\tilde{\chi}_1^\pm \rightarrow \tilde{\nu}l \rightarrow l\nu\tilde{\chi}_1^0$). Se asume que las tres generaciones de sleptones están degeneradas con una masa igual al promedio entre las

masas del chargino y del neutralino. También se asume que la fracción de desintegración de un chargino en un sleptón cargado o un neutrino es la misma. Los resultados son así mismo interpretados en términos un segundo modelo donde cada chargino decae en un neutralino y un bosón W .

Otras búsquedas de pares de chargino han sido previamente publicadas por la Colaboración CMS en el contexto del primer escenario de señal usando datos de colisiones protón-protón a 8 TeV [52] y por la Colaboración ATLAS en el contexto de ambos escenarios, usando datos de colisiones a 8 TeV [53–55] y a 13 TeV [56–58].

En el caso de los pares de top squarks, se considera como referencia un modelo donde el squark top decae en quark top (t) y un neutralino. La búsqueda se realiza en la zona comprimida del espectro de masas donde la diferencia de masas (Δm) entre el squark top y el neutralino se encuentra entre la masa del quark top y del bosón W , $m_W < \Delta m \lesssim m_t$. Los resultados son además interpretados en términos de otro modelo donde cada uno de los squark top producidos decae en un quark bottom (b) y un chargino que a su vez decae en un bosón W y un neutralino. En este modelo la masa del chargino se asume igual al promedio de las masas del squark top y del neutralino. Este trabajo complementa un resultado previo publicado por la Colaboración CMS [59] y cuyo objetivo era testear los mencionados modelos en la región de mayor masa $\Delta m > m_t$ donde los quarks top son producidos con mayor momento. Con respecto a ese análisis, esta búsqueda gana sensibilidad en la región de masas comprimida al suavizar las exigencias sobre los jets provenientes de la hadronización del quark bottom y mediante la optimización de la selección de eventos de señal enfocándola a la producción de neutralinos con bajo momento.

9.1 El gran colisionador de hadrones

La organización europea para la investigación nuclear, normalmente conocida por su denominación francesa *Conseil Européen pour la Recherche Nucléaire* (CERN), es una institución internacional dedicada a la investigación en física fundamental. Su misión se puede resumir en cuatro palabras: "Investigación", "Innovación", "Colaboración" e "Inspiración".

El CERN comprende hasta la fecha el mayor complejo de aceleradores, detectores e infraestructura computacional dedicado al estudio de física de altas energías del mundo. Siendo el gran colisionador de hadrones, en inglés *Large Hadron Collider* (LHC), su más reciente adquisición.

El propósito esencial del LHC es investigar las cuestiones más relevantes que actualmente se encuentran sin explicación dentro del marco propuesto por el Modelo Estándar. Entre ellas están, el recientemente descubierto bosón de Higgs (2012), la asimetría entre materia y antimateria, qué es la gravedad, o cuál es la naturaleza de la materia oscura.

En el LHC, dos haces hadrones, protones o iones pesados, son acelerados en direcciones opuestas a lo largo de una circunferencia de 26.7 km de longitud, alcanzando energías del orden del TeV. Estas partículas viajan al 99% de la velocidad de luz y se hacen colisionar creando las condiciones iniciales que se pudieron dar en las etapas tempranas del Universo tras el Big Bang.

Son cuatro los principales experimentos que operan en el LHC: ALICE, dedicado a las colisiones de iones pesados, LHCb, diseñado para el estudio de la física del quark b, ATLAS un detector con múltiples usos, y CMS cuyo diseño es en varios aspectos complementario al de ATLAS, y el experimento con el que se han recogido los datos de este análisis.

El detector CMS [51, 77] conocido por sus siglas en inglés, Compact Muon Solenoid, es un experimento de carácter general cuyos principales objetivos son la exploración de nueva física en escalas de energía del orden del TeV y las medidas de alta precisión de diferentes predicciones del SM. En la siguiente sección se pasa a introducir brevemente la estructura de este detector.

9.2 El detector CMS

El detector CMS tiene un diseño compacto y hermético con una arquitectura cilíndrica formada a base de capas concéntricas. Cada capa juega un papel particular en el proceso de detección y es parte constituyente de alguno de los subdetectores. El exterior de CMS tiene unas dimensiones totales de 21.6 m de largo por 14.6 m de diámetro, albergando un peso de 12500 toneladas, lo que le confiere su característica de detector compacto. Se pueden distinguir dos grandes regiones, la zona central o barril del cilindro y la zona delantera o tapas del cilindro. En el momento de la colisión, cada uno de los haces de protones entra en el detector por uno de los dos extremos y viaja a lo largo del eje central hasta encontrarse con el otro haz en el centro del aparato.

El elemento central, que da el nombre de CMS, es un solenoide superconductor de 6 m de diámetro interno que genera un campo magnético constante de 3.8 T. Es en términos de tamaño, peso y rigidez estructural el componente más grande de CMS y sirve como principal soporte para el resto de sistemas de detección. En su interior y rodeando el punto de colisión, se encuentra el sistema interno de detección de trazas o en inglés, tracker. Está compuesto por un detector de píxeles de silicio rodeado de un detector de mayor tamaño a base de tiras de silicio. Ambos distribuidos entre el barril y las tapas del cilindro proveen medidas en tres dimensiones con las que se reconstruyen las trayectorias de las partículas cargadas. A continuación, se ubican el calorímetro electromagnético (ECAL, del inglés Electromagnetic calorimeter) hecho con cristales centelleadores de tungstato de plomo (PbWO_4) y el calorímetro hadrónico (HCAL, del inglés Hadronic Calorimeter), concebido como un detector de muestro alternando laminas de metal absorbente y plástico centelleador. Ambos sistemas también se distribuyen en la parte central y las tapas del cilindro, midiendo la energía de las partículas en amplio ángulo de aceptación. Una parte delantera situada en la zona más cercana a la línea del haz, completa el revestimiento global ampliando la cobertura hasta prácticamente la totalidad del detector. La información de las trazas y la energía depositada en los calorímetros se combina para reconstruir los electrones y los jets hadrónicos, o chorros de partículas generados por el proceso de hadronización de algún hadrón.

Finalmente y en la parte más externa del detector, parcialmente incrustado en el hierro de retorno del campo magnético del solenoide, se encuentra el espectrómetro de muones. Es un sistema compuesto por tres subdetectores de tipo gaseoso que miden con una alta precisión el momento de los muones. Las diferentes tecnologías escogidas obedecen a los requisitos impuestos por la gran superficie a cubrir y los distintos

entornos de radiación. Así, en la zona central, donde el flujo de muones y el campo magnético residual son bajos se han situado los tubos de deriva (DTs, del inglés Drift Tube); mientras que las tapas del cilindro con niveles más altos de campo magnético y número de partículas están formadas por cámaras de tiras catódicas (CSCs, del inglés Cathode Strip Chamber). En ambos casos las DTs y CSCs son complementadas por cámaras de tiras resistivas (RPC, del inglés Resistive Plate Chambers) que ayudan en la reconstrucción e identificación de los muones. Este sistema de detección de muones es muy sofisticado y una de las características principales del detector CMS.

El detector es casi completamente hermético permitiendo una medida precisa del balance entre el momento total observado y el momento total faltante en el plano transversal a la dirección del haz. Esto da a conocer cuál es la energía de partículas como los neutrinos, que pasan casi invisibles por todas las capas del detector y además sirve para identificar posibles fallos entre otros, de reconstrucción o identificación.

9.3 Reconstrucción de eventos

En el LHC en condiciones normales de funcionamiento, se producen millones de colisiones por segundo, de las que sólo unas pocas serán colisiones inelásticas de interés relevante para la física de partículas. Durante 2016 la frecuencia de cruce de los haces fue de 25 ns dando un promedio de 20 interacciones inelásticas por cruce (interacciones de pileup). Los haces pasaron por el mismo punto entorno 40 millones de veces por segundo, dando un promedio de 10^8 colisiones efectivas por segundo. Esta gran cantidad de información no puede ser guardada y procesada al completo, ya que la velocidad y el espacio de almacenamiento no son suficientes para gestionar todos los sucesos producidos. Sin embargo, no todas las colisiones tienen información relevante desde el punto de vista físico. Las colisiones de baja energía o producidas en ángulos muy oblicuos, por ejemplo, pueden ser descartadas con mínima pérdida de contenido físico, manteniendo así solamente los eventos más esenciales.

El sistema que se encarga de filtrar y retener aquellos sucesos significativos se llama *trigger*. Está diseñado en varios niveles de hardware y software, comenzando desde el sistema de lectura en los distintos subdetectores de CMS. La selección se realiza en dos pasos principales: el *Level-1 (L1) trigger* y el *High-Level trigger (HLT)*. De un modo similar, en ambos pasos, se ejecutan algoritmos especializados que en función de las características físicas previamente fijadas, aceptan o rechazan sucesivamente cada evento. El factor de reducción total respecto al número de colisiones producidas por segundo es del orden de 10^5 .

En cada colisión protón-protón se produce un gran número de partículas, algunas provenientes del vértice primario donde se originó la interacción y otras generadas en cadenas de desintegración con origen en algún vértice secundario. El objetivo principal del análisis físico de los sucesos, es caracterizar, rastrear y contar todas las partículas así producidas, midiendo sus propiedades físicas, como el momento, la carga, y la masa, reconstruyendo de este modo el proceso completo. Las partículas susceptibles de ser detectadas por el experimento CMS son los fotones (γ), los electrones (e^+ , e^-), muones (μ^+ , μ^-), hadrones cargados tales como los protones (p) y piones (π), y hadrones neutros como los neutrones (n).

El proceso de reconstrucción de un evento se inicia con una intrincada colección de medidas de energía, distribución espacial de los impactos en el material ionizado, tiempos de llegada, posiciones..., que de una manera secuencial se va reduciendo para construir cada vez objetos más complejos. Los objetos finales como trazas, vértices, o parámetros de impacto, terminan en la identificación de las partículas subyacentes en el suceso. El algoritmo que se ha usado para reconstruir los objetos físicos de este análisis es el *particle-flow algorithm* [104]. Éste se basa en la información procedente de todos los detectores que forman parte de CMS para reconstruir e identificar cada partícula individual dentro del evento.

Empezando desde el punto de la interacción de los haces, el origen y trayectoria de las partículas cargadas se reconstruye usando las marcas (*hits*, en inglés) que dejan a su paso por las capas del tracker. La energía de los electrones y fotones es principalmente depositada en el ECAL a través de las cascadas electromagnéticas, mientras que los hadrones son sustancialmente absorbidos a su paso por el HCAL. Los muones atraviesan el detector hasta las cámaras de muones donde su interacción con los distintos subdetectores dejan un rastro suficientemente marcado para su detección.

La energía de los fotones se obtiene directamente de las medidas en el ECAL. Por su parte, la energía de los electrones es determinada a partir de la combinación del momento, medido en el vértice primario por el tracker, la energía depositada en el ECAL y la suma total de la energía cedida por todos los fotones producidos en el proceso de *bremsstrahlung* que son espacialmente compatibles con la traza del electrón. De una forma similar, la energía de los hadrones cargados se obtiene a partir del momento medido por el tracker en combinación con los correspondientes depósitos de energía en el ECAL y el HCAL corregidos por los efectos de ruido y por la función de respuesta de los calorímetros ante las casadas hadrónicas de partículas. La energía de los hadrones neutros se obtiene a través de las medidas corregidas de energía del ECAL y HCAL. El momento de los muones se determina a partir de la curvatura de su traza usando las medidas del tracker y del sistema de muones. Se considera vértice primario a aquel vértice reconstruido cuyas trayectorias asociadas suman el mayor momento transversal medido en el plano transversal al haz.

9.4 Identificación y selección de objetos

Tras el proceso de reconstrucción, se aplica un criterio de calidad para la identificación de las partículas genuinas, procedentes de los procesos físicos relevantes como la desintegración de un bosón o la hadronización de un quark, separándolas de las partículas procedentes de otras fuentes como los rayos cósmicos, o la conversión de fotones en electrones. Existen diversas técnicas ampliamente utilizadas en CMS y para cada una de ellas se definen diferentes puntos de operación, o en inglés *working points (WP)*, según los cuales se optimiza la relación entre pureza y eficiencia de identificación.

En este análisis se utiliza un criterio de selección en base a los valores de ciertas variables discriminadoras que permiten identificar electrones y muones con un alto grado de eficiencia. En el caso de los muones la selección combina las medidas en el tracker y en el sistema de muones, así como la calidad del ajuste de la traza global incluyendo los dos sistemas. La identificación de los electrones se apoya en el criterio

de calidad aplicado a la traza que combina las medidas en el ECAL con las medidas compatibles en el tracker.

Para los muones, se utiliza el WP intermedio que tiene una eficiencia de identificación del 98% en la zona central del detector y más del 95% en las tapas. Se ha escogido el WP más duro para electrones, con un 70% de eficiencia y alta pureza. Éste está recomendado en medidas donde el fondo es importante y la probabilidad de identificación errónea como electrón de otra partícula, ej. un hadrón, es alta.

Además de esto, al nivel de este análisis, la selección de leptones se optimiza para identificar aquellos muones o electrones que provienen exactamente de la desintegración de los bosones Z o W . Se utilizan observables como el aislamiento del leptón o el parámetro de impacto de su traza para separar la contribución de partículas genuinas de las producidas por otras causas. Los leptones no genuinos pueden aparecer debido a las interacciones de pileup, a los hadrones cargados que han sido mal identificados como electrones o muones, a leptones que provienen de la desintegración semileptónica de un quark b o c , o como se mencionó antes, pueden ser electrones producidos por efecto fotoeléctrico y muones de origen cósmico que pueden atravesar el detector dejando trazas similares a los muones producidos en una colisión pp .

Los jets de origen hadrónico, se identifican mediante el algoritmo denominado *anti- k_T algorithm* [112, 113] que agrupa todas las partículas reconstruidas anteriormente dentro de un cono de 0.4 como parámetro de distancia. En cada evento, el \vec{p}_T de los jets se calcula como la suma vectorial del \vec{p}_T de todas las partículas dentro del jet. Para evitar posteriores problemas de selección, se eliminan aquellos jets que solapan dentro de un cono de radio 0.4 con un leptón previamente seleccionado. Cada jet debe satisfacer la selección más suave de un criterio de calidad basado en la multiplicidad de las partículas dentro del jet y su composición de fracciones de energía.

Los jets originados por la hadronización de un quark bottom (b jets) son etiquetados por el algoritmo *combined secondary vertex v_2 b -tagging algorithm* (CSV v_2) [132] usando el WP intermedio, con una eficiencia del 50 al 70% para los jets con p_T de 20 a 100 GeV. Así mismo, la tasa de jets provenientes de la hadronización de un quark de sabor ligero o de un gluon identificados erróneamente como b jets es del 1% en el mismo rango de p_T .

El desequilibrio del momento total en el plano transversal al eje del haz se denomina momento faltante en el plano transversal (\vec{p}_T^{miss}) y es definido como la suma vectorial negativa de todas las partículas reconstruidas en el evento. La magnitud de \vec{p}_T^{miss} se denota p_T^{miss} .

9.5 Selección de eventos

Los eventos de interés en este análisis usan una selección de trigger que requiere la presencia de dos leptones (ee , $e\mu$, $\mu\mu$). El momento transversal (p_T) del leptón más energético debe satisfacer, en el caso de los triggers con ee o $e\mu$, un umbral mínimo de 23 GeV y, en el caso de los triggers con $\mu\mu$ el umbral debe ser de 17 GeV. El umbral para el p_T del leptón que le sigue es de 8 ó 12 GeV según sea éste un muón o un electrón. Además, con la intención de incrementar la eficiencia de selección, se

usan también eventos donde al menos un electrón (muón) con un $p_T > 25$ (24) GeV, ha sido identificado bajo un criterio de calidad más riguroso.

La muestra de sucesos analizados obedece a una estrategia de selección de trigger que combina mediante un OR lógico los eventos registrados con dos leptones y con al menos un leptón. La eficiencia combinada de trigger se encuentra en un rango entre el 90 y el 99% dependiendo del momento transversal (p_T) y la pseudorapidity (η) de los leptones.

Se requieren aquellos leptones aislados, con un $p_T > 20$ GeV un valor de $\eta < 2.4$, con parámetro de impacto longitudinal y transversal $|d_z| < 0.1$ cm y $|d_{xy}| < 0.05$ cm respectivamente. Los jets deben satisfacer $p_T > 20$ GeV, $\eta < 2.4$, y los b jets además deben ser etiquetados con el WP intermedio del algoritmo CSVv2. La eficiencia de reconstrucción, identificación aislamiento y selección de leptones es del 65-90% para muones y 26-70% para electrones dependiendo del p_T and η .

9.6 Simulaciones de Monte Carlo

Las muestras de eventos simulados mediante el método de Monte Carlo (MC) se utilizan para estudiar la contribución de los procesos del Modelo Estándar en los datos, así como la tasa de eventos esperada para los distintos modelos de señal estudiados.

La producción de eventos con pares de quarks top-antitop ($t\bar{t}$) son generados con POWHEG v2 [118, 138, 139] y normalizados a una sección eficaz calculada a segundo orden (NNLO) en teoría de perturbaciones de QCD. Los eventos con un quark top producido en asociación a un bosón W (tW) son generados con POWHEG v1 [141] y normalizados a una sección eficaz de precisión NNLO [142]. La producción de dibosones (WW, WZ, ZZ) via aniquilación quark-antiquark es simulada con POWHEG v2 hasta primer orden (NLO) de teoría de perturbaciones. Los eventos de WW son escalados al valor NNLO de la sección eficaz de producción del proceso.

