

**A SEARCH FOR ANOMALOUS PRODUCTION OF
MULTILEPTON EVENTS USING 19.5 fb^{-1} OF $\sqrt{S} = 8 \text{ TEV}$
LHC DATA**

By

EMMANUEL CONTRERAS-CAMPANA

**A dissertation submitted to the
Graduate School—New Brunswick
Rutgers, The State University of New Jersey
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
Graduate Program in Physics and Astronomy**

written under the direction of

Dr. Sunil Somalwar

and approved by

New Brunswick, New Jersey

May, 2015



© 2015

Emmanuel Contreras-Campana

ALL RIGHTS RESERVED

ABSTRACT OF THE DISSERTATION

A search for anomalous production of multilepton events using 19.5 fb^{-1} of $\sqrt{s} = 8 \text{ TeV}$ LHC data

By EMMANUEL CONTRERAS-CAMPANA

Dissertation Director:

Dr. Sunil Somalwar

An inclusive search for physics beyond the standard model in events with at least three leptons is presented. The search is based on a sample of pp collision data, corresponding to an integrated luminosity of 19.5 fb^{-1} , collected at a center-of-mass energy $\sqrt{s} = 8 \text{ TeV}$ with the CMS detector at the LHC in 2012. The data is divided into exclusive search channels characterized by the number and flavor of the leptons, the number of opposite-sign, same-flavor (OSSF) lepton pairs, the invariant mass of the OSSF lepton pair, the presence or absence of hadronically decaying τ -leptons, and the number of identified bottom-quark jets. Additionally, kinematic properties of the event, such as the scalar sum of jet transverse momenta and the magnitude of the missing transverse energy in the event, are considered. Standard model backgrounds are estimated using data control samples and from samples of simulated events. The number of observed events are consistent with the expectations from standard model processes. We therefore place 95% confidence level upper limits on various supersymmetric models that yield multilepton final states. In particular, scenarios that predict production of Higgs bosons arising from supersymmetric decays are examined.

Acknowledgments

First and foremost, I would like to thank my loving parents Victor Contreras and Jeannette Campaña. They supported and encouraged me throughout the years to pursue my dreams. Their guidance helped me be the person I am today. I will forever be thankful to them. *Muchas gracias queridos padres.* I thank my wife Megan and twin brother Christian for bringing so much joy to my life. They made this endeavor all the worth while. Thank you Megan for all your love, encouragement, and support. Thank you Christian for all the inspiration you have given me along the way. My life would not have been the same without you. I would also like to take the opportunity to express my deep gratitude to the Parker family. To Richard, Susan, and Karen, thanks for your love and patience, as well as for all the great Friday night dinners we have shared together. Family is what gives life meaning and, therefore, I would like to take the time to acknowledge Scott and Ann Parker. They have shared their home and family with me. Thank you for making every one of our family gatherings so memorable and special. A special thanks to my sister-in-law Claudia Seitz for always taking the time to answer all my experimental questions but more importantly for being a great friend.

This dissertation would not have been possible, if it were not for the many people who have assisted me along the way. I am very thankful to my advisor, Sunil Somalwar. He introduced me to the fascinating world of particle physics and guiding me on its path. Thank you Sunil for making me a better experimentalist. My thesis relied upon the foundations established by post doctoral researchers, Richard Gray and Matthew Walker. Richard began the work on the multilepton analysis and Matt was instrumental in making our group as efficient as possible. In addition, I would like to acknowledge the theory professor Scott

Thomas. Thank you Scott for all your SUSY models and expertise. I am very proud of how much we accomplished with Run I of the LHC. It is my sincere hope that the group may one day discover physics beyond the standard model. It has been a great pleasure working with my fellow graduate students Patrick, Peter, Shruti, and Sanjay. I have learned so much from everyone of you. Thank you Patrick for keeping me company late in to the night producing Monte Carlo simulations. Your sense of humor gave me the energy I needed to get through those long nights.

Lastly, I would also like to thank my dear friend Ignacio Rios Caviedes and his family for making the time I spent in Chile an amazing period of my life. They have always been in my thoughts and in my heart from the moment I left Chile to come to the US to study physics and follow my dreams. Thank you Lord for creating this incredible universe for us to explore.

Dedication

I dedicate this dissertation to my family, especially...
to my wife Megan for her patience and understanding,
to my brother Christian for opening my eyes to the world,
to my sister-in-law Claudia for her encouragement, and
to my parents Victor and Jeannette for instilling
the importance of hard work.

Table of Contents

Abstract	ii
Acknowledgments	iii
Dedication	v
List of Tables	xi
List of Figures	xv
1. Introduction	1
1.1. A brief history of particle physics	1
2. Overview of the Standard Model and Supersymmetry	3
2.1. Theoretical overview	3
2.2. The standard model of particle physics	5
2.2.1. Quantum Chromodynamics	6
2.2.2. Electroweak interactions and symmetry breaking	8
The Higgs Mechanism	10
2.3. Physics beyond the standard model	16
2.4. Supersymmetry	18
2.4.1. Gauge-mediated supersymmetry breaking	21
2.4.2. Simplified model spectrum scenarios	22
3. Experimental Apparatus	23

3.1. The Large Hadron Collider	23
3.1.1. LHC structure	25
3.2. The Compact Muon Solenoid detector	25
3.2.1. Detector coordinates	26
3.2.2. Superconducting Solenoid Magnet	28
3.2.3. Tracking system	28
Pixel detector	29
Silicon strip detector	30
3.2.4. Calorimetry system	31
Electromagnetic calorimeter	32
Hadron calorimeter	35
3.2.5. Muon system	37
3.3. Trigger system	39
3.4. Luminosity measurement	41
4. Collision Data and Simulations	44
4.1. Collision data	44
4.2. Simulated samples	46
4.2.1. Monte Carlo simulations	46
4.2.2. Background samples	50
4.2.3. Signal samples	51
5. Event Reconstruction	55
5.1. Particle-Flow event reconstruction	55
5.2. Track reconstruction	56
5.3. Primary vertex reconstruction	58
5.4. Photon reconstruction	59

5.5. Lepton reconstruction	60
5.5.1. Muon reconstruction	61
5.5.2. Electron reconstruction	62
5.5.3. Hadronic τ -lepton reconstruction	63
5.6. Jet reconstruction	65
5.6.1. Jet algorithm	65
5.6.2. b -tagged jets	67
5.7. E_T^{miss} reconstruction	68
6. Physics Object and Event Selections	70
6.1. Objection identification and selection	70
6.1.1. Muon selections	71
6.1.2. Electron selections	72
6.1.3. τ_h -lepton selections	74
6.1.4. Photon selections	75
6.1.5. Jet selections	75
b -jet identification	77
6.1.6. E_T^{miss} selection	78
6.2. Measurement of the lepton trigger efficiency	79
6.3. Measurement of the lepton isolation and identification efficiency	81
6.4. b -tagging scale factor	87
6.5. Event selection	88
7. Event Classification	89
7.1. Search strategy	89
7.2. Background estimation from simulation	91
7.2.1. Backgrounds from WZ and ZZ production	91

7.2.2.	Background from $t\bar{t}$ production	94
7.2.3.	Background from rare standard model processes	96
7.3.	Background estimation from data-derived methods	96
7.3.1.	General principles of data-derived methods	97
7.3.2.	Backgrounds from misidentified electrons and muons from jets	98
7.3.3.	Background from jets misidentified as τ_h leptons	104
7.3.4.	Background from asymmetric internal photon conversions	107
8.	Experimental Uncertainties	110
8.1.	An overview on experimental uncertainties	110
8.2.	Sources of systematic uncertainties from data-derived methods	111
8.3.	Sources of systematic uncertainties for simulations	112
8.4.	Summary of systematic uncertainties	120
9.	Statistical Analysis Method	121
9.1.	Statistical technique	121
9.2.	Limit setting procedure	122
9.3.	LHC-type CL_s method	123
9.3.1.	Likelihood function	124
9.3.2.	Test-statistics	125
9.3.3.	Observed limit	127
9.3.4.	Median expected limit and uncertainty bands	130
9.4.	Limit calculation with LandS	131
10.	Experimental Results	133
10.1.	Multilepton results	133
10.2.	Interpretation of the results	140
10.2.1.	The natural higgsino NLSP scenario	140

10.2.2. The wino NLSP with a Higgs boson scenario	145
10.2.3. The slepton co-NLSP scenario	148
10.2.4. The stau-(N)NLSP scenario	150
10.2.5. The third-generation SMS T1tttt scenario	153
10.2.6. The third-generation SMS T6ttWW scenario	154
10.3. Conclusion	157
Appendix A. Observation and SM Background Plots	158
Appendix B. Trigger List	168
Appendix C. Study of excess	170
C.1. Breakdown of background estimation by SM contribution	171
C.2. Lepton flavor breakdown of observations	171
C.2.1. Check total charge of leptons	172
C.2.2. Systematic misestimation	172
C.2.3. Allow background sources to float	173
C.3. Expected number of channels with deviations	174
C.4. $3\ell + 1\tau_h$ channel comparison plots	175
Appendix D. Additional wino NLSP with a Higgs boson results	180
Bibliography	182

List of Tables

2.1. Properties of an elementary particle in the standard model include charge, spin, lepton number, baryon number, and mass. The electric charge in the table is given in units of the elementary charge e and spin in its of $\hbar$	5
2.2. Electroweak quantum numbers for fermion doublets, where Y corresponds to the weak hypercharge, T the weak isospin, and T_3 the third component of T	12
2.3. MSSM particle content, both gauge and mass eigenstates are show, its assumed intergenerational mixing is negligible.	20
3.1. LHC design parameters for pp collisions in 2012.	26
4.1. Dataset names from the pp collisions.	45
4.2. List of simulation samples for background estimations. All samples are from the <code>Summer12_DR53X-PU_S10_START53_V7A-v1</code> MC production campaign with the exception of the fully leptonic $t\bar{t}$, semi-leptonic $t\bar{t}$, and tbZ samples which are from the <code>v2</code> , <code>ext-v1</code> , and <code>V7C-v1</code> versions, respectively.	51
4.3. Number of generated events and cross sections for background simulations given in Table 4.2.	52
6.1. Muon selection requirements.	72
6.2. Electron selection requirements. Several electron ID criteria are different for the barrel ($ \eta < 1.44$) and endcap ($1.56 < \eta < 2.4$) regions. Electrons in the gap region $1.4442 < \eta < 1.566$ are rejected.	73
6.3. τ_h -lepton selection requirements.	75
6.4. Photon selection requirements. Several photon ID criteria are different for the barrel ($ \eta < 1.44$) and endcap ($1.56 < \eta < 2.4$) regions.	76

6.5. Jet selection requirements.	77
6.6. Lepton trigger efficiency for the “OR” of all dielectron and dimuon triggers determined using different H_T triggers paths.	81
8.1. Summary of the systematic uncertainties. When applicable, the percentage correction (on σ_{vis} for normalization effects and on the total luminosity for the afterglow effect) is also reported.	113
8.2. Brief summary of systematic uncertainty values.	120
10.1. Observed (Obs.) number of events with exactly three leptons from 19.5 fb^{-1} of 2012 data along with the expected (Exp.) number of SM background events. All channels are exclusive and all uncertainties include both systematic and statistical components. The $\text{OSSF}n$ designation indicates the number of e^+e^- and $\mu^+\mu^-$ pairs in the event, “On-Z” refers to events with at least one OSSF lepton pair with dilepton mass between 75 and 105 GeV, “Below-Z” and “Above-Z” refer to events with an OSSF lepton pair with mass below 75 GeV or above 105 GeV, respectively. Some search channels have been combined into coarser E_T^{miss} bins for the purposes of presentation. The channels marked with an asterisk are used as control regions and are therefore excluded from the limit calculations.	136

10.2. Observed (Obs.) number of events with four or more leptons from 19.5 fb ⁻¹ of 2012 data along with the expected (Exp.) number of SM background events. All channels are exclusive and all uncertainties include both systematic and statistical components. The OSSF <i>n</i> designation indicates the number of e ⁺ e ⁻ and μ ⁺ μ ⁻ pairs in the event, “On-Z” refers to events with at least one OSSF lepton pair with dilepton mass between 75 and 105 GeV, and “Off-Z” refers to events where all possible OSSF lepton pairs fall outside this mass range. Some search channels have been combined into coarser E_T^{miss} bins for the purposes of presentation. The channel marked with an asterisk is used as a control region and is therefore excluded from the limit calculations . .	137
10.3. Observed (Obs.) number of events and expected (Exp.) number of SM background events, along with the number of signal (Sig.) events, in the top five most sensitive search channels for the model point $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = 130$ GeV, and $m_{\tilde{\chi}_1^0} = 1$ GeV. All channels shown have exactly three selected leptons, a veto on tagged b-jets, and $H_T < 200$ GeV. The results are categorized by the presence of an OSSF lepton pair with invariant mass below 75 GeV (above 105 GeV) given by the “below-Z” (“above-Z”) designation, the E_T^{miss} in the event, and the number of τ_h candidates.	147
C.1. Most discrepant channels for stau-NLSP scenario for the grid point with smuon/selectron mass of 150 GeV and stau mass of 50 GeV (lower half) on the observed lower exclusion contour. All channels have no b jets and $H_T < 200$ GeV.	170
C.2. Most discrepant channels for stau-NNLSP for the grid point with smuon/selectron mass of 170 GeV and stau mass of 230 GeV (upper half) on the observed upper exclusion contour. All channels have no b jets and $H_T < 200$ GeV.	170
C.3. SM background breakdown for the 4-lepton channel with OSSF1, off-Z, 1 τ_h -lepton, no b jets, $H_T < 200$ GeV, and E_T^{miss} between 0 – 50 GeV.	171

C.4. Flavor breakdown of the 4-lepton channel with OSSF1, off-Z, 1 τ_h -lepton, no b jets, $H_T < 200$ GeV, and E_T^{miss} between 0 – 50 GeV.	172
C.5. Check total charge of leptons. Need much more collision data to distinguish between the different cases based on charges.	172
C.6. The probability of having n out of 64 E_T^{miss} distributions with p-values $\leq 1.04\%$ for various values of n	175
D.1. Observed (Obs.) number of events and expected (Exp.) number of SM background events, along with the number of signal (Sig.) events, in the 5 best search channels for the model point $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = 150$ GeV, and $m_{\tilde{\chi}_1^0} = 1$ GeV. All channels shown have exactly three selected leptons, a veto on tagged b-jets, and $H_T < 200$ GeV. The results are categorized by the presence of an OSSF pair with invariant mass below 75 GeV (above 105 GeV) given by the “below-Z” (“above-Z”) designation, the E_T^{miss} in the event, and the number of τ_h candidates.	180
D.2. Results for the $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = 200$ GeV, $m_{\tilde{\chi}_1^0} = 1$ GeV model point. Details are the same as in Table D.1.	180
D.3. Results for the $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = 300$ GeV, $m_{\tilde{\chi}_1^0} = 1$ GeV model point. Details are the same as in Table D.1.	181
D.4. Results for the $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = 400$ GeV, $m_{\tilde{\chi}_1^0} = 1$ GeV model point. Details are the same as in Table D.1.	181

List of Figures

2.1. Standard model of elementary particles physics.	4
2.2. Feynman diagram of the perturbative expansion of the Higgs boson propagator.	17
3.1. Schematic representation of the LHC's injection chain composed of multiple smaller accelerators.	24
3.2. A schematic view of the CMS detector.	27
3.3. Schematic cross section through the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules which de- liver stereo hits. Disposition of the different detectors in the silicon tracker. PIXEL (red) refers to silicon pixel detectors while TIB, TID, TOB and TEC (blue) all refer to silicon strip detectors.	30
3.4. Layout of the pixel detectors in the CMS tracker showing the pixel barrel (TPB) and pixel endcaps (TPE), in black and pink, respectively.	31
3.5. Geometric view of one quarter of the ECAL system showing the locations of the electromagnetic barrel (EB), electromagnetic endcaps (EB), and electro- magnetic preshower (ES).	33
3.6. A transverse section of the ECAL system showing the arrangement of crystal modules, supermodules, preshower, and endcaps.	34
3.7. A longitudinal view of the CMS HCAL detector showing the locations of the hadron barrel (HB), hadron endcap (HE), hadron outer (HO), and hadron forward (HF) calorimeters.	36

3.8.	A layout of one quarter of the CMS muon system showing the location of the drift tube (DT), resistive plate chamber (RPC), and cathode strip chambers (CSC) systems.	39
3.9.	Level-1 trigger decision flow of the CMS detector before data is being transferred to the DAQ.	41
3.10.	Peak luminosity (top left), integrated luminosity (top right), and total integrated luminosity (bottom) per day, delivered (blue) to and recorded by the CMS detector (yellow), for pp collisions at 8 TeV center-of-mass energy in 2012.	43
6.1.	Double muon “OR” efficiency (left) and double electron “OR” efficiency (right) as a function of the sub-leading muon and electron p_T , respectively.	80
6.2.	The isolation (top left) and identification (top right) efficiency as a function of the probe muon p_T . The isolation (middle left) and identification (middle right) efficiency as a function of the probe electron p_T . The isolation efficiency (bottom) as a function of the probe τ_h -lepton p_T	83
6.3.	The isolation (top left) and identification (top right) efficiency scale factors as a function of the probe muon p_T . The isolation (middle left) and identification (middle right) efficiency as a function of the probe electron p_T . The isolation efficiency scale factor (bottom) as a function of the probe τ_h -lepton p_T	85
7.1.	Distribution of M_T (left) and E_T^{miss} (right) in a data control region enriched with WZ events. We obtain a normalization scale factor from the M_T distribution and validate it in the E_T^{miss} distribution.	92
7.2.	Distribution for the four lepton invariant mass in a data control region enriched with ZZ events.	93
7.3.	Distribution of H_T (left) and E_T^{miss} (right) in a data control region enriched in $t\bar{t}$ events.	95