Aquellos eventos procedentes de la producción $q\bar{q} \rightarrow ZZ$ son repesados por factores de corrección (K factors) NNLO/NLO, calculados en función de la masa del sistema ZZ [146]. Además se utilizan otros dos conjuntos adicionales de K factors, en función del p_T del sistema ZZ y de la separación azimutal ($\Delta\phi, \phi$ en radianes) entre los bosones Z , para evaluar la incertidumbre en las propiedades cinemáticas del proceso tras la aplicación de dicha corrección. Por otro lado, la producción de dibosones a través de la fusión de gluones se simula usando MCFM v7 [147], donde la sección eficaz calculada a primer orden de teoría de perturbaciones (LO) se corrige con K factors de tipo NNLO/LO [146, 148]

El proceso de Drell-Yan es generado con MADGRAPH5_aMC@NLOv2.2.2 [119] a LO, y los eventos son escalados a la sección eficaz a NNLO [149]. Los eventos de la producción $t\bar{t}W, t\bar{t}Z$, tribosones (WWW, WWZ, WZZ), and $H \rightarrow WW$ son generados a NLO [150, 151] con el generador MADGRAPH5_aMC@NLO.

Los eventos de señal, tanto en la producción de pares de chargino como de squark top, son generados usando MADGRAPH5_aMC@NLO a LO con hasta dos partones extra en los cálculos del elemento de matriz, y son normalizados a sus respectivas secciones eficaces calculadas a orden NLO [47, 152–159], asumiendo que el resto de partículas son pesadas y están desacopladas. En el caso de la producción de

chargino, los cálculos son hechos en límite de masa degenerada para los winos $\tilde{\chi}_2^0$ y $\tilde{\chi}_1^\pm$, y del bino ligero $\tilde{\chi}_1^0$.

Todos los procesos son generados usando el paquete de funciones de distribución partónicas, o *parton distribution function* (PDF) en inglés, NNPDF3.0 [160]. Para simular los procesos de hadronización, las cascadas partónicas y los sucesos de QCD y QED subyacentes de una colisión pp se utiliza PYTHIA 8.212 [123]. La reacción del detector a los distintos eventos generados es simulada con un modelo realista del detector CMS basada en GEANT4 [126] para los procesos del Modelo Estándar, mientras que para los procesos de señal se utiliza una simulación más sencilla y rápida del detector (FastSim) [164] basada en una parametrización del promedio de respuesta de CMS a las partículas estudiadas. Los eventos simulados son de este modo reconstruidos y estudiados usando los mismos algoritmos que se aplican a los datos.

Para modelar correctamente el efecto de pileup en la producción de los procesos del Modelo Estándar, la distribución de interacciones múltiples en eventos simulados se repesa por la correspondiente distribución observada en datos.

Debido a las diferencias observadas en la resolución de p_T^{miss} , modelada usando la simulación completa del detector y la simulación parametrizada o FastSim, el valor nominal de p_T^{miss} en los eventos de señal es calculada como el promedio del valor encontrado a nivel de generación (sin simulación del detector) y del valor obtenido tras la reconstrucción completa del evento (incluyendo la simulación paramétrica del detector). Como incertidumbre se coge la mitad de su diferencia. Las muestras de señal también se corrigen para mejorar el modelado de los jets procedentes de estados iniciales de radiación ocurridos antes de la colisión, en inglés *initial-state radiation* (ISR) jets. En los eventos de chargino la corrección se hace en función del p_T del ISR jet y en el caso de los eventos de squark top se hace en función de la multiplicidad de los ISR jets en el evento.

Finalmente, los eventos de MC son pesados por factores de corrección que sirven para rectificar las diferencias observadas respecto a los datos. Éstos, se denominan factores de escala y afectan principalmente a la reconstrucción, identificación y aislamiento de los leptones, así como a la identificación de los b jets. Los valores que toman en este análisis difieren de la unidad menos del 10% con valores típicos de corrección entre 2-3 (5)% para la eficiencia obtenida en la identificación de los leptones (b jets) con $p_T > 20$ GeV y $|\eta| < 2.4$. En el caso de los eventos de señal los factores de escala asociados a la identificación y aislamiento de los leptones así como la identificación de b jets son relativos al valor obtenido usando la simulación completa. Otros factores de escala son derivados propiamente en el estudio de los fondos del Modelo Estándar que contribuyen en la región del espacio de fases dónde se busca la señal.

9.7 Estrategia del análisis

La estrategia de búsqueda de este análisis se desarrolla para dos hipótesis de señal: la producción de parejas de chargino-antichargino, y la producción de pares de squarks top-antitop. Mientras que la primera se estudia a lo largo de todo el plano de masas ($m_{\tilde{\chi}_1^\pm}$, $m_{\tilde{\chi}_1^0}$), la segunda es optimizada en la región de masa comprimida, dónde la diferencia de masa entre el squark top y el LSP se encuentra entre las masas del quark top y del bosón W . En ambas búsquedas se emplean las mismas técnicas

para la estimación de fondos y para la extracción de señal, en tanto que el espacio de fases específico para cada hipótesis varía ligeramente en función de sus propiedades cinemáticas.

Los modelos de señal están caracterizados por aquellos estados finales con dos leptones de carga opuesta y un alto valor de p_T^{miss} debido a la contribución de los dos LSP en cada evento. Se define así, una selección de referencia común que requiere, dos leptones aislados de carga opuesta con $|\eta| < 2.4$ y $p_T > 25$ GeV para el leptón más energético siendo $p_T > 20$ GeV para el leptón que le sigue. Para reducir la contaminación de resonancias de baja masa, la producción de $z \rightarrow \tau\tau$, y los leptones no genuinos en jets hadrónicos, se exige que ambos leptones tengan una masa invariante, $m_{\ell\ell}$, mayor que 20 GeV; además, para reducir la contribución principal del proceso de Drell-Yan, se pide que cada par de leptones de carga opuesta y mismo sabor tengan una masa invariante que diste más de 15 GeV de la masa de bosón Z, m_Z , esto es, $|m_{\ell\ell} - m_Z| > 15$ GeV. En base al alto valor de p_T^{miss} que hay en los eventos de señal se requiere un corte mínimo correspondiente a $p_T^{miss} \geq 140$ GeV. Finalmente, se excluyen de la selección aquellos eventos con un tercer leptón identificado con el criterio de calidad más bajo y $pt > 15$ GeV, $|\eta| < 2.4$. Cabe mencionar que se tienen en cuenta aquellos eventos procedentes de la desintegración leptónica de un τ si satisfacen los criterios de selección del análisis.

Tras esta elección de sucesos, los procesos del Modelo Estándar que más contribuyen son las producciones de $t\bar{t}$, tW y WW . En todos ellos, el par de leptones y la p_T^{miss} provienen de la desintegración del bosón W. De este modo, la variable m_{T2} [166], que generaliza la masa transversa (m_T) para un sistema con dos partículas invisibles ($\vec{p}_T^{miss1}, \vec{p}_T^{miss2}$) se puede definir evento a evento usando el sistema visible de los dos leptones ($\vec{p}_T^{lep1}, \vec{p}_T^{lep2}$) como:

$$m_{T2}(ll) = \min_{\vec{p}_T^{miss1} + \vec{p}_T^{miss2} = \vec{p}_T^{miss}} (\max[m_T(\vec{p}_T^{lep1}, \vec{p}_T^{miss1}), (m_T \vec{p}_T^{lep2}, \vec{p}_T^{miss2})]). \quad (9.1)$$

El observable $m_{T2}(ll)$ presenta un límite cinemático en la masa del W para los fondos considerados. Sin embargo, los eventos de señal no obedecen este límite debido a la presencia del neutralino que contribuye al valor total de p_T^{miss} . Este hecho proporciona una clara discriminación entre los procesos de señal y los principales procesos del Modelo Estándar. Además la sensibilidad del análisis se refuerza dividiendo la región de señal, en inglés *signal region* (SR), en diferentes rangos o bins de p_T^{miss} : [140, 200), [200, 300), ≥ 300 GeV. Esto nos permite no sólo explotar las largas colas de la distribución de p_T^{miss} en los modelos de señal, sino también optimizar la sensibilidad a las distintos modelos con diferente separación de masa entre la partícula supersimétrica producida y el LSP. Cada rango o bin de p_T^{miss} es a su vez dividido en eventos con leptones de igual y diferente sabor, en inglés *same flavour* y *different flavour* (SF, DF), que permite aprovechar las diferencias entre los estados finales de los fondos WZ , ZZ y Drell-Yan.

Específicamente para la búsqueda de chargino, se aplica un veto a los eventos con b jets y se dividen los bins con $p_T^{miss} \leq 300$ GeV en función de la presencia o no de al menos un jet con $p_T > 20$ GeV y $|\eta| < 2.4$. Esto reduce la contribución de los procesos $t\bar{t}$, tW y $t\bar{t}W$, permitiendo una mejor discriminación de la señal. Se utilizan así mismo los eventos con al menos un b jet como región de control, en inglés *control region* (CR), para la normalización de la producción de $t\bar{t}$ y tW .

Los estados finales producidos por la desintegración de una pareja de squarks top-antitop se caracterizan por la presencia de dos quarks bottom. Cuando la diferencia entre la masa del squark top y del neutralino (Δm) está próxima a la masa del bosón W (m_W), en el límite de la región comprimida, $\Delta m \lesssim m_W$, los quarks bottom que se producen son poco energéticos y dan lugar a jets de bajo momento y con baja probabilidad de ser etiquetados. En este caso, aplicar un veto a los eventos con al menos un b jet, como en el caso del chargino, reduce la contribución de los principales fondos de top ($t\bar{t}$, tW , $t\bar{t}W$) a la vez que fortalece la presencia de señal. Por otro lado, para escenarios de señal con Δm mayor, los b jets adquieren un mayor momento y conforme ésta se aproxima a la masa del top (m_t), el espacio de fases de la señal se hace muy similar al del fondo de $t\bar{t}$. Aquí, seleccionar aquellos eventos con al menos un b jet, reduce la tasa de fondos como WW y Drell-Yan, aumentando la contribución de la señal. Además de estas consideraciones, se encuentra que la presencia de ISR jets en el bin con $p_T^{miss} \geq 300$ GeV mejora la sensibilidad del análisis. Estos eventos requieren que el jet más energético no sea etiquetado como un b jet, tenga $p_T > 150$ GeV, y que el $\Delta\phi$ entre el jet y \vec{p}_T^{miss} sea mayor que 2.5 rad favoreciendo la topología donde el jet sufre un retroceso respecto a la dirección de propagación del sistema y confiriere así un extra de p_T^{miss} al proceso. En la Figura 9.1 se muestran las distribuciones observadas y simuladas para algunos de los observables usados en la definición de las regiones señal de este análisis.

En cada región de señal, se estudia la distribución de $m_{T2}(ll)$ en eventos con pares de leptones de igual o diferente sabor. Dichas distribuciones son divididas en siete bins de 20 GeV cada uno empezando en 0 GeV y terminando en 120 GeV, incluyendo éste último todos los eventos con $m_{T2}(ll) \geq 120$ GeV. En la Figura 9.1 se muestran las distribuciones observadas y simuladas de las principales variables usadas para definir las regiones de señal.

Para la estimación efectiva de la contribución de señal se realiza un ajuste de máxima verosimilitud, en inglés *maximum likelihood* (ML), a los datos en la distribución de $m_{T2}(ll)$ con eventos de diferente e igual sabor de manera simultanea en todas las regiones de señal.

El modelado de la variable $m_{T2}(ll)$ y la normalización de los fondos principales ($t\bar{t}$, tW , WW) junto con los secundarios ($t\bar{t}Z$, WZ , ZZ , Drell-Yan) se analiza en regiones de control características haciendo uso de las distribuciones de datos. En la siguiente sección se resume brevemente el procedimiento realizado.

La contribución de otros fondos considerados minoritarios ($H \rightarrow WW$, $t\bar{t}W$, tri-bosones) en las regiones de señal es estimada a partir de la normalización a sus respectivas secciones eficaces y 35.9 fb^{-1} de luminosidad. El resto de procesos del Modelo Estándar tienen una presencia despreciable en este análisis.

9.8 Estimación de los fondos

Las principales contribuciones del Modelo Estándar en las regiones de señal vienen dadas por la producción de los procesos $t\bar{t}$, tW , y WW . La normalización de estos fondos se determina a través del ajuste ML a los datos. Debido a la baja contribución de los eventos de señal en los primeros bins de la distribución de $m_{T2}(ll)$, se puede aprovechar esta región para constreñir la tasa de eventos de cada uno de estos procesos en las regiones de señal con al menos un b jet, dominadas por $t\bar{t}$, tW , y en las

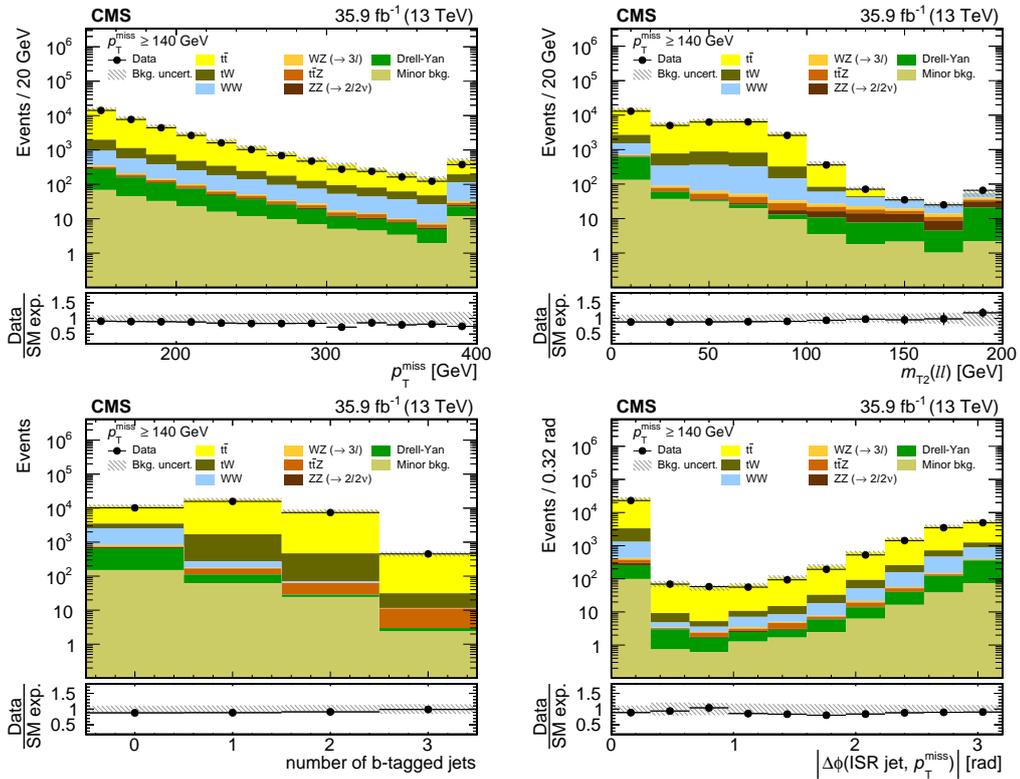


FIGURE 9.1: Distribuciones observadas y simuladas de algunos de los observables usados para definir las regiones de señal con eventos de dos leptones aislados de carga opuesta y $p_T^{\text{miss}} \geq 140$ GeV. Desde arriba a la izquierda y en el sentido de las agujas del reloj: p_T^{miss} , $m_{T2}(ll)$, $\Delta\phi$ entre \vec{p}_T^{miss} y el jet más energético (no etiquetado como b jet y con $p_T > 150$ GeV, los eventos que fallan estas condiciones se muestran en el primer bin de la distribución), y la multiplicidad de b jets en el evento. El último bin de cada distribución incluye las entradas con valores de la variable superiores al máximo representado en la figura. Las contribuciones de los fondos menores como $t\bar{t}W$, $H \rightarrow WW$, y la producción de tribosones son agrupados juntos. En el panel inferior se muestra el cociente entre el número de eventos observado y esperado. La banda sombreada representa la incertidumbre total (sistemática y estadística) en la previsión total de los procesos de fondo [19].

regiones de señal sin b jets donde el proceso WW se vuelve más relevante, mediante el ajuste.

Además de la normalización, se estudia el modelado de la distribución de $m_{T2}(ll)$ en diferentes regiones de control. Debido al límite natural en la masa de bosón W que manifiesta este observable, su forma es el principal discriminador entre la señal y estos procesos dominantes.

Se ha observado que aquellos eventos presentes en la región de $m_{T2}(ll)$ relevante para la extracción de señal ($m_{T2}(ll) \geq 80$ GeV) provienen principalmente de efectos de resolución del detector, siendo ésta una contribución difícil de modelar. Por esta razón se comparan las distribuciones de $m_{T2}(ll)$ simuladas para $t\bar{t}$, tW , y WW con las observadas en datos en diferentes regiones.

Se estudian las distribuciones de $m_{T2}(ll)$ simuladas para $t\bar{t}$, tW , y WW en dos regiones de control. La primera se construye modificando la selección común de referencia con el requisito $100 < p_T^{miss} < 140$ GeV. En esta CR los eventos se separan de acuerdo a la multiplicidad en el número de b jets, creando a su vez dos sub-regiones con diferente contenido de los procesos $t\bar{t}$, tW y de WW . Para eliminar el fondo de Drell-Yan, sólo se consideran eventos con pares de leptones de distinto sabor. La segunda CR se define en la región con $p_T^{miss} > 140$ GeV. Para no solapar con las regiones de señal, se seleccionan eventos con tres leptones utilizando la producción de $WZ \rightarrow 3\ell 1\nu$ para emular la distribución de $m_{T2}(ll)$ en eventos de WW y top ($t\bar{t}$, tW).

Se concluye que la simulación describe adecuadamente los datos en las regiones de control y se deriva una incertidumbre conservativa en base a la precisión estadística encontrada. Esta incertidumbre se aplica de manera descorrelacionada en los fondos de top y WW y afecta sólo a los últimos cuatro bins de la distribución de $m_{T2}(ll)$ siendo del 5, 10, 20 y 30%, respectivamente, sobre el fondo esperado.

Otra fuente potencial que puede afectar al modelado de la distribución de $m_{T2}(ll)$ es la presencia de leptones no genuinos procedentes entre otras causas de la desintegración semileptónica de un hadrón B dentro de un b jet o de un jet hadrónico erróneamente identificado como un lepton. Estos eventos no estarían limitados por el límite cinemático de la masa del bosón W y su contribución podría ser relevante en las colas de la distribución. Sin embargo, se ha encontrado que su valor es menor del 1 % del fondo esperado a lo largo de todas las regiones de señal, siendo más relevante sólo en altos valores de $m_{T2}(ll)$ y p_T^{miss} , dónde constituye hasta el 20% del fondo de $t\bar{t}$. Se ha estudiado la tasa de estos eventos en una región de control con dos leptones de misma carga y al menos un b jet. En base al acuerdo observado entre simulación y datos, se deriva un factor de corrección en la tasa de eventos simulados que tienen leptones no genuinos de 1.08 ± 0.21 .

La contribución de los fondos secundarios $t\bar{t}Z$, WZ , ZZ y Drell-Yan se estudia en regiones de control específicas donde se derivan los factores de corrección a sus respectivas normalizaciones.

La producción de eventos del proceso $t\bar{t}Z$ en los cuales los dos bosones W decaen leptónicamente y el bosón Z se desintegra en dos neutrinos, originan los mismos estados finales que la signatura experimental. Además, como en el caso de la señal, no presentan un límite cinemático en la distribución de $m_{T2}(ll)$, dada la contribución adicional de los neutrinos procedentes de la desintegración del bosón Z en la magnitud de \vec{p}_T^{miss} . La normalización de este fondo se estudia en una región de control con tres leptones, $p_T^{miss} > 140$ GeV, al menos dos jets de $p_T > 20$ GeV de los cuales

al menos uno es un b jet. Se requiere también la presencia de un par de leptones de carga opuesta y mismo sabor con una masa invariante que verifica $|m_{\ell\ell} - m_Z| \leq 10$ GeV. La comparación del número de eventos observados con el número de eventos esperados nos deja un factor de escala de normalización de 1.44 ± 0.36 para el proceso $t\bar{t}Z$.

Los eventos de la producción WZ pueden entrar en las regiones de señal cuando ambos bosones decaen leptónicamente y uno de los tres leptones producidos falla nuestra selección de veto en el tercer leptón. Queda así un estado final que puede ser aceptado si se satisfacen los requerimientos de selección básica del análisis. El modelado de esta fuente de fondo se estudia en una CR con tres leptones, $p_T^{miss} > 140$ GeV, y no b jets, de donde se deriva el factor de escala 0.97 ± 0.09 para la normalización del proceso.

El fondo ZZ está dominado por eventos con dos leptones y dos neutrinos provenientes de la desintegración de ambos bosones Z . Esta contribución se estudia en una región de control con cuatro leptones, $p_T^{miss} > 140$ GeV y sin b jets, donde la producción de $ZZ \rightarrow 2\ell 2\nu$ es emulada a través de sucesos $ZZ \rightarrow 4\ell$. Se requieren dos parejas de leptones de mismo sabor y carga opuesta, con masa invariante dentro de un rango de 30 GeV entorno la masa del boson Z y al menos uno de los dos pares en un rango de 15 GeV. Para imitar el sistema $ZZ \rightarrow 2\ell 2\nu$, el p_T de uno de los bosones Z reconstruidos, escogido aleatoriamente entre aquellos que satisfacen $|m_{\ell\ell} - m_Z| \leq 15$ GeV, es añadido vectorialmente al vector \vec{p}_T^{miss} . De la comparación entre el número de eventos observados y el número de eventos simulados se deriva el factor de escala de normalización 1.05 ± 0.12 usado en las regiones de señal de búsqueda del squark top. En las regiones de señal con y sin jets para la búsqueda del chargino, se estudia el modelado de la multiplicidad en jets del proceso ZZ . Debido a la dependencia encontrada entre el número de eventos en función del número de jets al comparar datos y simulación, se derivan de manera independiente, dos factores de escala en las regiones con al menos un jet, 1.21 ± 0.17 y sin jets, 0.74 ± 0.19 que corrigen la normalización del fondo ZZ en dicha búsqueda.

Los eventos de Drell-Yan pueden pasar la selección básica inicial debido a una mala medida del valor p_T^{miss} en el evento. El modelado de este fondo se estudia en eventos sin b jets y con exclusivamente dos leptones, de carga opuesta y mismo sabor, que satisfacen $|m_{\ell\ell} - m_Z| \leq 15$ GeV (eventos del bosón Z). Los sucesos con $100 < p_T^{miss} < 140$ GeV están dominados por la producción de Drell-Yan y son usados para derivar una corrección al modelado de la distribución $m_{T2}(ll)$, la cuál es posteriormente examinada en eventos con $p_T^{miss} > 140$ GeV. Dicha corrección varía entre un pequeño porcentaje a baja $m_{T2}(ll)$ hasta aproximadamente el 50% para $m_{T2}(ll) > 100$ GeV. Se establece también una incertidumbre global de normalización del 32% en base a la discrepancia observada entre el número de eventos observados y simulados con $100 < p_T^{miss} < 140$ GeV. Finalmente, las predicciones de Drell-Yan para eventos sin jets es verificada en eventos del bosón Z con $p_T^{miss} > 140$ GeV con cero jets. Se aplica una incertidumbre conservativa del 100% en esta contribución. Puesto que la producción de Drell-Yan es un fondo subdominante en las regiones de señal sin jets, esta incertidumbre tiene un impacto mínimo en la sensibilidad esperada para la producción de señal en este análisis.

A la estimación de los fondos minoritarios se le aplica conservativamente una incertidumbre de normalización del 50%. Las incertidumbres en los factores de escala incluyen la incertidumbre estadística de los eventos simulados y observados, además de la incertidumbre sistemática total en el número de eventos esperados de

las contribuciones de los procesos del Modelo Estándar residuales en las regiones de control. Se encuentra insignificante la contribución de señal en cualquiera de las CRs usadas.

9.9 Incertidumbres sistemáticas

Tanto a la normalización como a la forma de la distribución de $m_{T2}(ll)$ están afectadas por varias fuentes de incertidumbre sistemática. Aquellas que conciernen igualmente a eventos de fondo y señal son: la estimación de la luminosidad integrada, la medida de la eficiencia de selección de trigger, la eficiencia de identificación y aislamiento de los leptones y la eficiencia de identificación de los b jets. Así mismo la escala de energía de los jets, y de las partículas de bajo momento que no han sido agrupadas en ningún jet son también fuente de incertidumbre en la medida. El efecto del tamaño de las muestras está asociado a la incertidumbre estadística. Y finalmente son también incluidas las incertidumbres teóricas asociadas a la elección del conjunto de PDFs (no se aplica en eventos de señal, ver más abajo), a la escala de renormalización y a la escala de factorización. Son estas dos últimas, junto con la incertidumbre en la escala de energía de los jets, las que más afectan al número de eventos esperados en los procesos del Modelo Estándar. Mientras que en el número de eventos de señal se observa un mayor impacto de la incertidumbre relativa a las eficiencias de identificación y aislamiento de los leptones. Las fuentes de incertidumbre que más afectan al modelado de la distribución de $m_{T2}(ll)$, son la escala de energía de los jets y de las partículas de bajo momento, presumiblemente por su alta correlación con la variable \vec{p}_T^{miss} . El mayor impacto en la shape de $m_{T2}(ll)$ para la señal viene de la incertidumbre en el modelado de p_T^{miss} .

Las fuentes de incertidumbres sistemática que afectan específicamente a los procesos del Modelo Estándar incluyen además: la precisión obtenida en el número de eventos estimado (<4%) cuyo mayor impacto en la forma de $m_{T2}(ll)$ se encuentra en el proceso de Drell-Yan, el modelado en la forma de la distribución de $m_{T2}(ll)$ de los procesos $t\bar{t}$, tW y WW , siendo la asociada a los eventos de top ($t\bar{t}$, tW) la más importante (4-18%), y finalmente las correspondientes correcciones aplicadas en el caso de los fondos ZZ (<3%) y Drell-Yan (4-13%). De manera adicional, se deriva una incertidumbre relativa al espectro de p_T del quark top en eventos de $t\bar{t}$ para cuantificar la discrepancia observada en datos y simulación [168–170]. El efecto relativo tanto en el número de eventos como en la forma de la distribución $m_{T2}(ll)$ (1-4% y 1-8% respectivamente) es del orden de las fuentes mencionadas anteriormente.

Por último, la reconstrucción de eventos con FastSim afecta al modelado de la muestra de señal. Las fuentes de incertidumbres sistemáticas inducidas particularmente por este tipo de simulación afectan a la eficiencia de identificación y aislamiento de los leptones, a la eficiencia de identificación de los b jets, al modelado de la variable \vec{p}_T^{miss} , al modelado de las interacciones de pileup, y finalmente el modelado de los eventos con jets ISR. Las incertidumbres relacionadas con la elección del conjunto de PDFs no se aplican a los eventos de señal pues resultan redundantes tras incluir la incertidumbre asociada a los eventos con jets ISR. De todas ellas, la más importante respecto a la predicción del número de eventos es la obtenida en la medida de eficiencia de identificación y aislamiento de los leptones (4%), siendo la reconstrucción de la variable \vec{p}_T^{miss} la que más afecta a la forma de la distribución (6-25%).

9.10 Resultados e interpretación

Para la estimación del número de eventos de señal se utiliza el método de máxima verosimilitud empleando simultáneamente las distribuciones de $m_{T2}(ll)$ para datos y simulación en todas las regiones de señal.

Las incertidumbres que afectan a la medida son introducidas en el ajuste a través de parámetros de error que obedecen a una distribución de probabilidad inicial determinada. En el caso de las incertidumbres que afectan a la normalización de la señal y del fondo, la distribución de probabilidad es una normal logarítmica así como en el caso de las que afectan directamente a la forma de $m_{T2}(ll)$ se utiliza una distribución gaussiana. Los resultados del ajuste se muestran en las Tablas 9.1-9.2, donde el número de eventos esperados se ha obtenido considerando la hipótesis de ausencia de señal. No se observa ningún exceso en datos sobre la predicción del Modelo Estándar.

Se establece, así mismo, un límite superior en la sección eficaz de producción para los diferentes modelos considerados. El procedimiento se basa una aproximación asintótica del criterio CLs [171–173] que fija dicho límite con un nivel de confianza, en inglés *confidence level* (CL), del 95%.

$m_{T2}(ll)$ [GeV]		0–20	20–40	40–60	60–80	80–100	100–120	≥ 120
DF events								
SR1 ^{jets} _{0tag}	Predicted	1493 ± 32	558 ± 12	719 ± 16	730 ± 16	316 ± 10	45.1 ± 3.1	13.7 ± 2.8
	Observed	1484	532	732	725	298	47	13
SR1 ^{0jet} _{0tag}	Predicted	41.9 ± 5	27.4 ± 3.8	34.1 ± 4.8	42 ± 5.5	21.1 ± 3.4	6 ± 1.3	7.9 ± 2.1
	Observed	39	24	33	44	13	6	9
SR2 ^{jets} _{0tag}	Predicted	534 ± 15	158.6 ± 5.9	167.9 ± 6.1	157.9 ± 6.5	42.4 ± 2.9	5.9 ± 1	9 ± 1.7
	Observed	511	162	156	176	43	5	9
SR2 ^{0jet} _{0tag}	Predicted	10.3 ± 1.7	7 ± 1.5	6.5 ± 1.3	6.9 ± 1.3	2.19 ± 0.69	1.59 ± 0.7	7.8 ± 1.8
	Observed	10	4	4	6	2	2	7
SR3 _{0tag}	Predicted	127.9 ± 7.2	28.3 ± 2	30.2 ± 2.4	23.1 ± 2	4.96 ± 0.73	1.12 ± 0.38	4.5 ± 1.2
	Observed	116	35	29	21	3	1	5
SF events								
SR1 ^{jets} _{0tag}	Predicted	1310 ± 29	499 ± 12	623 ± 14	634 ± 15	271.7 ± 8.9	51.6 ± 3.5	48.6 ± 5.5
	Observed	1324	499	609	659	284	57	47
SR1 ^{0jet} _{0tag}	Predicted	44.1 ± 7.5	28.5 ± 4.1	33.5 ± 4.4	33.5 ± 4.5	18.6 ± 2.6	7.7 ± 1.6	12.5 ± 2.5
	Observed	43	40	39	33	17	6	12
SR2 ^{jets} _{0tag}	Predicted	474 ± 14	134.8 ± 5.1	155.1 ± 5.5	128.5 ± 5.5	37.1 ± 2.5	7.29 ± 0.91	23.9 ± 2.4
	Observed	493	123	166	118	33	7	25
SR2 ^{0jet} _{0tag}	Predicted	10.9 ± 1.9	7.8 ± 1.8	7.3 ± 1.4	7.9 ± 1.3	1.9 ± 0.52	1.28 ± 0.58	7.1 ± 1.4
	Observed	8	12	11	10	3	2	7
SR3 _{0tag}	Predicted	112.8 ± 6.3	27.9 ± 2.2	24.2 ± 1.8	22.5 ± 1.8	5.2 ± 1	1.36 ± 0.36	10.6 ± 1.2
	Observed	110	35	26	26	2	1	14

TABLE 9.1: Número observado y esperado de eventos con pares de leptones de distinto sabor (mitad superior de la tabla) y de igual sabor (mitad inferior) en las regiones de señal de búsqueda del chargino. La incertidumbre sobre el total de la predicción en los procesos del fondo alude al error estadístico y sistemático total [19].

En la Figura 9.2 (izq.) se muestran, con un nivel de confianza del 95%, los límites superiores de la sección eficaz de producción de una pareja de charginos decayendo en sleptones. Las cadenas $\tilde{\chi}_1^\pm \rightarrow \tilde{l} \nu \rightarrow l \nu \tilde{\chi}_1^0$ y $\tilde{\chi}_1^\pm \rightarrow \tilde{\nu} l \rightarrow l \nu \tilde{\chi}_1^0$ tienen una fracción de desintegración del 50% cada una y se asume que los sleptones están degenerados, con masa igual al promedio de las masas del chargino y del neutralino. Compa-

$m_{T2}(ll)$ [GeV]		0–20	20–40	40–60	60–80	80–100	100–120	≥ 120
DF events								
SR1 _{tags}	Predicted	3525 ± 80	1505 ± 31	1958 ± 42	2049 ± 46	897 ± 22	108.4 ± 7.3	13.4 ± 2.2
	Observed	3534	1494	1938	2068	879	111	15
SR1 _{0tag}	Predicted	1542 ± 33	588 ± 13	756 ± 15	771 ± 19	338.3 ± 9.3	50.6 ± 3.8	21 ± 3.8
	Observed	1523	556	765	769	311	53	22
SR2 _{tags}	Predicted	1036 ± 37	363 ± 13	415 ± 14	377 ± 14	105.1 ± 6.5	12.3 ± 2	5.02 ± 0.82
	Observed	1045	357	412	389	111	11	1
SR2 _{0tag}	Predicted	545 ± 18	164.3 ± 7.3	173.2 ± 6.2	165.1 ± 6.8	44.8 ± 3.1	7.1 ± 1.4	15.5 ± 3
	Observed	521	166	160	182	45	7	16
SR3 ^{ISR} _{tags}	Predicted	152.1 ± 9.9	35.5 ± 2.7	32.3 ± 2.3	25 ± 2.2	4.67 ± 0.77	0.41 ± 0.38	0.41 ± 0.26
	Observed	133	44	36	26	2	1	0
SR3 ^{ISR} _{0tag}	Predicted	103.9 ± 6.8	21.3 ± 1.9	22.2 ± 2.1	15.4 ± 1.6	3.51 ± 0.6	0.53 ± 0.21	0.53 ± 0.34
	Observed	100	27	22	12	3	0	1
SF events								
SR1 _{tags}	Predicted	2979 ± 68	1277 ± 30	1644 ± 35	1712 ± 37	762 ± 19	91.9 ± 6.1	18.1 ± 2.1
	Observed	3003	1266	1674	1671	798	85	16
SR1 _{0tag}	Predicted	1350 ± 33	526 ± 13	656 ± 15	670 ± 17	289.2 ± 7.6	57.9 ± 4.2	61.8 ± 5.8
	Observed	1367	539	648	692	301	63	59
SR2 _{tags}	Predicted	888 ± 30	319 ± 12	363 ± 14	323 ± 13	90.5 ± 5.5	10.8 ± 1.5	7.43 ± 0.98
	Observed	900	315	343	325	86	13	11
SR2 _{0tag}	Predicted	487 ± 16	140.7 ± 5.5	161.9 ± 5.9	134.5 ± 6.2	39.6 ± 2.7	8.1 ± 1.1	30.6 ± 3
	Observed	501	135	177	128	36	9	32
SR3 ^{ISR} _{tags}	Predicted	129.6 ± 8.9	29.6 ± 2.1	27.8 ± 2.1	22.2 ± 1.9	3.71 ± 0.57	0.47 ± 0.42	0.71 ± 0.38
	Observed	123	27	28	38	4	1	1
SR3 ^{ISR} _{0tag}	Predicted	91.5 ± 6.1	20.1 ± 1.8	16.5 ± 1.4	13.7 ± 1.4	3.14 ± 0.58	0.78 ± 0.36	1.63 ± 0.42
	Observed	92	26	17	12	1	1	2

TABLE 9.2: Número observado y esperado de eventos con pares de leptones de distinto sabor (mitad superior de la tabla) y de igual sabor (mitad inferior) en las regiones de señal de búsqueda del squark top. La incertidumbre sobre el total de la predicción en los procesos del fondo alude al error estadístico y sistemático total [19].

rando estos límites superiores con el valor teórico de la sección eficaz de producción $pp \rightarrow \tilde{\chi}_1^\pm \tilde{\chi}_1^\mp$ en cada modelo, se han determinado las regiones de exclusión observada y esperada en el plano de masas ($m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0}$). Se han logrado excluir masas con valores próximos a los 800 y 320 GeV para el chargino y el neutralino respectivamente, siendo los límites más altos que se han obtenido para este modelo hasta la fecha. Por otro, lado, se ha encontrado una sensibilidad limitada en modelos dónde el chargino decae en un bosón W y un neutralino, debido a la fracción de desintegración relativamente baja de la desintegración leptónica de un bosón W . Para este escenario de señal se han derivado los límites superiores de la sección eficaz de producción asumiendo que el neutralino más ligero tiene una masa de 1 GeV. Los límites superiores observados y esperados con un nivel de confianza del 95% se muestran en función de la masa del chargino en la figura 9.2 (dcha.) donde son comparados con la sección eficaz teórica. Se logran excluir masas de chargino en el rango de los 170 a 200 GeV.