7.4.	Distribution of relative isolation of non-prompt muons (left) and non-prompt electrons (right) in the single muon data control region enriched in $t\bar{t}$ events. A scale factor of 1.5 is applied to both distributions, which is obtained from measuring the discrepancy between simulation and data for the plot on the left in the relative isolation range of $0 - 0.15$	96
7.5.	Efficiency ratio vs R_{dxy}	103
7.6.	Illustration of isolation distribution for τ_h candidates in two different jet p_T spectra.	105
7.7.	Distribution of isolation for τ_h candidates for different ranges of $\sum p_T^{\text{track}}$. . .	106
7.8.	Distribution of f_τ versus f_{SB} for τ_h candidates with p_T between $20 - 40$ GeV (left) and $40 - 60$ GeV (right) in a data control region, with exactly one on-Z OSSF lepton pair, low E_T^{miss} , and low H_T	106
7.9.	A Feynman diagram of a Z boson decaying to an OSSF lepton pair, where one of the leptons undergoes asymmetric internal photon conversion into a pair of leptons such that one of them fails to be properly reconstructed or does not pass lepton selection requirements. The label ℓ' indicates that the leptons from the photon conversion can be of the same flavor, but not necessarily, as those leptons from the Z boson decay.	108
7.10.	Distribution for a three-body μ invariant mass in a tri- μ control region. . .	109
9.1.	Test statistic distributions for ensembles of pseudo-data generated for the alternative (red) and null (blue) hypotheses. See the text for definitions of the test statistic and methodology of generating pseudo-data.	129

9.2. Cumulative probability distribution for possible experimental outcomes in terms of $\mu^{95\%}$ for the case of a high background rate. The median expected limit is given by the intersection of the 50% quantile with the CDF curve. The $\pm 1\sigma$ (68%) uncertainties bands (green region) are given by the intersection of the 84% and 16% quantiles with the CDF curve, respectively. Similarly, the $\pm 2\sigma$ (95%) uncertainties bands (yellow region) are given by the intersection of the 97.5% and 2.5% quantiles with the CDF curve, respectively. The quantiles are given by the horizontal lines.	131
10.1. E_T^{miss} distribution of events with 4-leptons, one OSSF lepton pair off-Z, $1-\tau_h$, $H_T < 200$ GeV, and no b-jets.	135
10.2. E_T^{miss} distribution of events with 3-leptons, no OSSF lepton pairs, no τ_h , $H_T > 200$ GeV (top) or $H_T < 200$ GeV (bottom), and 0 b-jets (left) or ≥ 1 b-jet (right).	138
10.3. E_T^{miss} distribution of events with 3-leptons, one OSSF lepton pair below-Z, no τ_h , $H_T > 200$ GeV (top) or $H_T < 200$ GeV (bottom), and 0 b-jets (left) or ≥ 1 b-jet (right).	138
10.4. E_T^{miss} distribution of events with 3-leptons, one OSSF lepton pair on-Z, no τ_h , $H_T > 200$ GeV (top) or $H_T < 200$ GeV (bottom), and 0 b-jets (left) or ≥ 1 b-jet (right).	139
10.5. E_T^{miss} distribution of events with 3-leptons, one OSSF lepton pair above-Z, no τ_h , $H_T > 200$ GeV (top) or $H_T < 200$ GeV (bottom), and 0 b-jets (left) or ≥ 1 b-jet (right).	139
10.6. Feynman diagram and superpartner mass spectrum for the strong (top) and electroweak (bottom) SUSY production of the natural higgsino NLSP scenario, where the particles in parentheses (top left) have soft p_T spectrum. .	141

- 10.7. The 95% confidence level upper limits for the natural higgsino NLSP scenario on the cross section times branching fraction $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G}/Z\tilde{G})$. The three branching fractions of $\tilde{\chi}_1^0$ presented are $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G}) = 1.0$ (top), $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G}) = \mathcal{B}(\tilde{\chi}_1^0 \rightarrow Z\tilde{G}) = 0.5$ (middle), and $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow Z\tilde{G}) = 1.0$ (bottom). Both strong and electroweak SUSY production processes are taken into consideration. The $\tilde{\chi}_1^\pm$ mass is fixed at 150 GeV (left) and 350 GeV (right). The region of the bottom left of the contours is excluded. The region to the left of the vertical line is unphysical and limited by the $\tilde{\chi}_1^\pm$ mass. The green and yellow bands show the $\pm 1\sigma$ and $\pm 2\sigma$ variations on the expected limit due to experimental uncertainties, respectively. 143
- 10.8. The 95% confidence level upper limits for the natural higgsino NLSP scenario on the branching fraction $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G})$ with fixed $\tilde{\chi}_1^\pm$ mass of 150 GeV (upper left), 250 GeV (upper right), and 350 GeV (bottom). It is assumed that $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G}) + \mathcal{B}(\tilde{\chi}_1^0 \rightarrow Z\tilde{G}) = 1.0$, Both strong and electroweak SUSY production processes are taken into consideration. The region to the left of the vertical line is unphysical and limited by the $\tilde{\chi}_1^\pm$ mass. The green band shows the $\pm 1\sigma$ variations on the expected limit due to experimental uncertainties. 144
- 10.9. The 95% confidence level lower limits for the natural higgsino NLSP scenario in the top squark versus chargino mass plane. The three branching fractions of $\tilde{\chi}_1^0$ presented are $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G}) = 1.0$ (top left), $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow Z\tilde{G}) = 1.0$ (top right), and $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G}) = \mathcal{B}(\tilde{\chi}_1^0 \rightarrow Z\tilde{G}) = 0.5$ (bottom). Both strong and electroweak SUSY production processes are taken into consideration. The region to the left and below the contours is excluded. The region above the diagonal line is unphysical. The green band shows the $\pm 1\sigma$ variations on the expected limit due to experimental uncertainties. 145

10.10	Feynman diagram and superpartner mass spectrum for the electroweak SUSY production of chargino-neutralino pairs for the wino NLSP with a Higgs boson scenario.	146
10.11	The 95% confidence level upper limits for the wino NLSP with a Higgs boson scenario on the $\tilde{\chi}_1^\pm \tilde{\chi}_1^0$ cross section times branching fraction $\mathcal{B}(\tilde{\chi}_1^\pm \tilde{\chi}_2^0 \rightarrow W^\pm H + \tilde{\chi}_1^0)$. The green band shows the $\pm 1\sigma$ variations on the expected limit due to experimental uncertainties.	147
10.12	The E_T^{miss} distributions of the top five most sensitive search channels for the model point $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = 130$ GeV, and $m_{\tilde{\chi}_1^0} = 1$ GeV.	148
10.13	Feynman diagram and superpartner mass spectrum for the strong SUSY production of squark and/or gluino pairs for the slepton co-NLSP scenario.	149
10.14	The 95% confidence level lower limits for the slepton co-NLSP scenario in the gluino versus chargino mass plane. Masses to the left of the solid black curve are excluded. The green band shows the $\pm 1\sigma$ variations on the expected limit due to experimental uncertainties.	150
10.15	Feynman diagrams of right-handed smuon or right-handed selectron pair production (top left) and stau pair production (top right) in proton-proton collisions followed by decays leading to a final state with four leptons along with LSPs along with the a schematic superpartner mass spectrum for the GMSB stau-(N)NLSP scenario (bottom).	151
10.16	The 95% confidence level lower limits for the stau-NLSP (below diagonal line) and stau-NNLSP (above diagonal line) scenarios in the degenerate smuon and selectron versus stau mass plane. Masses to the left of the solid black curve are excluded. The green and yellow bands show the $\pm 1\sigma$ and $\pm 2\sigma$ variations on the expected limit due to experimental uncertainties, respectively.	152
10.17	Feynman diagram and superpartner mass spectrum for the SMS T1tttt scenario.	153

10.18	The 95% confidence level lower limits for the T1tttt scenario in the gluino versus neutralino mass plane. Masses to the left of the solid black curve and below the diagonal line are excluded. The green band shows the $\pm 1\sigma$ variations on the expected limit due to experimental uncertainties.	154
10.19	Feynman diagram and schematic superpartner mass spectrum for the SMS T6ttWW scenario.	155
10.20	The 95% confidence level lower limits for the T6ttWW scenario in the bottom squark versus chargino mass. Masses to the left of the solid black curve and below the diagonal line are excluded. The green band shows the $\pm 1\sigma$ variations on the expected limit due to experimental uncertainties.	156
A.1.	3-leptons + OSSF0 + Tau0 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	158
A.2.	3-leptons + OSSF0 + Tau0 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	158
A.3.	3-leptons + OSSF0 + Tau1 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	159
A.4.	3-leptons + OSSF0 + Tau1 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	159
A.5.	3-leptons + OSSF1 + on-Z + Tau0 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	159
A.6.	3-leptons + OSSF1 + on-Z + Tau0 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	160
A.7.	3-leptons + OSSF1 + on-Z + Tau1 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	160
A.8.	3-leptons + OSSF1 + on-Z + Tau1 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	160

A.9. 3-leptons + OSSF1 + above-Z + Tau0 + $H_T > 200$ GeV and 0 b-jets (left)	
or ≥ 1 b-jet (right)	161
A.10.3-leptons + OSSF1 + above-Z + Tau0 + $H_T < 200$ GeV and 0 b-jets (left)	
or ≥ 1 b-jet (right)	161
A.11.3-leptons + OSSF1 + above-Z + Tau1 + $H_T > 200$ GeV and 0 b-jets (left)	
or ≥ 1 b-jet (right)	161
A.12.3-leptons + OSSF1 + above-Z + Tau1 + $H_T < 200$ GeV and 0 b-jets (left)	
or ≥ 1 b-jet (right)	162
A.13.3-leptons + OSSF1 + below-Z + Tau0 + $H_T > 200$ GeV and 0 b-jets (left)	
or ≥ 1 b-jet (right)	162
A.14.3-leptons + OSSF1 + below-Z + Tau0 + $H_T < 200$ GeV and 0 b-jets (left)	
or ≥ 1 b-jet (right)	162
A.15.3-leptons + OSSF1 + below-Z + Tau1 + $H_T > 200$ GeV and 0 b-jets (left)	
or ≥ 1 b-jet (right)	163
A.16.3-leptons + OSSF1 + below-Z + Tau1 + $H_T < 200$ GeV and 0 b-jets (left)	
or ≥ 1 b-jet (right)	163
A.17.4-leptons + OSSF0 + Tau1 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	163
A.18.4-leptons + OSSF1 + on-Z + Tau0 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	164
A.19.4-leptons + OSSF1 + on-Z + Tau0 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	164
A.20.4-leptons + OSSF1 + on-Z + Tau1 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	164
A.21.4-leptons + OSSF1 + on-Z + Tau1 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	165