En la Figura 9.3 se muestran, con un nivel de confianza del 95%, los límites superiores de la sección eficaz de producción de una pareja squarks top usando los dos modelos de referencia. Mientras que la estrategia del análisis ha sido optimizada para la zona de masa comprimida, por completitud, los resultados se presentan en todo el plano de masas ($m_{\tilde{t}_1}, m_{\tilde{\chi}_1^0}$). Cuando consideramos la zona comprimida, donde Δm está entre la masa del quark top y del bosón W , y la desintegración del squark top en

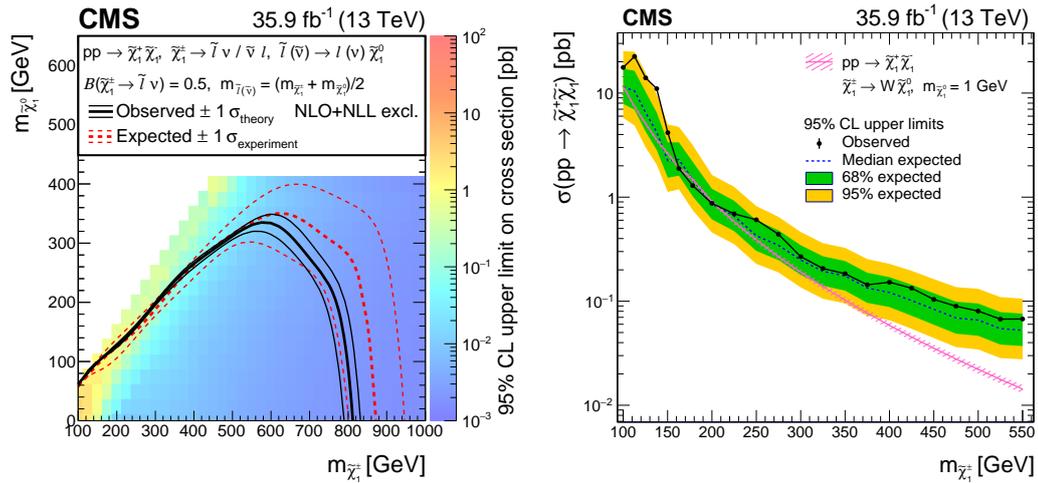


FIGURE 9.2: En la izquierda, los límites superiores a 95% CL en la sección eficaz de producción de una pareja de charginos como función de las masas del neutralino y del chargino, cuando el chargino tiene la cadena de desintegración $\tilde{\chi}_1^\pm \rightarrow \tilde{l} \nu (\bar{\nu} l) \rightarrow l \nu \tilde{\chi}_1^0$. Las regiones de exclusión en el plano $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0})$ son determinadas comparando los límites superiores con las secciones eficaces de producción calculadas a orden NLO+NLL. La línea punteada de color rojo más gruesa muestra la región esperada de exclusión. Las líneas punteadas de color rojo más finas muestran la variación de las regiones de exclusión debido a las incertidumbres experimentales. La línea negra continua más gruesa muestra la región de exclusión observada, mientras que las líneas continuas negras y más finas muestran la variación de las regiones de exclusión debido a las incertidumbres teóricas en la sección eficaz de producción. Derecha: límites superiores esperados y observados con un 95% CL como función de la masa del chargino siendo la masa del neutralino de 1 GeV y asumiendo que el chargino decae en un neutralino y un bosón W ($\tilde{\chi}_1^\pm \rightarrow W \tilde{\chi}_1^0$) [19].

un quark top y en un neutralino los valores de masa excluidos alcanzan los 420 y 360 GeV respectivamente. En el caso del modo de decaimiento $\tilde{t}_1 \rightarrow b \tilde{\chi}_1^\pm \rightarrow b W \tilde{\chi}_1^0$, se establece un valor mínimo $\Delta m \approx 2m_W$ debido a la asunción de que $m_{\tilde{\chi}_1^\pm} = (m_{\tilde{t}_1} + m_{\tilde{\chi}_1^0})/2$. Para $\Delta m \approx 2m_W$, la masa del quark top es excluida en el rango de 225-325 GeV. La región sin cubrir en la Figura 9.3 (dch.) entorno a la masa del squark top de 200 GeV corresponde a un espacio de fases similar al del fondo $t\bar{t}$ y donde la contribución de los neutralinos a la p_T^{miss} es muy baja. En esta situación, la incertidumbre en el modelado de la p_T^{miss} en eventos de FastSim se vuelve demasiado grande como para proveer de sensibilidad alguna a la región ante la posible existencia de señal.

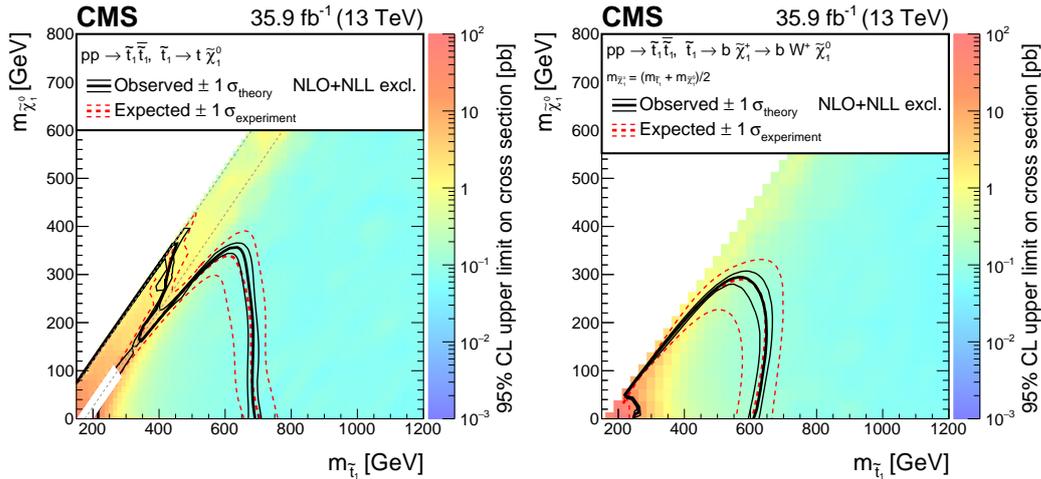


FIGURE 9.3: Límites superiores a 95% CL de la sección eficaz de producción de una pareja de squarks top en función de las masas de squark top y del neutralino. En la imagen izquierda, se muestran los resultados para el modelo de referencia donde el squark top decae en un quark top y un neutralino. Las dos líneas diagonales grises encierran la región de masa comprimida $m_W < m_{\tilde{t}_1} - m_{\tilde{\chi}_1^0} \lesssim m_t$. La imagen de la derecha muestra los límites superiores para el segundo modelo de referencia donde el squark top decae en un quark bottom y un chargino, el cuál a su vez decae en un bosón W y un neutralino. Se asume que la masa del chargino es igual al promedio de las masas del squark top y del neutralino. Las regiones de exclusión en el plano de masas $(m_{\tilde{t}_1}, m_{\tilde{\chi}_1^0})$ son determinadas comparando los límites superiores con las secciones eficaces de producción calculadas a orden NLO+NLL. La línea punteada de color rojo más gruesa muestra la región esperada de exclusión. Las líneas punteadas de color rojo más finas muestran la variación de las regiones de exclusión debido a las incertidumbres experimentales. La línea negra continua más gruesa muestra la región de exclusión observada, mientras que las líneas continuas negras y más finas muestran la variación de las regiones de exclusión debido a las incertidumbres teóricas en la sección eficaz de producción [19].

En resumen, con el estudio de los datos recogidos por el experimento CMS en 2016, esta búsqueda extiende los actuales límites de exclusión en la producción de parejas de charginos decayendo en sleptones [58], mejorando en aproximadamente 70 GeV el límite en la masa del chargino para un neutralino sin masa. Los límites de exclusión de la producción de parejas de squark top extienden los resultados obtenidos por la Colaboración CMS en estados finales con dos leptones de carga opuesta [59] en

la zona de masa comprimida, siendo éstos competitivos con los resultados obtenidos por la Colaboración ATLAS en el mismo canal de desintegración [65].

Durante 2017 y 2018 la Colaboración CMS ha incrementado sustancialmente la muestra de datos disponible. Se espera que este aumento de estadística junto con nuevas técnicas de predicción de fondos favorezcan a todas las búsquedas de nueva física. En particular, dentro del alcance de este análisis se espera que la combinación de los datos recogidos en 2016, 2017 y 2018 mejore los actuales límites de exclusión sino afirme el descubrimiento. Se han planteado diversas estrategias de optimización para incrementar la sensibilidad de las búsquedas. Estas están principalmente centradas en la selección de eventos y la definición de zonas de señal.

Appendix A

Data and Monte Carlo simulated samples

The data event samples used in the analysis have been collected from pp collisions at centre of energy mass $\sqrt{s}=13$ TeV by the CMS detector during the 2016 data taking period at LHC. The dataset is characterized by the presence of a single lepton (e, μ) or a lepton pair ($ee, e\mu$) in the event. The Table [A.1](#) lists all the collection according to the 2016 run periods. The integrated luminosity shown for each run period corresponds to a total integrated luminosity of 35.9 ± 0.9 fb⁻¹. From the oldest to newest these are: Run2016B, Run2016C, Run2016D, Run2016E, Run2016F, Run2016G, and Run2016H.

The MC simulated samples of the SM processes used as background and those considered as signals in the analysis are presented in Tables [A.2](#) and [A.3](#) respectively. The SM samples are simulated with MADGRAPH, POWHEG, and PYTHIA, while the signals are produced using only MADGRAPH. The underlying event tune for the processes is also shown.

These data and simulated samples have been stored in miniAOD format. This is high-level data tier introduced in Spring 2014 in order to satisfy the needs of the mainstream physics analyses while keeping a small event size (30-50 kb/event). For further details see [CMS 2016 MiniAOD Analysis Documentation](#) and [run period developments](#).

Period	Dataset	Events	Luminosity [fb^{-1}]
Run2016B	SingleMuon_ver2-v2	158,145,722	5.78
	SingleElectron_ver2-v2	246,440,440	
	MuonEG_ver2-v2	32,727,796	
	DoubleMuon_ver2-v2	82,535,526	
	DoubleEG_ver2-v2	143,073,268	
Run2016C	SingleMuon-v1	67,441,308	2.56
	SingleElectronC-v1	97,259,854	
	MuonEG-v1	15,405,678	
	DoubleMuon-v1	27,934,629	
	DoubleEG-v1	47,677,856	
Run2016D	SingleMuon-v1	98,017,996	4.25
	SingleElectron-v1	148,167,727	
	MuonEG-v1	23,482,352	
	DoubleMuon-v1	33,861,745	
	DoubleEG-v1	53,324,960	
Run2016E	SingleMuon-v1	90,963,495	4.01
	SingleElectron-v1	117,321,545	
	MuonEG-v1	22,519,303	
	DoubleMuon-v1	28,246,946	
	DoubleEG/Run2016E-v1	49,877,710	
Run2016F	SingleMuon-v1	65,489,554	3.10
	SingleElectron-v1	70,593,532	
	MuonEG-v1	16,002,165	
	DoubleMuon-v1	20,329,921	
	DoubleEG-v1	34,577,629	
Run2016G	SingleMuon-v1D	149,916,849	7.54
	SingleElectron-v1	153,330,123	
	MuonEG-v1	33,854,612	
	/DoubleMuon-v1	45,235,604	
	DoubleEG-v1	78,764,716	
Run2016H	SingleMuon_ver2-v1	169,642,135	8.39
	SingleElectron_ver2-v1	125,826,667	
	MuonEG_ver2-v1	28,466,022	
	DoubleMuon_ver2-v1	47,693,168	
	DoubleEG_ver2-v1	83,361,083	
Run2016H	SingleMuon_ver3-v1	4,393,029	0.22
	SingleElectron_ver3-v1	3,191,585	
	MuonEG_ver3-v1D	770,494	
	DoubleMuon_ver3-v1	1,219,644	
	DoubleEG_ver3-v1	2,027,651	

TABLE A.1: Data event samples with a single muon(SingleMuon), single electron (SingleElectron), or a lepton pair of electron and muon (MuonEG) , double muon (DoubleMuon), and double electron (DoubleEG) . The integrated luminosity of each run period corresponds to a total integrated luminosity of $35.9 \pm 0.9 \text{ fb}^{-1}$. From oldest to newest: Run2016B, C, D, E, F, G, and H.

Process	Sample	σ [pb]	Events/M
Top backgrounds			
$qq \rightarrow t\bar{t} \rightarrow 2\ell 2\nu 2b$	/TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_[1]	87.310	79
$b\bar{g} \rightarrow \bar{t}W^- \rightarrow X$	/ST_tW_antitop_5f_inclusiveDecays_[1]_TuneCUETP8M1 (_ext1-v1)	35.6	6.9
	/ST_tW_top_5f_inclusiveDecays_TuneCUETP8M1_[1] (_ext1-v1)	35.6	7.0
$qq \rightarrow t\bar{t}W+ \text{ Jets} \rightarrow \ell\nu$	/TTWJetsToLNu_TuneCUETP8M1_[3]_madspin (_ext1-v3)	0.2043	2
$t\bar{t}W+ \text{ Jets} (W \rightarrow qq')$	/TTWJetsToQQ_TuneCUETP8M1_[3]_madspin	0.40620	0.8
$g\bar{g} \rightarrow t\bar{t}Z (Z \rightarrow 2q)$	/TTZToQQ_TuneCUETP8M1_[2]	0.5297	0.7
$t\bar{t}Z \rightarrow 2\ell 2\nu$	/TTZToLLNuNu_M-10_TuneCUETP8M1_[2] (_ext1-v1)	0.253	2.0
Dibosons			
$qq \rightarrow WW \rightarrow 2\ell 2\nu$	/WWTo2L2Nu_13TeV-powheg	12.178	2
$g\bar{g} \rightarrow WW \rightarrow 2\ell 2\nu$	/GluGluWWTo2L2Nu_MCFM_13TeV	0.5905	0.5
$qq \rightarrow WZ \rightarrow 3\ell 1\nu$	/WZTo3LNu_TuneCUETP8M1_[1]	4.42965	2
$WZ \rightarrow 2\ell 2q$	/WZTo2L2Q_[3]_madspin	5.5950	26
$qq \rightarrow ZZ \rightarrow 2\ell 2q$	/ZZTo2L2Q_[3]_madspin	3.22	15
$ZZ \rightarrow 2\ell 2\nu$	/ZZTo2L2Nu_[1]	0.564	9
$g\bar{g} \rightarrow ZZ \rightarrow 2\ell 2\nu$	/GluGluToContinToZZTo2e2nu_13TeV_MCFM701_pythia8	0.001720	0.5
	/GluGluToContinToZZTo2mu2nu_13TeV_MCFM701_pythia8	0.001720	0.5
Drell-Yan			
$qq \rightarrow Z/\gamma^* \rightarrow \bar{\ell}\ell + \text{ Jets}$	/DYJetsToLL_M-10to50_TuneCUETP8M1_[3]	18610.0	30
	/DYJetsToLL_M-50_TuneCUETP8M1_[3] (_ext2-v1)	6025.2	122.0
	/DYJetsToLL_M-10to50_TuneCUETP8M1_[4]	18610.0	35
	/DYJetsToLL_M-5to50_HT-XXXtoYYY_TuneCUETP8M1_[4]	-	-
	/DYJetsToLL_M-50_TuneCUETP8M1_[4] (_ext1-v2)	6025.2	49
	/DYJetsToLL_M-50_HT-XXXtoYYY_TuneCUETP8M1_[4]	-	-
Tribosons			
$qq \rightarrow WWW \rightarrow X$	/WWW_4F_TuneCUETP8M1_[2]	0.18331	0.24
$qq \rightarrow WWZ \rightarrow X$	/WWZ_TuneCUETP8M1_[2]	0.16510	0.25
$qq \rightarrow WZZ \rightarrow X$	/WZZ_TuneCUETP8M1_[2]	0.05565	0.25
Higgs production			
$g\bar{g} \rightarrow H \rightarrow WW \rightarrow 2\ell 2\nu$	/GluGluHToWWTo2L2Nu_M125_[3]	0.3128	0.1
$H \rightarrow \tau\tau$	/GluGluHToTauTau_M125_[1]	2.7757	1.5
$VBF \rightarrow H \rightarrow WW \rightarrow 2\ell 2\nu$	/VBFHToWWTo2L2Nu_M125_[1]_JHUGenv628	0.0580	0.1
$H \rightarrow \tau\tau$	/VBFHToTauTau_M125_[1]	0.237	1.5
$qq \rightarrow HW^+ \rightarrow WWW^+$	/HWplusJ_HToWW_M125_[1]	0.1810	0.3
$qq \rightarrow HW^- \rightarrow WWW^-$	/HWminusJ_HToWW_M125_[1]	0.1160	0.3
Non-prompt leptons			
$t\bar{t} \rightarrow 1\ell 1\nu 2q$	/TTToSemilepton_TuneCUETP8M2_ttHtranche3_[1]	364.35	152
$qq \rightarrow W + \text{ Jets} \rightarrow \ell\nu + \text{ Jets}$	/WJetsToLNu_TuneCUETP8M1_[3]	61526.7	24.0

[1] = _13TeV-powheg-pythia8. [2] = _13TeV-amcatnlo-pythia8. [3] = _13TeV-amcatnloFXFX-pythia8. [4] = _13TeV-madgraphMLM-pythia8.

TABLE A.2: Summary of simulated SM processes used as background in this analysis, together with cross section and the number of processed events. For dedicated decay samples the cross section times branching ratio value is shown.

Process	Sample	Events/M
squark top pair ($pp \rightarrow \tilde{t}_1 \tilde{t}_1$)		
$\tilde{t}_1 \rightarrow t$	/SMS-T2tt_mStop-150to250_ [*]	32.28
	/SMS-T2tt_mStop-250to350_ [*]	33.86
	/SMS-T2tt_mStop-350to400_ [*]	30.41
	/SMS-T2tt_mStop-400to1200_ [*]	29.70
$\tilde{t}_1 \rightarrow b \tilde{\chi}_1^\pm \rightarrow b W \tilde{\chi}_1^0$	/SMS-T2bW_ [*]	30.65
chargino pair ($pp \rightarrow \tilde{\chi}_1^\pm \tilde{\chi}_1^\mp$)		
$\tilde{\chi}_1^\pm \rightarrow W \tilde{\chi}_1^0$	/SMS-TChipmWW_WWTo2LNu_ [*]	2.62
$\tilde{\chi}_1^\pm \rightarrow \tilde{l} \nu (\tilde{\nu} l) \rightarrow l \nu \tilde{\chi}_1^0$	/SMS-TChipmSlepSnu_ [*]	4.66
[*] = _TuneCUETP8M1_13TeV-madgraphMLM-pythia8.		

TABLE A.3: Simulated signal processes, together with the number of processed events. The samples are generated with MADGRAPH and PYTHIA. For dedicated decay samples the cross section times branching ratio value is shown.

Appendix B

Impact of the jet rate modeling on the chargino search

The chargino pair search relies on signal regions with no b-tagged jets split depending on the jet multiplicity into two sub-regions, with or without jets in the event. In the control region with no b-tagged jets and $100 < p_T^{miss} < 140$ GeV events, the observed jet multiplicity distribution is generally well described by simulation as it is shown in Figure B.1. However, after requiring higher values of $p_T^{miss} (> 140$ GeV), the rate modeling of events with different jet multiplicity is more likely to be compromised by the presence of extra jets from ISR emissions. These radiative processes are hard to simulate and are highly correlated with p_T^{miss} variable. Due to the fact that signal regions rely on the jet multiplicity and p_T^{miss} selections the effect of jet rate mismodeling on the normalization of the main backgrounds is considered.

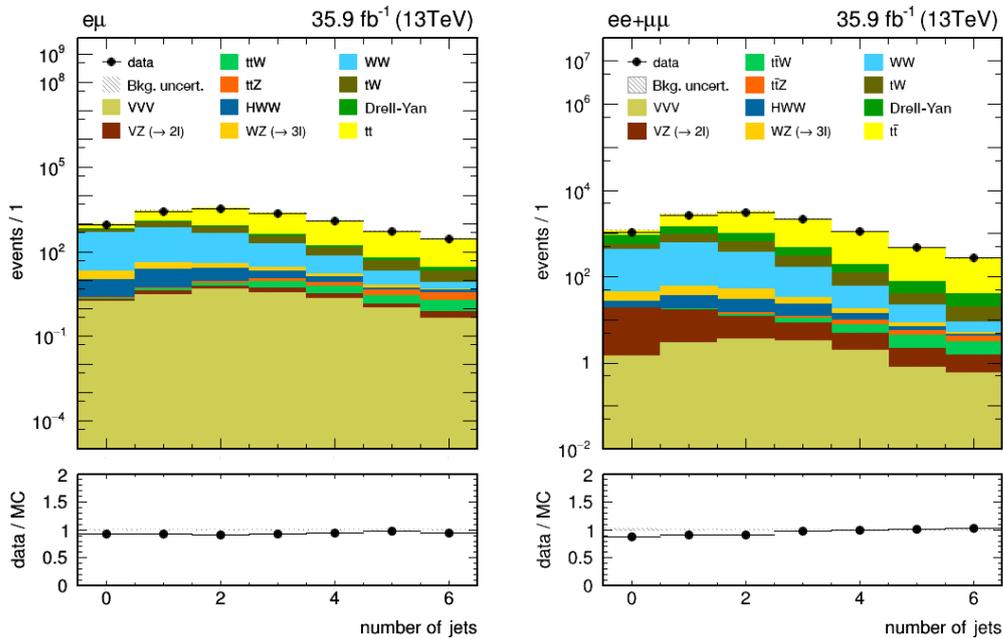


FIGURE B.1: Observed and expected jet multiplicity distributions in events with $100 < p_T^{miss} < 140$ GeV and no b-tagged jets for the $e\mu$ (left) and $ee + \mu\mu$ (right) channels. The names VVV and VZ stand for triboson and diboson production, respectively. The lower panel shows the ratio between data and total SM contribution in each bin. Only statistic uncertainties are shown.

The impact of the jet rate modeling of events with no jets can vary with the type of process under consideration. The normalization uncertainties to the expected events with and without jets for ZZ and Drell-Yan backgrounds have been already discussed in Sections 6.4.4 and 6.4.5 respectively. To study the effect of the jet multiplicity on diboson contributions (WW and WZ processes), this analysis resorts again to the modeling of WW background through $WZ \rightarrow 3l1\nu$ events, as describe in Section 6.4.1, to validate the $m_{T2}(ll)$ shape at $p_T^{miss} > 140$ GeV. In Figure B.2 the jet multiplicity distribution is shown for events with different content of jets. On the left, events with up to ten jets are arranged in ten different bins, while on the right, these events are grouped in only two categories, zero jets and at least one jet. From the value and uncertainty of data and simulation ratio in the 0-jet and at-least-one-jet bins (B.2, right), it can be observed that the fraction of the events with no jets is modeled with at worst a 30% inaccuracy in this control region.

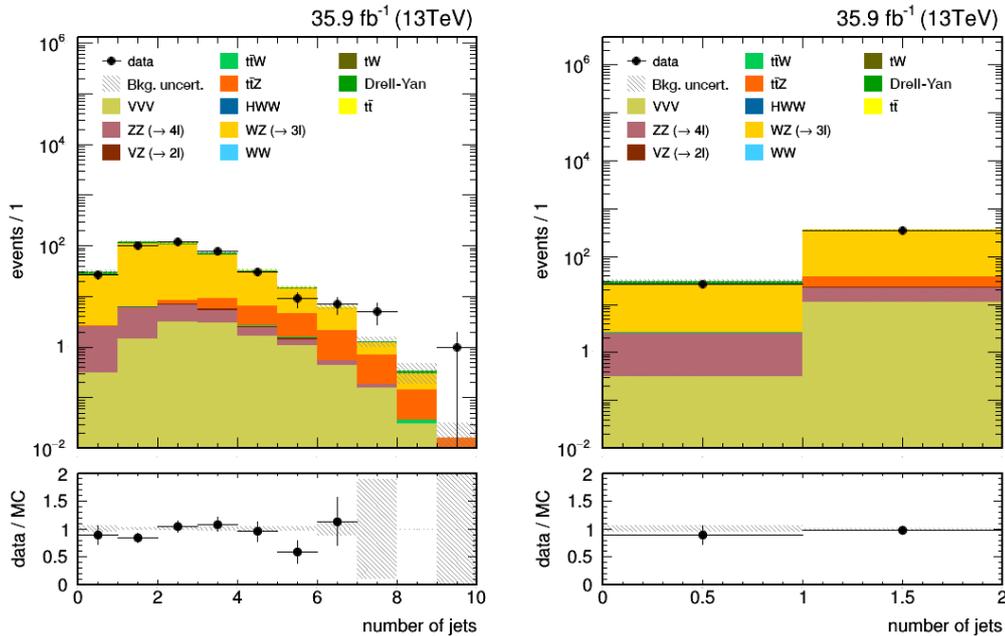


FIGURE B.2: Observed and expected jet multiplicity distributions of events with three leptons, no b-tagged jets and $p_T^{miss} > 140$ GeV. That lepton of the pair coming from the reconstructed Z boson decay with same charge as the third one has treated as lost particle and added vectorially to the p_T^{miss} of the event. On the left, the distribution shows bins up to ten jets in the event content, while on the right, all events are displayed in two categories, zero jets and at least one jet.

The lower panel shows the ratio between data and total SM contribution in each bin. Only statistic uncertainties are shown.

On the other hand, the contribution of events with no jets from top related backgrounds ($t\bar{t}$, tW , $t\bar{t}W$, $t\bar{t}Z$) is relatively rare but not negligible. Regarding the natural jet multiplicity of this type of backgrounds, finding specific CRs of events with no jet at high p_T^{miss} is not easily accessible for each one. In this analysis, the dominant of this kind of backgrounds is $t\bar{t}$, whose modeling in the 0-jet region has been studied and conservatively it has been estimated that the maximum discrepancy between the fraction of simulated events with no jets regarding the data events rises up to 50%. As a result, a common normalization uncertainty of 50% is considered for the jet enriched backgrounds.

The two uncertainties are introduced in the analysis by adding two normalization parameters in the ML fit as described in Section 7.1. The rate of events without jets is allowed to float within the estimated uncertainties, 30% for the diboson and 50% for jet enriched backgrounds. The variations are treated independently in the p_T^{miss} bins. The total sum of the expected yields for each process in the two regions (with and without jets) is constrained to remain invariant, so that, only a migration of events between the two regions is allowed.

Appendix C

Tables of systematic uncertainties

The range of the systematic uncertainties on the total SM prediction is given across the different signal regions for the top squark search in Tables C.5 and C.6 and for the chargino search in Tables C.1 and C.2. The range of statistical uncertainty along the $m_{T2}(ll)$ bins in each SR is also give for comparison.

The range of the specific systematic uncertainties on the predicted yields of top squark signal is given for a representative mass point in Tables C.8 and C.7, and for a representative chargino mass point in Tables C.3 to C.4.

Systematic	SR1 ^{0Jet} _{0Tag}	SR1 ^{Jets} _{0Tag}	SR2 ^{0Jet} _{0Tag}	SR2 ^{Jets} _{0Tag}	SR3 _{0Tag}
Integrated luminosity	2.5%	2.5%	2.5%	2.5%	2.5%
Trigger	2%	2%	2%	2%	2%
Pileup	3.7%	3.3%	3.6%	3.3%	1.2%
Lepton reconstruction	< 1%	< 1%	< 1%	< 1%	< 1%
Lepton ident./isolation	4.3%	4.2%	4.2%	4.1%	4.0%
Jet energy scale	5-9%	3-14%	3-14%	4-8%	1-14%
Unclustered energy	10-25%	5-17%	5-15%	2-7%	0-22%
b tagging	< 1%	1-3%	< 1%	0-3%	0-2%
Renorm./fact. scales	1-2%	0-2%	2-5%	4-5%	5-9%
PDFs	2-4%	0-4%	5.2%	2-4%	5-17%
$t\bar{t}Z$ normalization	< 1%	< 1%	< 1%	0-1%	0-2%
WZ normalization	< 1%	< 1%	< 1%	< 1%	< 1%
ZZ normalization	< 1%	< 1%	< 1%	< 1%	< 1%
Drell-Yan normalization	< 1%	< 1%	< 1%	< 1%	< 1%
$m_{T2}(ll)$ shape (top quark)	2-6%	8-16%	0-7%	3-13%	0-6%
$m_{T2}(ll)$ shape (WW)	4-26%	1-18%	3-28%	2-21%	4-23%
ZZ k -factors	< 1%	< 1%	< 1%	< 1%	< 1%
$m_{T2}(ll)$ shape (Drell-Yan)	< 1%	< 1%	< 1%	< 1%	< 1%
Nonprompt leptons	< 1%	< 1%	< 1%	< 1%	< 1%
$t\bar{t}$ p_T reweighting	< 1%	< 1%	0-3%	0-3%	0-6%
Statistical	5-6(16)%	1-4(10)%	14-20(45)%	1-14%	2-20(33)%

TABLE C.1: Range of systematic uncertainties on the total background prediction in the $e\mu$ channel. Statistical uncertainties are given for comparison. Extreme values found in the last bins of $m_{T2}(ll)$ are shown in parentheses.

Systematic	SR1 ^{0Jet} _{0Tag}	SR1 ^{Jets} _{0Tag}	SR2 ^{0Jet} _{0Tag}	SR2 ^{Jets} _{0Tag}	SR3 _{0Tag}
Integrated luminosity	2.5%	2.5%	2.5%	2.5%	2.5%
Trigger	2%	2%	2%	2%	2%
Pileup	3.6%	4.9%	2.6%	1.6%	1%
Lepton reconstruction	< 1%	< 1%	< 1%	< 1%	< 1%
Lepton ident./isolation	5.0%	4.9%	4-6%	5.0%	4-6%
Jet energy scale	1-8%	3-10%	0-7%	4-10%	2-10%
Unclustered energy	9-17%	7-11%	1-13%	2-9%	2-16%
b tagging	< 1%	0-3%	< 1%	0-3%	0-2%
Renorm./fact. scale	2.4%	0-4%	3-4%	4-5%	5-10%
PDFs	2-6%	0-5%	3-8%	2-6%	10-15%
$t\bar{t}Z$ normalization	< 1%	< 1%	< 1%	< 1%	0-1%
WZ normalization	0-1%	< 1%	0-1%	0-2%	0-1%
ZZ normalization	1-8%	0-2%	1-10%	0-4%	0-5%
Drell-Yan normalization	0-3%	0-15%	< 1%	0-6%	0-2%
$m_{T2}(ll)$ shape (top quark)	0-6%	3-13%	0-5%	1-10%	0-6%
$m_{T2}(ll)$ shape (WW)	3-12%	1-4%	4-11%	2-7%	3-10%
ZZ k -factors	0-5%	< 1%	0-4%	0-1%	< 1%
$m_{T2}(ll)$ shape (Drell-Yan)	< 1%	0-18%	< 1%	0-8%	0-2%
Nonprompt leptons	< 1%	< 1%	< 1%	< 1%	0-1%
$t\bar{t}$ p_T reweighting	< 1%	< 1%	0-2%	0-3%	0-5%
Statistical	5-12%	0-6%	10-20(42)%	1-11%	2-13(35)%

TABLE C.2: Same as Table C.1 but for the $ee+\mu\mu$ channel.

Systematic	SR1 ^{0Jet} _{0Tag}	SR1 ^{Jets} _{0Tag}	SR2 ^{0Jet} _{0Tag}	SR2 ^{Jets} _{0Tag}	SR3 _{0Tag}
Integrated luminosity	2.5%	2.5%	2.5%	2.5%	2.5%
Trigger	2%	2%	2%	2%	2%
Lepton reconstruction	0-1%	< 1%	0-1%	0-1%	0-1%
Lepton ident./isolation	3.9%	4.0%	3.8%	4.0%	3.8%
Jet energy scale	2-8%	2-3%	0-10%	2-21%	0-6%
Unclustered energy	0-6%	2-3%	0-10%	0-17%	2-25%
b tagging	< 1%	< 1%	< 1%	< 1%	< 1%
Renorm./fact. scales	< 1%	< 1%	< 1%	< 1%	< 1%
Lept. id./iso. (FastSim)	4.0%	4.0%	4.0%	4.0%	4.0%
b tagging (FastSim)	< 1%	< 1%	< 1%	< 1%	< 1%
\vec{p}_T^{miss} (FastSim)	0-4%	0-8%	0-10%	0-9%	1-24%
Pileup (FastSim)	0-6%	0-4%	0-1%	3-8%	1-8%
ISR reweighting	0-2%	0-1%	0-2%	1-2%	0-1%

TABLE C.3: Range of systematic uncertainties on the predicted yield for a representative signal with masses: $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0}) = (500, 200)$, signal in the $e\mu$ channel.

Systematic	SR1 ^{0Jet} _{0Tag}	SR1 ^{Jets} _{0Tag}	SR2 ^{0Jet} _{0Tag}	SR2 ^{Jets} _{0Tag}	SR3 _{0Tag}
Integrated luminosity	2.5%	2.5%	2.5%	2.5%	2.5%
Trigger	2%	2%	2%	2%	2%
Lepton reconstruction	0-1%	< 1%	< 1%	< 1%	0-2%
Lepton ident./isolation	4-5%	4.1%	4-5%	4.3%	3-5%
Jet energy scale	3-4%	3-5%	0-5%	0-4%	2-7%
Unclustered energy	0-5%	0-11%	0-39%	0-11%	2-7%
b tagging	< 1%	< 1%	< 1%	< 1%	< 1%
Renorm./fact. scales	< 1%	< 1%	< 1%	< 1%	< 1%
Lept. id./iso.(FastSim)	4.0%	4.0%	4.0%	4.0%	4.0%
b tagging (FastSim)	< 1%	< 1%	< 1%	< 1%	< 1%
\vec{p}_T^{miss} (FastSim)	4-11%	2-12%	2-60%	0-13%	2-25%
Pileup (FastSim)	0-6%	0-4%	0-1%	1-8%	4-11%
ISR reweighting	1-2%	0-2%	0-2%	0-2%	0-8%

TABLE C.4: Same as Table C.3 but for the $ee+\mu\mu$ channel.

Systematic	SR1 _{0Tag}	SR1 _{Tags}	SR2 _{0Tag}	SR2 _{Tags}	SR3 _{0Tag} ^{ISR}	SR3 _{Tags} ^{ISR}
Integrated luminosity	2.5%	2.5%	2.5%	2.5%	2.5%	2.5%
Trigger	2%	2%	2%	2%	2%	2%
Pileup	3.2%	2.4%	2.6%	2.1%	4.3%	3.4%
Lepton reconstruction	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
Lepton ident./isolation	4.2%	4.2%	4.1%	4.1%	3.9%	4-5%
Jet energy scale	0-13%	8-11%	4-8%	5-16%	2-19%	8-17%
Unclustered energy	2-18%	9-17%	0-7%	6-23%	3-17%	2-28%
b tagging	0-3%	1.2%	0-3%	1-2%	2.4%	< 1%
Renorm./fact. scales	0-2%	0-4%	3.8%	3-11%	8-12%	14-19%
PDFs	0-4%	1-2%	2-5%	2-4%	5-11%	5-16%
$t\bar{t}Z$ normalization	< 1%	0-2%	< 1%	0-10%	0-7%	0-5%
WZ normalization	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
ZZ normalization	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
Drell-Yan normalization	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
$m_{T2}(ll)$ shape (top quark)	6-15%	10-25%	2-12%	10-18%	0-9%	10-19%
$m_{T2}(ll)$ shape (WW)	1-21%	0-1%	2-24%	0-3%	4-15%	< 1%
ZZ k -factors	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
$m_{T2}(ll)$ shape (Drell-Yan)	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
Nonprompt leptons	< 1%	0-1%	< 1%	0-1%	< 1%	0-9%
$t\bar{t}$ p_T reweighting	< 1%	0-1%	0-3%	3-5%	0-5%	2-8%
Statistical	1-8%	< 1-7%	1-13%	1-16%	3-14(71)%	3-10(49)%

TABLE C.5: Range of systematic uncertainties on the total background prediction in the $e\mu$ channel. Statistical uncertainties are given for comparison. Extreme values found in the last bins of $m_{T2}(ll)$ are shown in parentheses.