A.22.4-leptons + OSSF1 + off-Z + Tau0 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	165
A.23.4-leptons + OSSF1 + off-Z + Tau0 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	165
A.24.4-leptons + OSSF1 + off-Z + Tau1 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	166
A.25.4-leptons + OSSF1 + off-Z + Tau1 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	166
A.26.4-leptons + OSSF2 + on-Z + Tau0 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	166
A.27.4-leptons + OSSF2 + on-Z + Tau0 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	167
A.28.4-leptons + OSSF2 + off-Z + Tau0 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	167
A.29.4-leptons + OSSF2 + off-Z + Tau0 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)	167
C.1. Investigation of excess in four leptons, one of which is a τ_h -lepton. The top row has a Z candidate and the bottom row does not. On the left is the H_T distribution with expected backgrounds. The right plot shows the H_T distribution where the background has been normalized to match the observation.	176
C.2. Investigation of excess in four leptons, one of which is a τ_h -lepton. The top row has a Z candidate and the bottom row does not. The left is the muon p_T distribution with expected backgrounds. The right plot shows the muon p_T distribution where the background has been normalized to match the observation.	177

C.3. Investigation of excess in four leptons, one of which is a τ_h -lepton. The top row has a Z candidate and the bottom row does not. On the left is the electron p_T distribution with expected backgrounds. The right plot shows the electron p_T distribution where the background has been normalized to match the observation.	178
C.4. Investigation of excess in four leptons, one of which is a τ_h -lepton. The top row has a Z candidate and the bottom row does not. On the left is the τ_h -lepton p_T distribution with expected backgrounds. The right plot shows the τ_h -lepton p_T distribution where the background has been normalized to match the observation.	179

Chapter 1

Introduction

1.1 A brief history of particle physics

Many experimental discoveries of new phenomena during the first half of the 20th century began to challenge existing theories. In 1936, Carl D. Anderson and Seth Neddermeyer observed a new particle produced by cosmic rays, referred to as the muon, which was later confirmed by cloud chamber experiments by Jabez C. Street and E. C. Stevenson in 1937 [1]. Wolfgang Pauli postulated the neutrino in 1930, in order to explain the shape of the electron energy spectrum in beta decays. In 1956, neutrinos were experimentally detected by Clyde L. Cowan and Frederick Reines [2], thereafter, a second neutrino type, the muon neutrino, was discovered a few years later by Leon M. Lederman, Melvin Schwartz and Jack Steinberger [3]. During the 1960's numerous new mesons and baryons were discovered creating a zoo of particles. Murray Gell-Mann [4] and, independently, George Zweig [5] proposed a classification scheme for all these particles by suggesting that hadrons were composed of smaller constituents, with new quantum numbers. This led to the development of the “quark model”, which predicted the existence of mesons and baryons with strange quark constituents, exemplified by the Ω^- baryon discovered at Brookhaven in 1964.

Quantum field theories, developed to describe the dynamics of particle interactions, suffered from divergences. A new gauge theory was proposed by Chen N. Yang and Robert Mills in 1950's based on the principle of “local gauge invariance” [6], whereby symmetry transformations performed at one point of space-time do not affect any other. Peter Higgs, François Englert, and Robert Brout [7, 8] introduced a scalar field into the theory, resulting

in the Yang-Mills theory known as the standard model of particle physics incorporating the weak interaction. A unified perturbative model independently developed by Abdus Salam, Sheldon Glashow, and Steve Weinberg [9, 10] faced a major challenge, in order to solve the equations in the theory, the calculation needed the equations to be expanded in a power series, but only the leading order term was free of ultraviolet divergences. In the 1970's, Gerard 't Hooft demonstrated the renormalizability of the theory. The introduction of the “renormalization group” revealed that Yang-Mills gauge theories had negative β -functions. This led to the discovery of “asymptotic freedom”, a property needed to make the Yang-Mills theory adequate in describing the strong interactions and reflect an observed experimental effect known as “Bjorken scaling”. In 1975, the τ -lepton was observed and, shortly thereafter, in 1983 the $W^\pm$ and Z bosons were discovered setting the scene for the standard model of particle physics. More recently, a Higgs-like boson was discovered in 2012 [11, 12, 13, 14] completing the list of particles first proposed by the standard model.

Chapter 2

Overview of the Standard Model and Supersymmetry

In the first part of this chapter we give an overview of the standard model of particle physics by introducing a field theoretic description of particles and their interactions. In the second part of the chapter we conclude with possible extensions to the standard model, where the basic ideas of supersymmetry are introduced.

2.1 Theoretical overview

The basic constituents of matter in the known universe and their fundamental interactions can be described by a set of particles known as fermions and bosons. Elementary matter in nature is comprised of half integer spin particles, referred to as fermions, which are grouped into three generation of chiral doublets and singlets for each of the leptons and quarks. These particles also have charge conjugate partners corresponding to their anti-particle. The other category of particles are the force carriers consisting of integer spin bosons, arising from local gauge invariance of the standard model. Matter consists of elementary fermions, which are point-like structures. An overview of the fundamental particles observed in nature can be found in Figure 2.1.

Leptons are classified based on their electric charge, as electrically charged and massive or electrically neutral and almost massless. The electron is the lightest charged particle, the others are the muon and τ -lepton. The three electrically neutral particles are the respective electron, muon, and tau neutrinos, which only interact weakly with matter. Fermions can be further sub-categorized as quarks and leptons. The quarks interact strongly, weakly

mass →	$\approx 2.3 \text{ MeV}/c^2$	$\approx 1.275 \text{ GeV}/c^2$	$\approx 173.07 \text{ GeV}/c^2$	0	$\approx 126 \text{ GeV}/c^2$
charge →	$2/3$	$2/3$	$2/3$	0	0
spin →	$1/2$	$1/2$	$1/2$	1	0
	u up	c charm	t top	g gluon	H Higgs boson
QUARKS	$\approx 4.8 \text{ MeV}/c^2$	$\approx 95 \text{ MeV}/c^2$	$\approx 4.18 \text{ GeV}/c^2$	0	
	$-1/3$	$-1/3$	$-1/3$	0	
	$1/2$	$1/2$	$1/2$	1	
	d down	s strange	b bottom	γ photon	
	$0.511 \text{ MeV}/c^2$	$105.7 \text{ MeV}/c^2$	$1.777 \text{ GeV}/c^2$	$91.2 \text{ GeV}/c^2$	
	-1	-1	-1	0	
	$1/2$	$1/2$	$1/2$	1	
	e electron	μ muon	τ tau	Z Z boson	
LEPTONS	$< 2.2 \text{ eV}/c^2$	$< 0.17 \text{ MeV}/c^2$	$< 15.5 \text{ MeV}/c^2$	$80.4 \text{ GeV}/c^2$	
	0	0	0	± 1	
	$1/2$	$1/2$	$1/2$	1	
	ν_e electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino	W W boson	
					GAUGE BOSONS

Figure 2.1: Standard model of elementary particles physics [15].

and electromagnetically. These particles are grouped into fermion generations. The first generation of fermions are the building blocks of matter, of atoms, and molecules. The next two generations are heavier but do not form stable bound states since they decay far too quickly. Each generation of quarks consists of a pair of electrically charged quarks, a positively charged up-type quark (up, charm, top) and a negatively charged down-type quark (down, strange, bottom). The atomic nucleus is composed of protons and neutrons which are bound states consisting of first generation quarks. Fermions, including leptons and quarks, form baryonic matter.

In nature there are four fundamental forces between these particles, the electromagnetic interaction mediated by the photon, the weak interaction carried by the W and Z bosons, the strong interaction mediated by gluons, and the gravitational force. The standard model describes three of the four interactions, the exception being gravity, whose strength at the electroweak scale is negligible. Thus far there is no agreed upon quantum field theory of gravity, which is presumably mediated by a so-called graviton particle.

A summary of the standard model fermions and bosons with their corresponding quantum numbers, such as spin S , electric charge Q , baryon number B , and mass M can be found in Table 2.1. The next section provides a brief formalism of the standard model

description.

Table 2.1: Properties of an elementary particle in the standard model include charge, lepton number, baryon number, and mass. The electric charge in the table is given in units of the elementary charge e and the spin in units of $\hbar$.

Particle	Charge Q	Spin S	Lepton number L	Baryon number B	Mass (MeV)
Leptons					
Electron (e)	-1	1/2	1	0	0.511
Electron neutrino (ν_e)	0	1/2	-1	0	$< 2.2 \times 10^{-6}$
Muon (μ)	-1	1/2	1	0	105.7
Muon neutrino (ν_μ)	0	1/2	-1	0	< 0.17
Tau (τ)	-1	1/2	1	0	1.77×10^3
Tau neutrino (ν_τ)	0	1/2	-1	0	< 15.5
Quarks					
Up (u)	2/3	1/2	0	1/3	2.4
Down (d)	-1/3	1/2	0	1/3	4.8
Charm (c)	2/3	1/2	0	1/3	1.27×10^3
Strange (s)	-1/3	1/2	0	1/3	104
Top (t)	2/3	1/2	0	1/3	171.2×10^3
Bottom (b)	-1/3	1/2	0	1/3	4.2×10^3
Gauge bosons					
Photon (γ)	0	1	0	0	0
W boson ($W^\pm$)	± 1	1	0	0	80.4×10^3
Z boson (Z)	0	1	0	0	91.2×10^3
Gluon (g)	0	1	0	0	0
Higgs boson (H)	0	0	0	0	125.09×10^3

2.2 The standard model of particle physics

The standard model (SM) of particle physics is one of the most precise quantum mechanical theories of all time. Therefore, its excellent agreement with experimental data has served as the benchmark for the level of success by which any other proposed theory of particle physics is measured. To briefly review, the SM is a relativistic quantum field theory with an internal symmetry group,

$$\text{SU}(3)_C \otimes \text{SU}(2)_L \otimes \text{U}(1)_Y. \quad (2.1)$$

Quantum Chromodynamics is represented by the symmetry group $\text{SU}(3)_C$ of color charge and describes the strong interaction; that is the fundamental interactions between quarks

and gluons which make up hadrons. The Electroweak interaction is represented by the symmetry group $SU(2)_L \otimes U(1)_Y$ of weak isospin and weak hypercharge, respectively, and describes the weak and electromagnetic interactions; that is the fundamental interactions which leptons, quarks, photons, $W^\pm$ and Z bosons may undergo. In the standard model, the $W^\pm$ and Z bosons, and the photon, are produced by the spontaneous breaking of the electroweak symmetry group from $SU(2)_L \otimes U(1)_Y \rightarrow U(1)_{EM}$, through a process known as the Higgs mechanism. Thereby a complex scalar field, the Higgs field, receives a vacuum expectation value, such that the $W^\pm$ and Z bosons become massive while the photon remains massless. The addition of Yukawa couplings of the Higgs field gives rise to the masses of the fermions.

2.2.1 Quantum Chromodynamics

The structure of the protons colliding at the LHC is governed by the strong force and is described by the theory of quantum chromodynamics (QCD). In the SM, the strong force arises by requiring local gauge invariance under the non-Abelian group $SU(3)_C$ between the colored states of quarks, whereby the free fermion Lagrangian, given by,

$$\mathcal{L}_0 = \bar{\psi}(i\cancel{\partial} - m)\psi \quad (2.2)$$

is modified with the covariant derivative,

$$\partial_\mu \rightarrow D_\mu \equiv \partial_\mu + ig_3 t_a G_\mu^a, \quad (2.3)$$

where g_3 is a gauge coupling constant, $t_a = \lambda_a/2$ are the matrix representation for the generators of the $SU(3)_C$ gauge group, such that λ_a are the Gell-Mann matrices, ψ is the quark field in the fundamental representation of the $SU(3)_C$ gauge group, while G_μ^a are the gluon fields in the adjoint representation of the $SU(3)_C$ gauge group. We use the Feynman slash notation, for example $\cancel{\partial} = \gamma^\mu \partial_\mu$, where γ^μ are the so-called Dirac or gamma matrices.

The Dirac adjoint is defined as $\bar{\psi} \equiv \psi^\dagger \gamma^0$.

Therefore, the QCD Lagrangian [16], including terms related to the vector boson, is expressed as,

$$\mathcal{L}_{\text{QCD}} = \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}G_a^{\mu\nu}G_{\mu\nu}^a, \quad (2.4)$$

where $\not{D} = \gamma^\mu D_\mu$ and the gluon field is given by,

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_3 f_{bc}^a G_\mu^b G_\nu^c, \quad (2.5)$$

such that f_{bc}^a are the structure constants of $\text{SU}(3)_C$. It can be seen from equation 2.4 that the vector boson couples to fermion–anti-fermion pairs, as well as to itself, specifically, in three- and four-point interactions. Quarks, in addition to having electric charge, carry color charge, either r (red), g (green), or b (blue) color, or their anti-colors. Unlike photons, which are electrically neutral, gluons carry color charge, which allows for gauge self-interactions.

The strong interaction exhibits “asymptotic freedom”, such that at low energies the QCD coupling, α_S , is too large to permit perturbative calculations but at the energy scale of the LHC collisions the coupling is small enough to allow it. Non-perturbative models are needed to describe the behavior of the underlying event and hadronization process, further details are found in Sections 4.2.1 and 4.2.1. The one-loop strong running coupling constant of QCD is given by,

$$\alpha_S(Q^2) \equiv \frac{g_3^2(Q^2)}{4\pi} = \frac{\alpha_S(\mu^2)}{1 + \beta_0 \alpha_S(\mu^2) \ln\left(\frac{Q^2}{\mu^2}\right)}, \quad \text{with } \beta_0 = \frac{1}{12\pi}(11N_C - 2n_f) \quad (2.6)$$

where n_f and N_C are the number of flavors and colors, respectively, active at the energy scale Q^2 [17]. The reference scale μ^2 is called the renormalization scale, which is an unphysical parameter. The dependence of the running coupling constant α_S on Q^2 leads to asymptotic freedom at large Q^2 , where strong interactions become weaker at short distances. While at

at low Q^2 it leads to “color confinement”, where strong force increases at large distances.

The expression for the running coupling constant can be defined in terms of the the scale parameter Λ_{QCD} , where the second term in the denominator of Equation 2.6 dominates over the first term, giving,

$$\alpha_S(Q^2) \equiv \frac{1}{\beta_0 \ln \left(\frac{Q^2}{\Lambda_{\text{QCD}}^2} \right)}. \quad (2.7)$$

The range of the strong interaction is confined to nuclear length scales even though gluons are massless. This is due to the phenomenon known as “color confinement”, where color singlet states are the only free states which are allowed. Color confinement is responsible for the so-called hadronization process, which is a non-perturbative effect yielding color singlet states of quark–antiquark pairs or trios of quarks. In particle colliders, energetic quarks and gluons produced in high energy collisions hadronize into sprays of collimated particles that are only observed as reconstructed jets. As a consequence of the hadronization process information about the initial state partons, such as quark flavors, is concealed. A lot of effort goes into experimentally determining as much of the information as possible that is lost during hadronization process.

2.2.2 Electroweak interactions and symmetry breaking

Electroweak interactions arise from the gauge theory of $\text{SU}(2)_L \otimes \text{U}(1)_Y$. Besides electromagnetic interactions, fermions are subject to weak interactions. Both forces are manifestations of the unified electroweak theory that is described by the gauge symmetry $\text{SU}(2)_L \otimes \text{U}(1)_Y$. The electromagnetic force is described by the theory of quantum electrodynamics (QED) which emerges from electroweak interactions. The electromagnetic force governs the interactions among electrically charged particles. The emission of light from excited atoms is an example of electromagnetic interactions. The force is mediated by photons with an infinite range effect since they are electrically neutral and, therefore, do not experience photon-self

coupling. The weak force has a range of $\mathcal{O}(10^{-16} \text{ m})$ and is mediated by $W^\pm$ bosons and Z bosons. The beta decay of atomic nuclei is an example of a process that is mediated by weak interactions.

The gauge groups $SU(2)_L$ and $U(1)_Y$ cannot be factorized from each other since this leads to massless bosons, which is in clear contradiction with the observed masses of the $W^\pm$ and Z bosons. However, the gauge group $SU(3)_C$ can be decomposed from $SU(2)_L \otimes U(1)_Y$. In addition, gauge invariance also forbids the introduction of explicit mass terms for gauge bosons. The electroweak theory is constructed in such a way that it distinguishes between “left-handed” and “right-handed” fermions. Under the local gauge symmetry group, $SU(2)_L$, the left- and right-handed chiral fields $\psi_{L,R} = (1 \mp \gamma^5)/2$, where $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$, transform differently. Therefore, explicit mass terms for fermions, such as

$$\mathcal{L}_{\text{mass}} = -m\bar{\psi}\psi = -m(\bar{\psi}_L\psi_R + \psi_L\bar{\psi}_R), \quad (2.8)$$

are not invariant under the gauge group and, as a result, are prohibited in the same manner as in the case for gauge bosons.

The electroweak theory proposed by Glashow [10], Weinberg [9], and Salam [18] provides a unified description of the electromagnetic and weak interactions. The theory introduces a local $SU(2)_L \otimes U(1)_Y$ gauge symmetry that undergoes spontaneous symmetry breaking through the Higgs mechanism, which resolves the gauge boson and fermion mass issue while preserving gauge invariance.

The Higgs Mechanism

The simplest realization of the Higgs mechanism in the SM assumes the existence of a complex scalar $SU(2)_L$ doublet, given by,

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_3 + i\phi_4 \\ \phi_1 + i\phi_2 \end{pmatrix}, \quad (2.9)$$

where the four scalar degrees of freedom $\phi_{i=1,2,3,4}$ correspond to the real and imaginary parts of the two complex fields ϕ^+ and ϕ^- . The field ϕ is referred to as the Higgs field.