Systematic	SR1 _{0Tag}	SR1 _{Tags}	SR2 _{0Tag}	SR2 _{Tags}	SR3 _{0Tag} ^{ISR}	SR3 _{Tags} ^{ISR}
Integrated luminosity	2.5%	2.5%	2.5%	2.5%	2.5%	2.5%
Trigger	2%	2%	2%	2%	2%	2%
Pileup	4.5%	3.5%	1.8%	1.9%	1.6%	2.2%
Lepton reconstruction	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
Lepton ident./isolation	4.8%	4.8%	5.0%	4.9%	4-5%	5-6%
Jet energy scale	3-10%	2-11%	3-10%	5-15%	4-21%	2-23%
Unclustered energy	5-12%	5-17%	2-9%	8-14%	3-28%	4-24%
b tagging	0-3%	1.3%	0-3%	1.3%	2-3%	< 1%
Renorm./fact. scales	0-3%	0-5%	4-5%	2-8%	9-14%	14-21%
PDFs	0-6%	2-6%	3-6%	3-7%	9-15%	5-16%
$t\bar{t}Z$ normalization	< 1%	0-2%	< 1%	0-8%	0-5%	0-9%
WZ normalization	< 1%	< 1%	0-2%	< 1%	0-2%	0-1%
ZZ normalization	0-2%	< 1%	0-4%	< 1%	0-5%	< 1%
Drell-Yan normalization	0-12%	0-8%	0-5%	0-5%	0-2%	0-7%
$m_{T2}(ll)$ shape (top quark)	3-12%	10-19%	1-9%	10-18%	0-13%	3-17%
$m_{T2}(ll)$ shape (WW)	1-6%	< 1%	2-8%	< 1%	0-4%	< 1%
ZZ k -factors	< 1%	< 1%	< 1%	< 1%	0-4%	< 1%
$m_{T2}(ll)$ shape (Drell-Yan)	0-15%	0-10%	0-6%	0-6%	0-3%	0-9%
Nonprompt leptons	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
$t\bar{t}$ p_T reweighting	< 1%	0-1%	0-3%	2-5%	0-5%	1-9%
Statistical	1-6%	1-6%	2-11%	1-9%	3-18(35)%	2-10(31)%

TABLE C.6: Same as Table C.5 but for the $ee+\mu\mu$ channel.

Systematic	SR1 _{0Tag}	SR1 _{Tags}	SR2 _{0Tag}	SR2 _{Tags}	SR3 _{0Tag} ^{ISR}	SR3 _{Tags} ^{ISR}
Integrated luminosity	2.5%	2.5%	2.5%	2.5%	2.5%	2.5%
Trigger	2%	2%	2%	2%	2%	2%
Lepton reconstruction	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
Lepton ident./isolation	4.3%	4.2%	4.2%	4.4%	4.3%	4.4%
Jet energy scale	0-2%	3-9%	3-9%	4-21%	5-13%	7-15%
Unclustered energy	2-5%	4-8%	3-4%	9-11%	5-6%	< 1%
b tagging	0-2%	2.9%	0-2%	2-3%	1.9%	1-2%
Renorm./fact. scales	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
Lept. id./iso.(FastSim)	4.0%	4.0%	4.0%	4.0%	4.0%	4.0%
b tagging (FastSim)	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
\vec{p}_T^{miss} (FastSim)	2-5%	0-14%	5-11%	2-38%	9-16%	0-13%
Pileup (FastSim)	0-5%	0-8%	0-8%	0-2%	2-18%	0-4%
ISR reweighting	0-3%	0-2%	3-4%	1-4%	4-7%	3-9%

TABLE C.7: Range of systematic uncertainties on the predicted yield of a representative signal with masses: $(m_{\tilde{\tau}_1}, m_{\tilde{\chi}_1^0}) = (350, 225)$, in the $e\mu$ channel.

Systematic	SR1 _{0Tag}	SR1 _{Tags}	SR2 _{0Tag}	SR2 _{Tags}	SR3 _{0Tag} ^{ISR}	SR3 _{Tags} ^{ISR}
Integrated luminosity	2.5%	2.5%	2.5%	2.5%	2.5%	2.5%
Trigger	2%	2%	2%	2%	2%	2%
Lepton reconstruction	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
Lepton ident./isolation	4.9%	5-6%	4-5%	4-5%	3-5%	4-6%
Jet energy scale	1-8%	2-15%	0-7%	1-15%	5-16%	2-20%
Unclustered energy	0-6%	0-4%	1-3%	4-8%	5-9%	0-18%
b tagging	0-1%	3.1%	1.8%	2.8%	1-2%	1-3%
Renorm./fact. scales	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
Lept. id./iso.(FastSim)	4.0%	4.0%	4.0%	4.0%	4.0%	4.0%
b tagging (FastSim)	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
\vec{p}_T^{miss} (FastSim)	3-19%	8-32%	3-22%	0-5%	0-10%	9-16%
Pileup (FastSim)	0-4%	0-8%	0-8%	0-3%	2-19%	0-5%
ISR reweighting	0-2%	0-2%	1-4%	4-7%	5-7%	7-12%

TABLE C.8: Same as Table C.7 but for the $ee+\mu\mu$ channel.

Appendix D

Pull of the nuisance parameters

The impact of the uncertainties on the result will depend on how the nuisances are treated and how much they are correlated with the parameter of interest. In order to verify that the quoted uncertainties in Section 6.5 do not present any pathology, the pulls of their associated nuisance parameters (Section 7.1) after the fit have been studied for both the background-only and signal+background hypotheses.

The results using one reference mass point for the top squark and chargino searches are shown respectively in Tables D.1- D.2. For the sake of readability, only the pull of nuisance parameters whose central value has been changed ($\Delta x / \sigma_{in}$) by more than $0.3\sigma_{in}$ and/or whose error ($\sigma_{out} / \sigma_{in}$) changes by more than 0.10 in the fit are reported. These sources of uncertainty include (listed from top to bottom): the b-tagging efficiency scale factors, data-driven shape corrections of Drell-Yan events, lepton identification and isolation efficiency scale factors, jet energy scale and unclustered energy, the modelling of the $m_{T2}(ll)$ shape for $t\bar{t}$, tW and WW backgrounds, the PDFs set, the QCD scales, the statistical uncertainties, the top p_T reweighting, the Drell-Yan normalization, the uncertainty on the shape correction of $m_{T2}(ll)$ for Drell-Yan events, and only in the chargino SRs, the modelling of the $m_{T2}(ll)$ shape for ZZ events and the Drell-Yan normalization uncertainty in the 0-jet SRs. In the last column, $\rho(\theta, \mu)$ shows the correlation coefficient between each nuisance parameter (θ) and the signal strength (μ).

In the chargino search the most affected nuisance are related to the b-tagging efficiency, the lepton identification and isolation efficiency, the uncertainty in the modelling of $m_{T2}(ll)$ for the top background ($t\bar{t}$, tW), and the WW or Drell-Yan statistical uncertainty. The parameters most correlated with the signal strength are the modelling uncertainty on bin number seven of the $m_{T2}(ll)$ shape for the WW background in events with $200 \leq p_T^{miss} < 300$ GeV (SR2), and the b-tagging efficiency. The most shifted nuisance parameters for the top squark search are those describing the b-tagging efficiency, the lepton identification and isolation efficiency, the $m_{T2}(ll)$ modelling of the top background and the PDFs uncertainties. In terms of their correlation with the signal strength the most important nuisances are those related to the b-tagging efficiency scale factors and the normalization of the Drell-Yan process.

Finally, Tables D.3 and D.4 list respectively the different rate parameters that we use in the chargino and top squark searches. As described in Section 6.4, the normalization of the main backgrounds from top ($t\bar{t}$ and tW) and WW production is left to be determined in the fit. The fitted values of the corresponding rate parameters give the normalization scale factor for these backgrounds and its uncertainty. In the case of the top background, the rate parameters through the different p_T^{miss} bins show a

mild tendency to decrease. For the WW background, the variations through the different p_T^{miss} bins are larger. The results are consistent between the two searches, but it is worth to mention the difference in the $p_T^{miss} > 300$ GeV bin where the selection is different since for the top squark search we apply an ISR jet requirement.

As described in Appendix B, in the case of the chargino search, additional rate parameters are included in order to take into account a possible bad modelling of the rate of events with no-jets for diboson and jet enriched backgrounds ($t\bar{t}$, tW , $t\bar{t}W$, $t\bar{t}Z$). The total number of yields expected for each process in the two regions with and without jets is constrained to remain invariant, so that only a migration of events from the two signal regions is allowed. The fitted values of these rate parameters for the jet enriched backgrounds are close to the unity, while for the diboson backgrounds they indicates a higher fitted fraction of events without jets. In particular, for the region with $200 < p_T^{miss} < 300$ GeV (SR2) the rate parameter reaches its maximum allowed value, in correspondence with the excess of observed events with no jets in SF and DF final states (see Figures 7.3-7.1, middle left).

name	<i>b</i> -only fit	<i>s + b</i> fit	$\rho(\theta, \mu)$
	$\Delta x / \sigma_{\text{in}}, \sigma_{\text{out}} / \sigma_{\text{in}}$	$\Delta x / \sigma_{\text{in}}, \sigma_{\text{out}} / \sigma_{\text{in}}$	
b-tag efficiency (b)	+0.62, 0.72	+0.59, 0.73	-0.18
IdIso leptons	+0.69, 0.87	+0.66, 0.87	-0.02
JES	-0.76, 0.58	-0.76, 0.59	-0.09
Unclustered energy	+0.20, 0.56	+0.16, 0.58	-0.21
Top $m_{T2}(ll)$ - Bin4 SR1	-0.25, 0.37	-0.25, 0.37	-0.03
Top $m_{T2}(ll)$ - Bin4 SR2	+1.06, 0.65	+1.07, 0.65	+0.02
Top $m_{T2}(ll)$ - Bin4 SR3	+0.55, 0.93	+0.55, 0.93	-0.00
Top $m_{T2}(ll)$ - Bin5 SR1	+0.32, 0.44	+0.34, 0.43	+0.11
Top $m_{T2}(ll)$ - Bin5 SR2	-0.10, 0.62	-0.10, 0.62	-0.03
Top $m_{T2}(ll)$ - Bin5 SR3	-0.45, 0.96	-0.46, 0.96	-0.03
Top $m_{T2}(ll)$ - Bin6 SR1	+0.09, 0.49	+0.09, 0.48	-0.02
Top $m_{T2}(ll)$ - Bin6 SR2	+0.21, 0.76	+0.18, 0.77	-0.21
Top $m_{T2}(ll)$ - Bin7 SR1	+0.06, 0.69	+0.05, 0.70	-0.04
WW $m_{T2}(ll)$ - Bin7 SR1	+0.07, 0.83	+0.07, 0.84	+0.02
WW $m_{T2}(ll)$ - Bin7 SR2	+0.12, 0.77	+0.12, 0.77	+0.03
PDFs	+0.52, 0.91	+0.51, 0.91	-0.13
QCD scale	+0.06, 0.60	+0.05, 0.60	-0.02
WW stat - SR2 Tag $e\mu$ Bin7	-0.30, 1.00	-0.30, 1.00	+0.02
WW stat - SR3 Tag $e\mu$ Bin5	-0.01, 0.59	-0.01, 0.55	+0.00
WW stat - SR3 Tag $e\mu$ Bin6	+0.17, 1.19	+0.13, 1.18	-0.13
WW stat - SR3 Tag $e\mu$ Bin7	-0.01, 0.60	-0.01, 0.46	+0.00
WZ ($\rightarrow 3l\nu$) stat - SR3 Tag $ee+\mu\mu$ Bin6	+0.08, 1.14	+0.07, 1.12	-0.04
$t\bar{t}$ stat - SR3 Tag $ee+\mu\mu$ Bin4	+0.55, 0.98	+0.55, 0.98	-0.01
tW stat - SR3 Tag $ee+\mu\mu$ Bin4	+0.43, 0.99	+0.43, 0.99	-0.00
tW stat - SR3 Tag $ee+\mu\mu$ Bin6	+0.27, 1.09	+0.23, 1.13	-0.14
WW stat - SR3 Tag $ee+\mu\mu$ Bin6	+0.12, 1.20	+0.10, 1.18	-0.07
WZ ($\rightarrow 3l\nu$) stat - SR3 Veto $e\mu$ Bin6	-0.02, 0.24	-0.02, 0.42	-0.00
WZ ($\rightarrow 3l\nu$) stat - SR3 Veto $e\mu$ Bin7	+0.07, 1.10	+0.05, 1.08	-0.06
tW stat - SR3 Veto $e\mu$ Bin7	+0.20, 1.13	+0.15, 1.15	-0.18
HWW stat - SR3 Veto $e\mu$ Bin4	-0.00, 0.66	-0.01, 0.72	-0.00
Top p_T reweighting	+0.03, 0.88	+0.01, 0.91	-0.02
Drell-Yan normalization	-0.22, 0.65	-0.19, 0.66	+0.18
Drell-Yan shape	+0.16, 0.67	+0.15, 0.68	-0.14

TABLE D.1: Pull values of the nuisance parameters for the for top squark search under the only background hypothesis (*b*-only) and for comparison under the background + signal hypothesis (*s + b*) using the signal mass point $(m_{\tilde{t}_1}, m_{\tilde{\chi}_1^0}) = (350, 225)$ GeV. The correlation coefficient between the each nuisance and the signal strength is also shown in the last column.