The electroweak Lagrangian before spontaneous symmetry breaking [16], which includes all scalar, gauge, and fermion fields of the SM, is given by,

$$\mathcal{L}_{EW} = |D_\mu \phi|^2 - V(\phi) - \frac{1}{4} B^{\mu\nu} B_{\mu\nu} - \frac{1}{4} W_a^{\mu\nu} W_{\mu\nu}^a + \mathcal{L}_{\text{fermions}} + \mathcal{L}_{\text{Yukawa}}, \quad (2.10)$$

where D_μ is the covariant derivative defined as,

$$D_\mu \equiv \partial_\mu + \frac{i}{2} g_1 Y B_\mu + i g_2 t_i W_\mu^i, \quad (2.11)$$

g_1 and g_2 are the gauge coupling constants for $U(1)_Y$ and $SU(2)_L$, respectively, Y is the weak hypercharge, which is the generator of the $U(1)_Y$ gauge group, and $t_i = \sigma_i/2$ are the generators of the weak isospin $SU(2)_L$ gauge group, where σ_i are the Pauli matrices.

The vector field B_μ is associated with the $U(1)_Y$ gauge symmetry and $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ is the corresponding field strength tensor, while W_μ^i are the vector fields associated with the $SU(2)_L$ gauge symmetry and $W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g \epsilon_{ijk} W_\mu^j W_\nu^k$, are the corresponding field strength tensors, where ϵ_{ijk} (i.e. Levi-Civita symbol) is the structure constant of $SU(2)_L$.

The fermion Lagrangian is defined in terms of chiral fields as,

$$\mathcal{L}_{\text{fermions}} = \bar{L}_j i \not{D} L^j + \bar{E}_j i \not{D} E^j + \bar{Q}_j i \not{D} Q^j + \bar{U}_j i \not{D} U^j + \bar{D}_j i \not{D} D^j, \quad (2.12)$$

where j is the generation of lepton, up-type quark, and down-type quark, respectively.

The left-handed electroweak doublets for leptons is,

$$L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L \quad (2.13)$$

and for quarks it is,

$$Q = \begin{pmatrix} u_L \\ d_L \end{pmatrix} = \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L \quad (2.14)$$

While the right-handed electroweak singlet for leptons is given by,

$$E = (e_R) = (e^-)_R, (\nu_\mu)_R, (\tau^-)_R \quad (2.15)$$

and for quarks it is,

$$U = (u_R) = (u)_R, (c)_R, (t)_R \quad (2.16)$$

$$D = (d_R) = (d)_R, (s)_R, (b)_R. \quad (2.17)$$

The Yukawa Lagrangian [16] is defined as,

$$\mathcal{L}_{\text{Yukawa}} = -y_{ij}^\ell \bar{L}^i \phi E^j - y_{ij}^u \epsilon^{ab} \bar{Q}_a^i (\phi^\dagger)_b U^j - y_{ij}^d \bar{Q}^i \phi D^j + h.c., \quad (2.18)$$

where y_ℓ , y_u , and y_d are the trilinear Yukawa couplings between the Higgs field and the left- and right-handed chiral fields of the charged leptons, up-type quarks, and down-type quarks, respectively, and h.c. refers to the hermitian conjugate. This Lagrangian is used later in the section to derive masses for the fermions. Table 2.2 shows a list of fermion doublets and their respective electroweak quantum numbers.

The Higgs mechanism uses the most general renormalizable scalar potential allowed by

Table 2.2: Electroweak quantum numbers for fermion doublets, where Y corresponds to the weak hypercharge, T the weak isospin, and T_3 the third component of T .

Particle type	1 st Gen.	2 nd Gen.	3 rd Gen.	T	T_3	Y
Leptons	$(\begin{smallmatrix} \nu_e \\ e \end{smallmatrix})_L$	$(\begin{smallmatrix} \nu_\mu \\ \mu \end{smallmatrix})_L$	$(\begin{smallmatrix} \nu_\tau \\ \tau \end{smallmatrix})_L$	1/2	$(\begin{smallmatrix} +1/2 \\ -1/2 \end{smallmatrix})_L$	-1
Quarks	$(\begin{smallmatrix} u \\ d \end{smallmatrix})_L$	$(\begin{smallmatrix} c \\ s \end{smallmatrix})_L$	$(\begin{smallmatrix} t \\ b \end{smallmatrix})_L$	1/2	$(\begin{smallmatrix} +1/2 \\ -1/2 \end{smallmatrix})_L$	+1/3

the SM gauge group,

$$V(\phi) = -\mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2, \quad (2.19)$$

where $\lambda > 0$.

The ground state of the Higgs field occurs at the minimum of the potential $V(\phi)$. If $\mu^2 < 0$, then the state of minimum energy is $\phi = 0$, and the potential preserves the $SU(2)_L \otimes U(1)_Y$ symmetry of the SM Lagrangian. However, if $\mu^2 > 0$, then the state of minimum energy is not $\phi = 0$, and the potential does not preserve the SM gauge group due to the spontaneous symmetry breaking. The Higgs field also acquires a vacuum expectation value (VEV). The electroweak symmetry breaking (EWSB) condition implies that,

$$\phi^\dagger\phi = |\phi| = \frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}, \quad (2.20)$$

Only the magnitude of Higgs field is fixed, defining a continuous spectrum of minima, hence the ground state is degenerate, and the gauge freedom allows for the simple choice of,

$$\langle\phi\rangle_0 = \langle 0|\phi|0\rangle = \begin{pmatrix} \langle 0|\phi^+|0\rangle \\ \langle 0|\phi^0|0\rangle \end{pmatrix}, \quad (2.21)$$

such that,

$$\langle 0 | \phi^+ | 0 \rangle = 0 \quad (2.22)$$

$$\langle 0 | \phi^0 | 0 \rangle = \frac{\mu}{\sqrt{2\lambda}} \equiv \frac{v}{\sqrt{2}}, \quad (2.23)$$

where $v/\sqrt{2}$ is the VEV of the Higgs field. So that the expectation value of the Higgs field with respect to the ground state that minimizes the Higgs potential is,

$$\langle \phi \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (2.24)$$

Working in the unitary gauge, the Higgs field is expanded around the VEV, such that the scalar doublet is written in the form,

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}, \quad (2.25)$$

where $H = H(x)$ is the Higgs boson.

All the original generators of the $SU(2)_L \otimes U(1)_Y$ symmetry are broken once the Higgs field acquires a VEV,

$$Y \langle \phi \rangle_0 \neq 0, \quad (2.26)$$

$$\sigma_i \langle \phi \rangle_0 \neq 0, \quad (2.27)$$

but a linear combination, corresponding to the electric charge,

$$Q \langle \phi \rangle_0 = \frac{1}{2}(Y + \sigma_3) \langle \phi \rangle_0 = 0, \quad (2.28)$$

remains invariant, implying that the Higgs mechanism in the SM contains a massless boson.

The electroweak Lagrangian after symmetry breaking [16] is given by,

$$\begin{aligned}
\mathcal{L}_{\text{EW}}^{\text{kinetic}} = & \frac{1}{2}(\partial^\mu H)(\partial_\mu H) - (\lambda v^2)H^2 \\
& - \frac{1}{2}W_{\mu\nu}^+ W^{-\mu\nu} + \left(\frac{vg_2}{2}\right)^2 W_\mu^+ W^{-\mu} \\
& - \frac{1}{4}Z^{\mu\nu} Z_{\mu\nu} + \frac{1}{2}\left(\frac{v\sqrt{g_1^2 + g_2^2}}{2}\right)^2 Z^\mu Z_\mu \\
& - \frac{1}{4}A^{\mu\nu} A_{\mu\nu}.
\end{aligned} \tag{2.29}$$

Only the kinetic terms of the Lagrangian are shown since these are the terms responsible for the mass spectrum of the gauge bosons. Also, the fields are written in terms of the physical gauge bosons and defined by,

$$W_\mu^\pm \equiv \frac{1}{\sqrt{2}}(W_\mu^1 \mp W_\mu^2), \tag{2.30}$$

$$Z_\mu \equiv \frac{1}{\sqrt{g_1^2 + g_2^2}}(g_2 W_\mu^3 - g_1 B_\mu), \tag{2.31}$$

$$A_\mu \equiv \frac{1}{\sqrt{g_1^2 + g_2^2}}(g_2 W_\mu^3 + g_1 B_\mu). \tag{2.32}$$

The spontaneous symmetry breaking of electroweak interactions rotates the plane of the original W^0 and B^0 vector bosons producing, as a result, the Z boson and photon. The Weinberg angle, also referred to as weak mixing angle, of the rotation is given by,

$$\sin \theta_W = \frac{g_1}{\sqrt{g_1^2 + g_2^2}} \tag{2.33}$$

After the electroweak symmetry breaking the particle spectrum now consists of massive

$W^\pm$, Z , and Higgs bosons, and one massless photon. The masses are given by,

$$m_W = \frac{1}{2}vg_2, \quad (2.34)$$

$$m_Z = \frac{1}{2}v\sqrt{g_1^2 + g_2^2}, \quad (2.35)$$

$$m_H = \mu\sqrt{2} = v\sqrt{2\lambda}, \quad (2.36)$$

$$m_\gamma = 0. \quad (2.37)$$

The Higgs mechanism in the SM has so far been shown to generate masses for the gauge bosons but the method can also produce mass terms for fermions. If the Higgs field is expanded once again about the physical vacuum $v/\sqrt{2}$, with oscillations governed by H , and substituted into Equation 2.18, then the Yukawa terms [16] become,