name	<i>b</i> -only fit	<i>s</i> + <i>b</i> fit	$\rho(\theta, \mu)$
	$\Delta x / \sigma_{\text{in}}, \sigma_{\text{out}} / \sigma_{\text{in}}$	$\Delta x / \sigma_{\text{in}}, \sigma_{\text{out}} / \sigma_{\text{in}}$	
b-tag efficiency (b)	+0.71, 0.77	+0.58, 0.85	-0.41
IdIso leptons	+0.80, 0.86	+0.78, 0.87	-0.05
JES	-0.02, 0.52	-0.02, 0.51	+0.02
Unclustered energy	-0.68, 0.50	-0.67, 0.51	+0.06
Top $m_{T2}(ll)$ - Bin4 SR1	-0.30, 0.35	-0.30, 0.35	-0.03
Top $m_{T2}(ll)$ - Bin4 SR2	+0.98, 0.64	+0.98, 0.64	-0.01
Top $m_{T2}(ll)$ - Bin4 SR3	+0.82, 0.91	+0.82, 0.91	+0.01
Top $m_{T2}(ll)$ - Bin5 SR1	+0.43, 0.32	+0.42, 0.32	-0.07
Top $m_{T2}(ll)$ - Bin5 SR2	-0.08, 0.59	-0.09, 0.59	-0.02
Top $m_{T2}(ll)$ - Bin5 SR3	-0.43, 0.92	-0.43, 0.92	-0.01
WW $m_{T2}(ll)$ - Bin5 SR1	-0.55, 0.96	-0.53, 0.97	+0.07
Top $m_{T2}(ll)$ - Bin6 SR1	+0.30, 0.37	+0.30, 0.37	-0.02
Top $m_{T2}(ll)$ - Bin6 SR2	+0.12, 0.74	+0.12, 0.74	-0.00
Top $m_{T2}(ll)$ - Bin6 SR3	+0.30, 0.96	+0.30, 0.96	+0.00
Top $m_{T2}(ll)$ - Bin7 SR1	+0.04, 0.69	+0.04, 0.68	-0.02
WW $m_{T2}(ll)$ - Bin7 SR1	+0.08, 0.79	+0.01, 0.83	-0.25
WW $m_{T2}(ll)$ - Bin7 SR2	-0.19, 0.74	-0.34, 0.85	-0.47
WW $m_{T2}(ll)$ - Bin7 SR3	+0.32, 0.92	+0.25, 0.95	-0.19
PDFs	+0.24, 0.70	+0.20, 0.69	-0.16
QCD scale	-0.13, 0.55	-0.12, 0.55	+0.02
WZ ($\rightarrow 3lv$) stat - SR1 NoJet $e\mu$ Bin2	-0.01, 0.79	-0.01, 0.81	+0.00
WW stat - SR1 NoJet $e\mu$ Bin5	-0.59, 0.97	-0.56, 0.98	+0.09
tW stat - SR1 NoJet $ee+\mu\mu$ Bin7	-0.00, 0.75	-0.00, 0.70	-0.00
WW stat - SR1 NoJet $ee+\mu\mu$ Bin2	+0.68, 0.97	+0.65, 0.98	-0.07
Drell-Yan stat - SR1 NoJet $ee+\mu\mu$ Bin1	+0.06, 0.85	+0.08, 0.85	+0.04
Drell-Yan stat - SR1 NoJet $ee+\mu\mu$ Bin2	+0.66, 1.05	+0.69, 1.02	+0.09
Drell-Yan stat - SR1 NoJet $ee+\mu\mu$ Bin6	-0.00, 0.14	-0.00, 0.31	+0.01
WZ ($\rightarrow 3lv$) stat - SR2 NoJet $e\mu$ Bin2	-0.02, 0.52	-0.02, 0.55	+0.00
WW stat - SR2 NoJet $e\mu$ Bin2	-0.39, 0.96	-0.37, 0.96	+0.05
WW stat - SR2 NoJet $e\mu$ Bin3	-0.43, 0.94	-0.42, 0.95	+0.04
WW stat - SR2 NoJet $e\mu$ Bin6	+0.16, 0.88	+0.16, 0.89	+0.00
Drell-Yan stat - SR2 NoJet $e\mu$ Bin2	-0.01, 0.83	-0.01, 0.97	+0.00
WZ ($\rightarrow 3lv$) stat - SR2 NoJet $ee+\mu\mu$ Bin3	+0.06, 1.12	+0.06, 1.12	-0.00
$t\bar{t}$ stat - SR2 NoJet $ee+\mu\mu$ Bin7	+0.00, 0.83	+0.00, 0.77	-0.00
tW stat - SR2 NoJet $ee+\mu\mu$ Bin6	+0.11, 1.21	+0.12, 1.22	+0.03
tW stat - SR2 NoJet $ee+\mu\mu$ Bin7	+0.00, 1.01	-0.00, 0.69	-0.01
WW stat - SR2 NoJet $ee+\mu\mu$ Bin1	-0.33, 0.97	-0.32, 0.97	+0.04
WW stat - SR2 NoJet $ee+\mu\mu$ Bin2	+0.57, 0.95	+0.56, 0.95	-0.04
WW stat - SR2 NoJet $ee+\mu\mu$ Bin3	+0.57, 0.94	+0.55, 0.95	-0.05
WW stat - SR2 NoJet $ee+\mu\mu$ Bin4	+0.33, 0.95	+0.33, 0.95	-0.03
WW stat - SR2 NoJet $ee+\mu\mu$ Bin6	+0.33, 0.87	+0.33, 0.88	-0.01
WW stat - SR2 NoTag $e\mu$ Bin4	+0.31, 0.97	+0.30, 0.97	-0.02
WW stat - SR2 Tag $e\mu$ Bin7	-0.33, 0.99	-0.31, 1.00	+0.04
WZ ($\rightarrow 3lv$) stat - SR3 Tag $e\mu$ Bin6	+0.12, 1.28	+0.12, 1.28	-0.00
$t\bar{t}$ stat - SR3 Tag $e\mu$ Bin6	+0.36, 0.97	+0.36, 0.97	+0.00
WW stat - SR3 Tag $e\mu$ Bin7	-0.01, 0.75	-0.01, 0.45	+0.00
HWW stat - SR3 Tag $e\mu$ Bin6	+0.06, 1.11	+0.06, 1.11	+0.00
WZ ($\rightarrow 3lv$) stat - SR3 Tag $ee+\mu\mu$ Bin5	-0.01, 0.93	-0.01, 0.83	+0.00
$t\bar{t}$ stat - SR3 Tag $ee+\mu\mu$ Bin4	+0.50, 0.98	+0.50, 0.98	-0.00
tW stat - SR3 Tag $ee+\mu\mu$ Bin4	+0.45, 0.99	+0.45, 0.99	+0.00
tW stat - SR3 Tag $ee+\mu\mu$ Bin7	+0.10, 1.15	+0.10, 1.15	+0.00
WW stat - SR3 Tag $ee+\mu\mu$ Bin6	-0.00, 0.74	-0.00, 0.73	+0.00
WW stat - SR3 Veto $ee+\mu\mu$ Bin5	-0.30, 0.99	-0.29, 0.99	+0.03
Top p_T reweighting	-0.03, 0.78	-0.03, 0.80	-0.00
Drell-Yan normalization	-0.14, 0.65	-0.10, 0.65	+0.15
Drell-Yan shape	+0.35, 0.60	+0.30, 0.61	-0.22
ZZ shape	+0.06, 1.29	+0.07, 1.29	+0.01
Drell-Yan nojet normalization	-0.16, 0.81	-0.13, 0.81	+0.08

TABLE D.2: Same as Table D.1 but for the chargino search using the signal mass point $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0}) = (500, 200)$ GeV.

name	pre fit	b -only fit	$s + b$ fit	$\rho(\theta, \mu)$
Top normalization SR1	[-, -]	+0.88 ± 0.09	+0.88 ± 0.09	+0.09
Top normalization SR2	[-, -]	+0.81 ± 0.09	+0.81 ± 0.09	+0.11
Top normalization SR3	[-, -]	+0.76 ± 0.09	+0.76 ± 0.09	+0.13
WW normalization SR1	[-, -]	+0.87 ± 0.25	+0.84 ± 0.26	-0.33
WW normalization SR2	[-, -]	+1.08 ± 0.21	+1.05 ± 0.23	-0.40
WW normalization SR3	[-, -]	+0.63 ± 0.16	+0.60 ± 0.17	-0.33
NoJet Rate Diboson backgrounds SR1	[0.70, 1.30]	+1.24 ± 0.48	+1.21 ± 0.40	-0.14
NoJet Rate Diboson backgrounds SR2	[0.70, 1.30]	+1.30 ± 0.44	+1.30 ± 0.46	-0.00
NoJet Rate Jet enriched backgrounds SR1	[0.50, 1.50]	+0.91 ± 0.21	+0.94 ± 0.22	+0.38
NoJet Rate Jet enriched backgrounds SR2	[0.50, 1.50]	+1.02 ± 0.34	+1.05 ± 0.35	+0.26

TABLE D.3: Variation of the SM rate parameters for the chargino search under the only background hypothesis (b -only) and for comparison under the background + signal hypothesis ($s + b$) using the signal mass point $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0}) = (500, 200)$ GeV. The correlation coefficient between the each nuisance and the signal strength is also shown in the last column.

name	pre fit	b -only fit	$s + b$ fit	$\rho(\theta, \mu)$
Top normalization SR1	[-, -]	+0.91 ± 0.10	+0.91 ± 0.10	+0.08
Top normalization SR2	[-, -]	+0.83 ± 0.10	+0.83 ± 0.10	+0.08
Top normalization SR3	[-, -]	+0.82 ± 0.13	+0.82 ± 0.13	+0.00
WW normalization SR1	[-, -]	+0.86 ± 0.23	+0.85 ± 0.24	-0.29
WW normalization SR2	[-, -]	+1.02 ± 0.19	+1.00 ± 0.21	-0.34
WW normalization SR3	[-, -]	+0.43 ± 0.18	+0.42 ± 0.19	-0.30

TABLE D.4: Variation of the SM rate parameters for the top squark search under the only background hypothesis (b -only) and for comparison under the background + signal hypothesis ($s + b$) using the signal mass point $(m_{\tilde{t}_1}, m_{\tilde{\chi}_1^0}) = (350, 225)$ GeV. The correlation coefficient between the each nuisance and the signal strength is also shown in the last column.

Bibliography

- [1] E. Gildener and S. Weinberg. Symmetry breaking and scalar bosons. *Phys. Rev. D*, 13:3333, 1976. doi: 10.1103/PhysRevD.13.3333.
- [2] G. 't Hooft. Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking. *NATO Sci. Ser. B*, 59:135, 1980. doi: 10.1007/978-1-4684-7571-5_9.
- [3] G. Bertone, D. Hooper, and J. Silk. Particle dark matter: evidence, candidates and constraints. *Phys. Rept.*, 405:279, 2005. doi: 10.1016/j.physrep.2004.08.031.
- [4] J. L. Feng. Dark matter candidates from particle physics and methods of detection. *Ann. Rev. Astron. Astrophys.*, 48:495, 2010. doi: 10.1146/annurev-astro-082708-101659.
- [5] T. A. Porter, R. P. Johnson, and P. W. Graham. Dark matter searches with astroparticle data. *Ann. Rev. Astron. Astrophys.*, 49:155, 2011. doi: 10.1146/annurev-astro-081710-102528.
- [6] P. Ramond. Dual theory for free fermions. *Phys. Rev. D*, 3:2415, 1971. doi: 10.1103/PhysRevD.3.2415.
- [7] Y. A. Golfand and E. P. Likhtman. Extension of the algebra of Poincaré group generators and violation of P invariance. *JETP Lett.*, 13:323, 1971.
- [8] A. Neveu and J. H. Schwarz. Factorizable dual model of pions. *Nucl. Phys. B*, 31:86, 1971. doi: 10.1016/0550-3213(71)90448-2.
- [9] D. V. Volkov and V. P. Akulov. Possible universal neutrino interaction. *JETP Lett.*, 16:438, 1972.
- [10] J. Wess and B. Zumino. A Lagrangian model invariant under supergauge transformations. *Phys. Lett. B*, 49:52, 1974. doi: 10.1016/0370-2693(74)90578-4.
- [11] J. Wess and B. Zumino. Supergauge transformations in four dimensions. *Nucl. Phys. B*, 70:39, 1974. doi: 10.1016/0550-3213(74)90355-1.
- [12] P. Fayet. Supergauge invariant extension of the Higgs mechanism and a model for the electron and its neutrino. *Nucl. Phys. B*, 90:104, 1975. doi: 10.1016/0550-3213(75)90636-7.
- [13] H. P. Nilles. Supersymmetry, supergravity and particle physics. *Phys. Rep.*, 110:1, 1984. doi: 10.1016/0370-1573(84)90008-5.
- [14] S. P. Martin. A Supersymmetry primer. *Advanced Series on Directions in High Energy Physics*, 21:1, 2010. doi: 10.1142/9789814307505_0001.
- [15] E. Witten. Dynamical breaking of supersymmetry. *Nucl. Phys. B*, 188:513, 1981. doi: 10.1016/0550-3213(81)90006-7.

- [16] S. Dimopoulos and H. Georgi. Softly broken supersymmetry and SU(5). *Nucl. Phys. B*, 193:150, 1981. doi: 10.1016/0550-3213(81)90522-8.
- [17] Kaul, R. K. and Majumdar, P. Cancellation of quadratically divergent mass corrections in globally supersymmetric spontaneously broken gauge theories. *Nucl. Phys. B*, 199:36, 1982. doi: 10.1016/0550-3213(82)90565-X.
- [18] G. R. Farrar and P. Fayet. Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry. *Phys. Lett. B*, 76:575, 1978. doi: 10.1016/0370-2693(78)90858-4.
- [19] CMS collaboration. Searches for pair production of charginos and top squarks in final states with two oppositely charged leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 11:079, 2018. doi: 10.1007/JHEP11(2018)079.
- [20] F. Halzen and A. D. Martin. *Quarks and leptons: An introductory course in modern particle physics*. John Wiley and Sons, New York, 1984. ISBN 0471887412, 9780471887416.
- [21] M. Thomson. *Modern particle physics*. Cambridge University Press, New York, 2013. ISBN 9781107034266.
- [22] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev.*, D98:030001, 2018. doi: 10.1103/PhysRevD.98.030001.
- [23] T. W. B. Kibble. The Standard Model of Particle Physics. In *Invited talk at 25 Anniversary Meeting of Academia Europaea (arXiv:1412.4094)*, 2014.
- [24] A. Pich. The Standard model of electroweak interactions. In *The Standard model of electroweak interactions*, pages 1–49, 2008.
- [25] N. Cabibbo. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10:531–533, 1963. doi: 10.1103/PhysRevLett.10.531. [648(1963)].
- [26] M. Kobayashi and T. Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49:652–657, 1973. doi: 10.1143/PTP.49.652.
- [27] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961. doi: 10.1016/0029-5582(61)90469-2.
- [28] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967. doi: 10.1103/PhysRevLett.19.1264.
- [29] A. Salam. Weak and Electromagnetic Interactions. In *8th Nobel Symposium Lerum, Sweden, May 19-25, 1968*, volume C680519, pages 367–377, 1968.
- [30] D. J. Gross and F. Wilczek. Ultraviolet Behavior of Nonabelian Gauge Theories. *Phys. Rev. Lett.*, 30:1343–1346, 1973. doi: 10.1103/PhysRevLett.30.1343.
- [31] H. D. Politzer. Reliable Perturbative Results for Strong Interactions? *Phys. Rev. Lett.*, 30:1346–1349, 1973. doi: 10.1103/PhysRevLett.30.1346.
- [32] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, 1964. doi: 10.1103/PhysRevLett.13.321.
- [33] P. W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964. doi: 10.1103/PhysRevLett.13.508.

- [34] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett.*, 13:585–587, 1964. doi: 10.1103/PhysRevLett.13.585.
- [35] ATLAS Collaboration. Observation of a new particle in the search for the standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012. doi: 10.1016/j.physletb.2012.08.020.
- [36] CMS Collaboration. Observation of a new boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012. doi: 10.1016/j.physletb.2012.08.021.
- [37] G. 't Hooft. Renormalization of Massless Yang-Mills Fields. *Nucl. Phys.*, B33: 173–199, 1971. doi: 10.1016/0550-3213(71)90395-6.
- [38] J. Goldstone. Field Theories with Superconductor Solutions. *Nuovo Cim.*, 19: 154–164, 1961. doi: 10.1007/BF02812722.
- [39] J. Ellis. Higgs Physics. In *Proceedings, 2013 European School of High-Energy Physics (ESHEP 2013): Paradfurdo, Hungary, June 5-18, 2013*, pages 117–168, 2015. doi: 10.5170/CERN-2015-004.117.
- [40] A. B. Balantekin and W. C. Haxton. Neutrino Oscillations. *Prog. Part. Nucl. Phys.*, 71:150–161, 2013. doi: 10.1016/j.pnpnp.2013.03.007.
- [41] Y. Gershtein et al. Working Group Report: New Particles, Forces, and Dimensions. In *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*, 2013.
- [42] H. Georgi and S. L. Glashow. Unity of All Elementary Particle Forces. *Phys. Rev. Lett.*, 32:438–441, 1974. doi: 10.1103/PhysRevLett.32.438.
- [43] H. Miyazawa. Baryon Number Changing Currents. *Progress of Theoretical Physics*, 36:1266–1276, 1966. doi: 10.1143/PTP.36.1266.
- [44] H. Miyazawa. Spinor currents and symmetries of baryons and mesons. *Phys. Rev.*, 170:1586–1590, Jun 1968. doi: 10.1103/PhysRev.170.1586.
- [45] P. Fayet. Introduction to Supersymmetry and Its Applications to Particle Interactions. In *Fundamentals of Microprocessing San Francisco, California, May 23-25, 1977*, 1977.
- [46] P. Fayet. Supersymmetry and Weak, Electromagnetic and Strong Interactions. In *Proceedings, XVIII International Conference on High-Energy Physics Volume 2: July 15-21, 1976 Tbilisi, USSR*, pages T8–T11, 2017.
- [47] C. Borschensky et al. Squark and gluino production cross sections in pp collisions at $\sqrt{s} = 13, 14, 33$ and 100 TeV. *Eur. Phys. J.*, C74:3174, 2014. doi: 10.1140/epjc/s10052-014-3174-y.
- [48] J. Alwall, P. Schuster, and N. Toro. Simplified models for a first characterization of new physics at the LHC. *Phys. Rev. D*, 79:075020, 2009. doi: 10.1103/PhysRevD.79.075020.
- [49] J. Alwall, M.-P. Le, M. Lisanti, and J. G. Wacker. Model-independent jets plus missing energy searches. *Phys. Rev. D*, 79:015005, 2009. doi: 10.1103/PhysRevD.79.015005.

- [50] D. Alves et al. Simplified models for LHC new physics searches. *J. Phys. G*, 39: 105005, 2012. doi: 10.1088/0954-3899/39/10/105005.
- [51] CMS Collaboration. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008. doi: 10.1088/1748-0221/3/08/S08004.
- [52] CMS Collaboration. 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. C*, 74:3036, 2014. doi: 10.1140/epjc/s10052-014-3036-7.
- [53] ATLAS Collaboration. Search for direct production of charginos, neutralinos and sleptons in final states with two leptons and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *JHEP*, 05:071, 2014. doi: 10.1007/JHEP05(2014)071.
- [54] ATLAS Collaboration. Search for the direct production of charginos, neutralinos and staus in final states with at least two hadronically decaying taus and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *JHEP*, 10:096, 2014. doi: 10.1007/JHEP10(2014)096.
- [55] ATLAS Collaboration. Search for the electroweak production of supersymmetric particles in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector. *Phys. Rev. D*, 93:052002, 2016. doi: 10.1103/PhysRevD.93.052002.
- [56] ATLAS Collaboration. Search for the direct production of charginos and neutralinos in final states with tau leptons in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector. *Eur. Phys. J. C*, 78:154, 2018. doi: 10.1140/epjc/s10052-018-5583-9.
- [57] ATLAS Collaboration. Search for electroweak production of supersymmetric states in scenarios with compressed mass spectra at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 97:052010, 2018. doi: 10.1103/PhysRevD.97.052010.
- [58] ATLAS Collaboration. Search for electroweak production of supersymmetric particles in final states with two or three leptons at $\sqrt{s} = 13$ TeV with the ATLAS detector. Submitted to *Eur. Phys. J. C*, 2018.
- [59] CMS Collaboration. Search for top squarks and dark matter particles in opposite-charge dilepton final states at $\sqrt{s} = 13$ TeV. *Phys. Rev. D*, 97:032009, 2018. doi: 10.1103/PhysRevD.97.032009.
- [60] CMS collaboration. Search for the pair production of light top squarks in the $e\mu$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 2019:101, 2019. doi: 10.1007/JHEP03(2019)101.
- [61] CMS Collaboration. Search for top squark pair production in pp collisions at $\sqrt{s} = 13$ TeV using single lepton events. *JHEP*, 10:019, 2017. doi: 10.1007/JHEP10(2017)019.
- [62] CMS Collaboration. 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*, 10:005, 2017. doi: 10.1007/JHEP10(2017)005.
- [63] ATLAS Collaboration. Search for a scalar partner of the top quark in the jets plus missing transverse momentum final state at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 12:085, 2017. doi: 10.1007/JHEP12(2017)085.