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} = & -\frac{1}{\sqrt{2}}(y_{ij}^\ell v) \bar{e}_L^i e_R^j - \frac{1}{\sqrt{2}}y_{ij}^\ell \bar{e}_L^i H e_R^j \\ & -\frac{1}{\sqrt{2}}(y_{ij}^u v) \bar{u}_L^i u_R^j - \frac{1}{\sqrt{2}}y_{ij}^u \bar{u}_L^i H u_R^j \\ & -\frac{1}{\sqrt{2}}(y_{ij}^d v) \bar{d}_L^i d_R^j - \frac{1}{\sqrt{2}}y_{ij}^d \bar{d}_L^i H d_R^j + h.c. \end{aligned} \quad (2.38)$$

The fermion masses are then determined from the Yukawa interaction Lagrangian, after symmetry breaking, and calculated with the expression,

$$m_{\text{fermion}} = \frac{1}{\sqrt{2}}y_f v. \quad (2.39)$$

A primary goal of the LHC was to search for the Higgs boson. A particle consistent with the SM Higgs boson was discovered in 2012 by both the CMS [12, 13] and ATLAS [14] collaborations. The measured mass of the Higgs boson [11], combining the results from both experiments in the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ \rightarrow 4\ell$ decay channels, is

$$m_H = 125.09 \pm 0.21 \text{ (stat)} \pm 0.11 \text{ (syst)} \text{ GeV}. \quad (2.40)$$

2.3 Physics beyond the standard model

Even though the SM has been tested and experimentally confirmed to a high accuracy, there still remains a number of unanswered questions and shortcomings. As previously mentioned, the SM does not include gravity, which is one of the four fundamental interactions of nature, as there are no agreed upon quantum field theory for general relativity. Also in the SM, neutrinos are defined as massless, nevertheless, it has been established experimentally that neutrinos have mass. Therefore, the SM must be adapted accordingly, but there is currently insufficient data to establish one particular description for the masses of neutrinos over another. Another shortcoming of the SM is that it is inconsistent with the observations of dark matter. Dark matter is estimated to constitute 84.5% of the total matter in the universe [19]. The existence and properties of dark matter cannot be directly observed with telescopes but can only be inferred from its gravitational effects on visible matter by measuring deviations from predictions using General Relativity. A stable particle that only interacts by means of the weak force and gravity can be a candidate for dark matter.

Gravitational interactions have a negligible effect at energies scales achieved in particle colliders but at the Planck scale, $\mathcal{O}(10^{19} \text{ GeV})$, they become significant and, as a consequence, new physics beyond the standard model (BSM) is expected to occur before or at this scale. We also expect new physics to appear at lower energies, $\mathcal{O}(1 \text{ TeV})$, due to the so-called Hierarchy Problem. In the SM, the Higgs boson mass is given by,

$$m_{\text{H}}^2 = m_0^2 + \Sigma_{\text{H}}, \quad (2.41)$$

where m_{H}^2 is the renormalized Higgs boson mass, m_0^2 is bare mass of the Higgs boson, which is an unphysical (i.e. unobservable) parameter of the quantum field theory, while Σ_{H} is the real part of the self interacting energy of the Higgs boson.

Figure 2.2 shows the propagator for the Higgs particle, represented by Feynman diagrams, used to estimate the Higgs boson mass. The first diagram on the right-hand side of

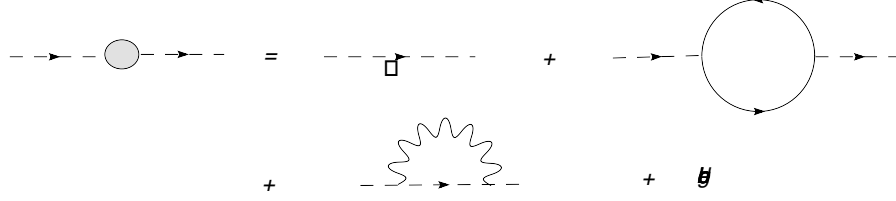


Figure 2.2: Feynman diagram of the perturbative expansion of the Higgs boson propagator [20].

the equality represents the tree level contribution, while the following diagrams represent the one-loop corrections to the Higgs boson propagator, due to two different types of particles. The first loop diagram shows the Higgs boson decaying to a fermion–anti-fermion pair then recombining to form the Higgs boson once again.

In order to evaluate the first loop diagram, we integrate over all possible momenta for the virtual particles, giving the expression [20],

$$-i\Sigma_H = -i\frac{\lambda_f^2}{8\pi^2} \left(\int_0^1 dx \, 2 \cdot \Delta \cdot \ln \frac{\Lambda^2 + \Delta}{\Delta} - \Lambda^2 \right), \quad (2.42)$$

where λ_f is the Yukawa coupling between the Higgs boson and fermion f , Λ is an ultraviolet cutoff on the Euclidean momentum, introduced for the purposes of renormalisation, and $\Delta = -x(1-x)m_H^2 + m_f^2$, such that m_f is the mass of the fermion. The second loop diagram from the gauge bosons gives similar contributions to the Higgs boson mass. Therefore, the quantum corrections to the Higgs boson mass is given by,

$$\Delta m_H^2 = \frac{\lambda_f^2}{8\pi^2} (\Lambda^2 + \dots), \quad (2.43)$$

where $\Delta m_H^2 = m_H^2 - m_0^2$. The logarithmic terms are dropped since they are much smaller than the quadratic contributing terms.

Unless no new physics beyond the standard model enters at lower energies then the ultraviolet cutoff can be set to the Planck scale, M_{Pl} . This means that quantum correction

to the Higgs boson mass will be huge due to quadratic term in Equation 2.43 being large. The bare mass of the Higgs boson would have to be very precisely “fine tuned” to be close to the ultraviolet cutoff in order for the weak scale masses to be around $\mathcal{O}(100 \text{ GeV})$. It is considered unnatural to require this level of fine-tuning and, therefore, it is speculated that a more natural mechanism should exist to produce the observed hierarchical structure of scales. This is referred to as the Hierarchy Problem. A solution to this problem proposes that new physics exists that can exactly cancel out the quadratic terms that appear in the Higgs boson mass correction. Supersymmetry is a theory that can meet this requisite, as will be discussed in the next section.

2.4 Supersymmetry

Supersymmetry (SUSY) is one of the most notable candidates for a theory beyond the SM because it resolves many issues, such as the hierarchy problem, allows for the unification of the gauge coupling constants, and may provide a candidate particle for dark matter [21, 22, 23]. Supersymmetry is a type of graded Lie algebra, or superalgebra. The generators S^a of the super algebra obey anti-commutation relations, $\{S^a, S^b\} = if^{abc}S^c$, as opposed to the typical commutation relations, $[T^a, T^b] = if^{abc}T^c$, corresponding to the generators defining a Lie algebra.

The supersymmetric charge generators, Q , change the spin of a state by a half integer. Therefore, these operators transform a fermion into a boson and vice-versa,

$$Q |fermion\rangle = |boson\rangle, \quad (2.44)$$

$$\bar{Q} |boson\rangle = |fermion\rangle. \quad (2.45)$$

In essence, the theory introduces a symmetry between fermions and bosons, predicting that every SM fermion (boson) has a corresponding bosonic (fermionic) partner. The supersymmetric partners of SM elementary particles are called “superpartners”. For example, the

superpartner of the top quark, the gluon, the $W^\pm$ and Z bosons are called top squark $\tilde{t}$, gluino $\tilde{g}$, and electroweak gauginos $\tilde{\chi}^\pm$ and $\tilde{\chi}^0$, respectively.

The supersymmetry algebra is given by,

$$\{Q_\alpha^A, \bar{Q}_{\dot{\beta}B}\} = 2\sigma_{\alpha\dot{\beta}}^\mu P_\mu \delta_B^A, \quad (2.46)$$

$$\{Q_\alpha^A, Q_\beta^B\} = \{\bar{Q}_{\dot{\alpha}A}, \bar{Q}_{\dot{\beta}B}\} = 0, \quad (2.47)$$

$$[P_\mu, Q_A^\alpha] = [P_\mu, \bar{Q}_{\dot{\alpha}A}] = 0, \quad (2.48)$$

where P_μ are the four-momentum generators of space-time translations, σ^μ are the Pauli matrices ($\mu = 0, 1, 2, 3$), which include the identity matrix (i.e. $\sigma^0 = I$). All other anti-commutation relations between the Q 's and P 's vanish. The labeling for the different copies of the generators is done with use of the Roman indices A and B, which have a range from 1 to N . The charge operator Q , as well as its charge conjugate $\bar{Q}$ carry Weyl spinor indices, which use the Van der Waerden notation (i.e. $\alpha, \beta = 1, 2$ and $\dot{\alpha}, \dot{\beta} = \dot{1}, \dot{2}$). The supersymmetric generators are allowed to have more than one copy. As a result, supersymmetries are classified according to the number of copies, N , of the generators.

The single-particle states of the supersymmetry theory are described by chiral supermultiplets, which are irreducible representations of the supersymmetry algebra. Additionally, supersymmetry requires that each supermultiplet contain both a fermion and a boson state, i.e the number of bosons, n_b , equals the number of fermions, n_f . Particles in a supermultiplet share the same representation of the gauge group and quantum numbers, such as electric charge, weak isospin, and color charge, since the generators of the supersymmetry transformations commute with the generators of the gauge transformations [24, 25].