- [64] ATLAS Collaboration. Search for top-squark pair production in final states with one lepton, jets, and missing transverse momentum using 36 fb^{-1} of $\sqrt{s} = 13 \text{ TeV}$ pp collision data with the ATLAS detector. *JHEP*, 06:108, 2018. doi: 10.1007/JHEP06(2018)108.
- [65] ATLAS Collaboration. Search for direct top squark pair production in final states with two leptons in $\sqrt{s} = 13 \text{ TeV}$ pp collisions with the ATLAS detector. *Eur. Phys. J. C*, 77:898, 2017. doi: 10.1140/epjc/s10052-017-5445-x.
- [66] R. K. Ellis, H. Georgi, M. Machacek, H. David Politzer, and Graham G. Ross. Factorization and the Parton Model in QCD. *Phys. Lett.*, 78B:281–284, 1978. doi: 10.1016/0370-2693(78)90023-0.
- [67] Y. L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and $e^+ e^-$ Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov. Phys. JETP*, 46:641–653, 1977.
- [68] G. Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys.*, B126:298–318, 1977. doi: 10.1016/0550-3213(77)90384-4.
- [69] L. N. Lipatov. The parton model and perturbation theory. *Sov. J. Nucl. Phys.*, 20:94–102, 1975.
- [70] V. N. Gribov and L. N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:438–450, 1972.
- [71] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J.*, C63:189–285, 2009. doi: 10.1140/epjc/s10052-009-1072-5.
- [72] O. S. Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004.
- [73] CERN. *LEP design report*, volume 1,2. CERN, Geneva, 1983, 1984.
- [74] ALICE Collaboration. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008. doi: 10.1088/1748-0221/3/08/S08002.
- [75] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. doi: 10.1088/1748-0221/3/08/S08003.
- [76] LHCb Collaboration. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008. doi: 10.1088/1748-0221/3/08/S08005.
- [77] CMS Collaboration. Technical proposal for the upgrade of the CMS detector through 2020. Technical Report CERN-LHCC-2011-006. LHCC-P-004, 2011.
- [78] TOTEM Collaboration. The TOTEM experiment at the CERN Large Hadron Collider. *JINST*, 3:S08007, 2008. doi: 10.1088/1748-0221/3/08/S08007.
- [79] LHCf Collaboration. The LHCf detector at the CERN Large Hadron Collider. *JINST*, 3:S08006, 2008. doi: 10.1088/1748-0221/3/08/S08006.
- [80] J. L. Pinfold. The MoEDAL Experiment at the LHC – a New Light on the Terascale Frontier. *J. Phys. Conf. Ser.*, 631(1):012014, 2015. doi: 10.1088/1742-6596/631/1/012014.
- [81] B Muratori. Luminosity and luminous region calculations for the lhc. Technical Report LHC-PROJECT-NOTE-301, 2002.

- [82] J. Wenninger. Approaching the Nominal Performance at the LHC. In *Proc. of International Particle Accelerator Conference (IPAC'17), Copenhagen, Denmark, 14-19 May, 2017*, number 8 in International Particle Accelerator Conference, pages 13–18, Geneva, Switzerland, 2017. JACoW. doi: 10.18429/JACoW-IPAC2017-MOYAA1.
- [83] CMS Collaboration. Public CMS Luminosity Information, 2016. URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2016_proton_proton_collisions. (r162 - 2019-09-02).
- [84] CMS Collaboration. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006.
- [85] CMS Collaboration. *CMS Physics: Technical Design Report Volume 2: Physics Performance*, volume 34. 2007.
- [86] CMS Collaboration. *The CMS magnet project: Technical Design Report*. Number CERN-LHCC-97-010 in Technical Design Report CMS. CERN, Geneva, 1997.
- [87] CMS Collaboration. *The CMS tracker system project: Technical Design Report*. Number CERN-LHCC-98-006 in Technical Design Report CMS. CERN, 1997.
- [88] CMS Collaboration. Tag and Probe Technique CMS Collaboration Twiki, 2015. URL <https://twiki.cern.ch/twiki/bin/view/CMSPublic/TagAndProbe>. (r15 - 2014-02-15).
- [89] CMS Collaboration. Tracking Efficiency for 2016 CMS Collaboration Twiki, 2016. URL <https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackingPOGPlots2016>. (r3-2019-09-13).
- [90] CMS Collaboration. *The CMS electromagnetic calorimeter project: Technical Design Report*. Number CERN-LHCC-97-033 in Technical Design Report CMS. CERN, Geneva, 1997.
- [91] CMS Collaboration. CMS ECAL first results with 2016 data. Technical Report CMS-DP-2016-031, 2016.
- [92] *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.
- [93] Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector. Technical Report CMS-PAS-JME-16-004, 2016.
- [94] CMS Collaboration. *The CMS muon project: Technical Design Report*. Number CERN-LHCC-97-032 in Technical Design Report CMS. CERN, Geneva, 1997.
- [95] CMS Collaboration. Performance of muon reconstruction including Alignment Position Errors for 2016 Collision Data. Technical Report CMS-DP-2016-067, 2016.
- [96] CMS Collaboration. *TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*. Number CERN-LHCC-2000-038 in Technical Design Report CMS. 2000.
- [97] CMS Collaboration. *The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Number CERN-LHCC-2002-026 in Technical Design Report CMS. CERN, Geneva, 2002.

- [98] CMS Collaboration. The CMS trigger system. *JINST*, 12:P01020, 2017. doi: 10.1088/1748-0221/12/01/P01020.
- [99] A Tapper and D. Acosta. CMS Technical Design Report for the Level-1 Trigger Upgrade. Technical Report CERN-LHCC-2013-011. CMS-TDR-12, 2013.
- [100] CMS Luminosity Measurements for the 2016 Data Taking Period. Technical Report CMS-PAS-LUM-17-001, 2017.
- [101] S. van der Meer. Calibration of the effective beam height in the ISR. Technical Report CERN-ISR-PO-68-31. ISR-PO-68-31, 1968.
- [102] CMS Collaboration. Measurement of CMS Luminosity. Technical Report CMS-PAS-EWK-10-004, CERN, Geneva, 1900.
- [103] J. Knobloch et al. *LHC computing Grid: Technical Design Report*. Technical Design Report LCG. CERN, Geneva, 2005.
- [104] CMS Collaboration. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12:P10003, 2017. doi: 10.1088/1748-0221/12/10/P10003.
- [105] CMS Collaboration. Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10:P06005, 2015. doi: 10.1088/1748-0221/10/06/P06005.
- [106] W. Adam, B. Mangano, T. Speer, and T. Todorov. Track Reconstruction in the CMS tracker. Technical Report CMS-NOTE-2006-041, CERN, 2006.
- [107] W. Adam et al. Reconstruction of electrons with the gaussian-sum filter in the CMS tracker at the LHC. *Journal of Physics G: Nuclear and Particle Physics*, 31:N9–N20, 2005. doi: 10.1088/0954-3899/31/9/n01.
- [108] CMS collaboration. Performance of CMS Muon Reconstruction in pp Collision Events at $\sqrt{s} = 7$ TeV. *JINST*, 7:P10002, 2012. doi: 10.1088/1748-0221/7/10/P10002.
- [109] CMS Collaboration. Muon Identification and Isolation efficiency on full 2016 dataset. Technical report, CERN, Geneva, Mar 2017.
- [110] CMS Collaboration. Electron and photon performance in CMS with the full 2016 data sample. Technical report, CERN, Geneva, 2017.
- [111] CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9:P10009, 2014. doi: 10.1088/1748-0221/9/10/P10009.
- [112] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- k_t jet clustering algorithm. *JHEP*, 04:063, 2008. doi: 10.1088/1126-6708/2008/04/063.
- [113] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *Eur. Phys. J. C*, 72:1896, 2012. doi: 10.1140/epjc/s10052-012-1896-2.
- [114] Nicholas Metropolis and S. Ulam. The monte carlo method. *Journal of the American Statistical Association*, 44(247):335–341, 1949.
- [115] R. K. Ellis and W. J. Stirling. QCD and collider physics. In *The 1990 CERN School of Physics Mallorca, Spain, September 16-29, 1990*, pages 135–236, 1992.

- [116] P. Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004. doi: 10.1088/1126-6708/2004/11/040.
- [117] S. Frixione, P. Nason, and C. Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007. doi: 10.1088/1126-6708/2007/11/070.
- [118] S. Alioli, P. Nason, C. Oleari, and E. Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010. doi: 10.1007/JHEP06(2010)043.
- [119] J. Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014. doi: 10.1007/JHEP07(2014)079.
- [120] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003. doi: 10.1088/1126-6708/2003/07/001.
- [121] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter. Event generation with SHERPA 1.1. *JHEP*, 02:007, 2009. doi: 10.1088/1126-6708/2009/02/007.
- [122] T. Sjostrand, S. Mrenna, and P. Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006. doi: 10.1088/1126-6708/2006/05/026.
- [123] T. Sjöstrand et al. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159, 2015. doi: 10.1016/j.cpc.2015.01.024.
- [124] M. Bahr et al. Herwig++ Physics and Manual. *Eur. Phys. J.*, C58:639–707, 2008. doi: 10.1140/epjc/s10052-008-0798-9.
- [125] J. Bellm and et al. Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J.*, C76(4):196, 2016. doi: 10.1140/epjc/s10052-016-4018-8.
- [126] S. Agostinelli et al. GEANT 4:a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003. doi: 10.1016/S0168-9002(03)01368-8.
- [127] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss. TMVA - Toolkit for Multivariate Data Analysis with ROOT: Users guide. Technical Report physics/0703039, CERN, Geneva, 2007.
- [128] CMS Collaboration. Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett.*, B791:96, 2019. doi: 10.1016/j.physletb.2018.12.073.
- [129] CMS Collaboration. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST*, 12:P02014, 2017. doi: 10.1088/1748-0221/12/02/P02014.
- [130] CMS Collaboration. Jet energy scale and resolution performance with 13 TeV data collected by CMS in 2016. Technical report, 2018.
- [131] CMS collaboration. Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS. *JINST*, 6:P11002, 2011. doi: 10.1088/1748-0221/6/11/P11002.

- [132] CMS collaboration. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *JINST*, 13:P05011, 2018. doi: 10.1088/1748-0221/13/05/P05011.
- [133] B. Chazin. Machine learning techniques for heavy flavour identification. In *proceedings of The sixth annual Large Hadron Collider Physics (LHCP18)*, number PoS(LHCP2018)066. PoS, 2018. doi: <https://doi.org/10.22323/1.321.0066>.
- [134] W. Waltenberger. Adaptive Vertex Reconstruction. Technical Report CMS-NOTE-2008-033, CERN, Geneva, 2008.
- [135] CMS Collaboration. Measurement of $B\bar{B}$ Angular Correlations based on Secondary Vertex Reconstruction at $\sqrt{s} = 7$ TeV. *JHEP*, 03:136, 2011. doi: 10.1007/JHEP03(2011)136.
- [136] CMS collaboration. Missing transverse energy performance of the CMS detector. *JINST*, 6:P09001, 2011. doi: 10.1088/1748-0221/6/09/P09001.
- [137] CMS collaboration. Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV. *JINST*, 10:P02006, 2015. doi: 10.1088/1748-0221/10/02/P02006.
- [138] P. Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004. doi: 10.1088/1126-6708/2004/11/040.
- [139] S. Frixione, P. Nason, and C. Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 11:070, 2007. doi: 10.1088/1126-6708/2007/11/070.
- [140] M. Czakon and A. Mitov. Top++: A program for the calculation of the top-pair cross-section at hadron colliders. *Comput. Phys. Commun.*, 185:2930, 2014. doi: 10.1016/j.cpc.2014.06.021.
- [141] E. Re. Single-top Wt-channel production matched with parton showers using the POWHEG method. *Eur. Phys. J. C*, 71:1547, 2011. doi: 10.1140/epjc/s10052-011-1547-z.
- [142] Kidonakis, N. NNLL threshold resummation for top-pair and single-top production. *Phys. Part. Nucl.*, 45:714, 2014. doi: 10.1134/S1063779614040091.
- [143] Melia, T. and Nason, P. and Rontsch, R. and Zanderighi, G. W^+W^- , WZ and ZZ production in the POWHEG BOX. *JHEP*, 11:078, 2011. doi: 10.1007/JHEP11(2011)078.
- [144] P. Nason and G. Zanderighi. W^+W^- , WZ and ZZ production in the POWHEG-BOX-V2. *Eur. Phys. J. C*, 74:2702, 2014. doi: 10.1140/epjc/s10052-013-2702-5.
- [145] T. Gehrmann et al. W^+W^- production at hadron colliders in NNLO QCD. *Phys. Rev. Lett.*, 113:212001, 2014. doi: 10.1103/PhysRevLett.113.212001.
- [146] CMS Collaboration. Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 11:047, 2017. doi: 10.1007/JHEP11(2017)047.
- [147] J. M. Campbell and R. K. Ellis. MCFM for the Tevatron and the LHC. *Nucl. Phys. Proc. Suppl.*, 205:10, 2010. doi: 10.1016/j.nuclphysbps.2010.08.011.

- [148] F. Caola, K. Melnikov, R. Rötsch, and L. Tancredi. QCD corrections to W^+W^- production through gluon fusion. *Phys. Lett. B*, 754:275, 2016. doi: 10.1016/j.physletb.2016.01.046.
- [149] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush. FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order. *Comput. Phys. Commun.*, 182:2388, 2011. doi: 10.1016/j.cpc.2011.06.008.
- [150] M. V. Garzelli, A. Kardos, C. G. Papadopoulos, and Z. Trocsanyi. $t\bar{t}W^\pm$ and $t\bar{t}Z$ hadroproduction at NLO accuracy in QCD with parton shower and hadronization effects. *JHEP*, 11:056, 2012. doi: 10.1007/JHEP11(2012)056.
- [151] LHC Higgs Cross Section Working Group. Handbook of LHC Higgs cross sections: 3. Higgs Properties. *CERN*, 2013. doi: 10.5170/CERN-2013-004.
- [152] W. Beenakker et al. Production of charginos, neutralinos, and sleptons at hadron colliders. *Phys. Rev. Lett.*, 83:3780, 1999. doi: 10.1103/PhysRevLett.83.3780. [Erratum: 10.1103/PhysRevLett.100.029901].
- [153] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering. Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV. *JHEP*, 10:081, 2012. doi: 10.1007/JHEP10(2012)081.
- [154] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering. Precision predictions for electroweak superpartner production at hadron colliders with Resummino. *Eur. Phys. J. C*, 73:2480, 2013. doi: 10.1140/epjc/s10052-013-2480-0.
- [155] W. Beenakker, R. Hopker, M. Spira, and P. M. Zerwas. Squark and gluino production at hadron colliders. *Nucl. Phys. B*, 492:51, 1997. doi: 10.1016/S0550-3213(97)80027-2.
- [156] A. Kulesza and L. Motyka. Threshold resummation for squark-antisquark and gluino-pair production at the LHC. *Phys. Rev. Lett.*, 102:111802, 2009. doi: 10.1103/PhysRevLett.102.111802.
- [157] A. Kulesza and L. Motyka. Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC. *Phys. Rev. D*, 80:095004, 2009. doi: 10.1103/PhysRevD.80.095004.
- [158] W. Beenakker, S. Brensing, M. Kramer, A. Kulesza, E. Laenen, and I. Niessen. Soft-gluon resummation for squark and gluino hadroproduction. *JHEP*, 12:041, 2009. doi: 10.1088/1126-6708/2009/12/041.
- [159] W. Beenakker, S. Brensing, M. Kramer, A. Kulesza, E. Laenen, L. Motyka, and I. Niessen. Squark and gluino hadroproduction. *Int. J. Mod. Phys. A*, 26:2637, 2011. doi: 10.1142/S0217751X11053560.
- [160] NNPDF collaboration, R. D. Ball, et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015. doi: 10.1007/JHEP04(2015)040.
- [161] CMS Collaboration. Event generator tunes obtained from underlying event and multiparton scattering measurements. *Eur. Phys. J. C*, 76:155, 2016. doi: 10.1140/epjc/s10052-016-3988-x.
- [162] CMS Collaboration. Investigations of the impact of the parton shower tuning in PYTHIA 8 in the modelling of $t\bar{t}$ at $\sqrt{s} = 8$ TeV and 13 TeV. CMS Physics Analysis Summary CMS-PAS-TOP-16-021, CERN, 2016.

- [163] A. Kalogeropoulos and J. Alwall. The SysCalc code: A tool to derive theoretical systematic uncertainties. 2018.
- [164] CMS Collaboration. The Fast Simulation of the CMS detector at LHC. In *J. Phys. Conf. Ser.*, volume 331, page 032049, 2011. doi: 10.1088/1742-6596/331/3/032049.
- [165] CMS Collaboration. Search for top-squark pair production in the single-lepton final state in pp collisions at $\sqrt{s} = 8$ TeV. *Eur. Phys. J. C*, 73:2677, 2013. doi: 10.1140/epjc/s10052-013-2677-2.
- [166] C. G. Lester and D. J. Summers. Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders. *Phys. Lett.*, B463:99–103, 1999. doi: 10.1016/S0370-2693(99)00945-4.
- [167] O. Behnke, K. Kröniger, G. Schott, and T. Schörner-Sadenius. *Data analysis in high energy physics: a practical guide to statistical methods*. Wiley-VCH, Weinheim, 2013.
- [168] CMS collaboration. Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV. *Phys. Rev.*, D95:092001, 2017. doi: 10.1103/PhysRevD.95.092001.
- [169] CMS collaboration. Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV. *Eur. Phys. J.*, C75:542, 2015. doi: 10.1140/epjc/s10052-015-3709-x.
- [170] CMS collaboration. Measurement of the $t\bar{t}$ production cross section in the all-jets final state in pp collisions at $\sqrt{s} = 8$ TeV. *Eur. Phys. J.*, C76:128, 2016. doi: 10.1140/epjc/s10052-016-3956-5.
- [171] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011. doi: 10.1140/epjc/s10052-011-1554-0. [Erratum: 10.1140/epjc/s10052-013-2501-z].
- [172] T. Junk. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A*, 434:435, 1999. doi: 10.1016/S0168-9002(99)00498-2.
- [173] A. L. Read. Presentation of search results: The CL_s technique. *J. Phys. G*, 28:2693, 2002. doi: 10.1088/0954-3899/28/10/313.