The minimal extension to the standard model that realizes supersymmetry is the so-called minimal supersymmetric standard model (MSSM), where $N = 1$. The superpartner of the gluon (g) is the gluino ($\tilde{g}$). The superpartners of the neutral electroweak field W^0 of $SU(2)_Y$ and B of $U(1)_Y$ are the neutral “wino” ($\tilde{W}^0$) and “bino” ($\tilde{B}$), respectively.

Table 2.3: MSSM particle content, both gauge and mass eigenstates are show, its assumed intergenerational mixing is negligible.

Particle	Gauge eigenstates	Mass eigenstates
Higgs bosons		
Higgs bosons (H)	H_u^0 H_d^0 H_u^+ H_d^-	h^0 H^0 A^0 $H^\pm$
Sleptons		
Charged sleptons ($\tilde{\ell}$)	$\tilde{e}_L$ $\tilde{e}_R$ $\tilde{\mu}_L$ $\tilde{\mu}_R$ $\tilde{\tau}_L$ $\tilde{\tau}_R$	$\tilde{e}_1$ $\tilde{e}_2$ $\tilde{\mu}_1$ $\tilde{\mu}_2$ $\tilde{\tau}_1$ $\tilde{\tau}_2$
Sneutrinos ($\tilde{\nu}$)	$\tilde{\nu}_e$ $\tilde{\nu}_\mu$ $\tilde{\nu}_\tau$	$\tilde{\nu}_e$ $\tilde{\nu}_\mu$ $\tilde{\nu}_\tau$
Squarks		
Up squarks ($\tilde{U}$)	$\tilde{u}_L$ $\tilde{u}_R$ $\tilde{c}_L$ $\tilde{c}_R$ $\tilde{t}_L$ $\tilde{t}_R$	$\tilde{u}_1$ $\tilde{u}_2$ $\tilde{c}_1$ $\tilde{c}_2$ $\tilde{t}_1$ $\tilde{t}_2$
Down squarks ($\tilde{D}$)	$\tilde{d}_L$ $\tilde{d}_R$ $\tilde{s}_L$ $\tilde{s}_R$ $\tilde{b}_L$ $\tilde{b}_R$	$\tilde{d}_1$ $\tilde{d}_2$ $\tilde{s}_1$ $\tilde{s}_2$ $\tilde{b}_1$ $\tilde{b}_2$
Electroweak gauginos/Higgsinos		
Neutralinos (χ)	$\tilde{B}$ $\tilde{W}^0$ $\tilde{H}_u^0$ $\tilde{H}_d^0$	$\tilde{\chi}_1^0$ $\tilde{\chi}_2^0$ $\tilde{\chi}_3^0$ $\tilde{\chi}_4^0$
Charginos ($\chi^\pm$)	$\tilde{W}^\pm$ $\tilde{H}_u^\pm$ $\tilde{H}_d^\pm$	$\tilde{\chi}_1^\pm$ $\tilde{\chi}_2^\pm$
Strong gaugino		
Gluino ($\tilde{g}$)	$\tilde{g}$	$\tilde{g}$

While the charged electroweak fields $W^\pm$ have the corresponding “charged winos” ($\tilde{W}^\pm$) as superpartners. Table 2.3 shows the MSSM particles.

The Higgs sector in the MSSM and, in general, for any SUSY theory, unlike in the SM, needs to be extended with two Higgs $SU(2)_L$ doublets to give masses to the up-type and down-type quarks. The superpartners of the neutral Higgs fields (H_u^0 , H_d^0) are the “neutral higgsinos” ($\tilde{H}_u^0$, $\tilde{H}_d^0$), while the charged Higgs fields (H_u^+ , H_d^-) have the “charged higgsinos” ($\tilde{H}_u^+$, $\tilde{H}_d^-$) as their superpartners. Similar to the SM, once electroweak symmetry is broken the gauge and mass eigenstates of these particles are no longer the same. The Higgs fields H_u^0 , H_d^0 , H_u^+ , H_d^- combine to form five Higgs bosons, specifically, two CP-even neutral states, h^0 and H^0 , a CP-odd neutral state, A^0 , and two charged states $H^\pm$. The lightest of the neutral Higgs particles has very similar properties to those of the SM Higgs boson. The bino, neutral wino, and neutral higgsinos combine to form the neutralino mass eigenstates $\tilde{\chi}_{1,2,3,4}^0$. The charged winos and charged higgsinos mix to form the chargino mass eigenstates $\tilde{\chi}_{1,2}^\pm$. As an example, the neutralinos and charginos can be “bino-like”, “wino-like”, or “higgsino-like” depending on the mixing. The exact mixing depends on the

specific breaking mechanism considered, which is discussed further in the next section.

A discrete symmetry called R-parity is typically postulated in SUSY models. R-parity is a multiplicative quantum number defined as $(-1)^{3B+L+2s}$, where B is baryon number, L is lepton number, while s is the particle spin. All SM particles, such as leptons and quarks, including the Higgs boson, have $R = +1$ (i.e. even-parity), while all supersymmetric particles, such as sleptons, squarks, gauginos, and higgsinos, have $R = -1$ (i.e. odd-parity). R-parity conserving supersymmetric models contain a natural dark matter candidate, referred to as the lightest supersymmetric particle (LSP). The LSP is, in most cases, the lightest neutralino $\tilde{\chi}_1^0$. We only investigate R-parity conserving MSSM models.

2.4.1 Gauge-mediated supersymmetry breaking

Supersymmetry in combination with EWSB requires that superpartners have an identical mass to their SM counterparts, which is in direct conflict with experimental evidence. No superpartners of the SM particles with equal mass have been observed. Therefore, if SUSY does exist it must be broken in such a way that the masses of the SUSY particles are heavy enough to evade current experimental limits. Soft SUSY breaking is introduced by adding mass terms to the Lagrangian that involve only the SUSY particles and, as a result, explicitly break the symmetry. There are various models that can generate the mass terms that arise from the soft SUSY breaking, which occurs at some energy scale far above the EWSB scale.

In the MSSM, a form of spontaneous symmetry breaking of SUSY, similar to the Higgs mechanism in the SM, is not possible since any field that has a non-zero vacuum expectation value to break SUSY will unintentionally spoil the SM gauge symmetry [24]. In order to avoid this problem and have a viable mechanism for soft SUSY breaking it is common to introduce a hidden SUSY sector, containing additional fields that lead to the breaking of supersymmetry, which is secluded from the visible SUSY sector where the MSSM resides. Consequently, supersymmetry is broken at some high energy in the hidden sector and then

mediated to the visible sector through the exchange of weakly interacting “messenger” fields.

Gauge-mediated supersymmetry breaking (GSMB) is one possible method of communicating supersymmetry breaking in the hidden sector to the MSSM sector through massive messenger fields that are charged under the SM gauge group. These messenger fields induce a gaugino mass at one loop level and then this is transmitted on to the scalar superpartners at two-loop level [26].

2.4.2 Simplified model spectrum scenarios

The simplified model spectrum (SMS) is an effective Lagrangian description of beyond SM interactions, in which a small number of SUSY particles are involved [27, 28, 29]. Under the SMS assumption all the properties of BSM physics can be reduced to its mass spectrum, production cross sections, and decay branching fraction. The SMS scenarios allow for comparisons between topological signatures that are more sensitive to the final state and kinematic selections than the assumptions made by more physically motivated models.

Chapter 3

Experimental Apparatus

This chapter introduces the Large Hadron Collider (LHC) apparatus and the Compact Muon Solenoid (CMS) experiment. The components relevant for this analysis are briefly described. We performed the analysis using collision data collected with the CMS detector. A more detail description can be found in the provided references.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a two-ring superconducting synchrotron hadron collider, producing either proton-proton or heavy ion (Pb-Pb) collisions. The LHC has a circumference of 27 km and is situated about 100 m underneath the French-Swiss border. The LHC is designed to probe physics at the tera-scale by colliding protons, with a center-of-mass energy $\sqrt{s} = 14 \text{ TeV}$ at a frequency of 40 MHz, reaching a peak luminosity of $\mathcal{L} = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. A detailed account of the accelerator can be found in Reference [30]. The LHC did not reach the design specification during the data collection period but operated at a lower energy and luminosity. In the first half of Run I, the LHC operated at a center-of-mass energy of $\sqrt{s} = 7 \text{ TeV}$ with a peak instantaneous luminosity of $\mathcal{L} = 3.5 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. For the second half of Run I, the center-of-mass energy was increased to $\sqrt{s} = 8 \text{ TeV}$ with the highest instantaneous luminosity reaching $\mathcal{L} = 7.7 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. The proton particles are accelerated and brought to collisions at different interaction points. Four major experimental detectors are installed around the LHC ring aimed at measuring the results of collisions occurring at the impact point at the center of the detector. These four experiments are: the Compact Muon Solenoid (CMS), A Toroidal LHC Apparatus (ATLAS), A Large Ion

Collider Experiment (ALICE), and LHC beauty (LHCb). A schematic of the LHC complex and its four major experiments is shown in Figure 3.1. The CMS and ATLAS experiments are general purpose detectors with the goal of searching for the Higgs boson, improving SM measurements, investigating electroweak symmetry breaking processes, and looking at possible physics beyond the standard model by studying the product of collisions. The LHCb is optimized for decays of mesons (e.g. B-hadron) with the goal of performing precise measurements of CP violation, while ALICE, specialized in heavy ion collisions, is designed to focus on physics of strongly interacting particles including the so-called quark-gluon plasma, a state expected to exist at extremely high temperature and energy density.

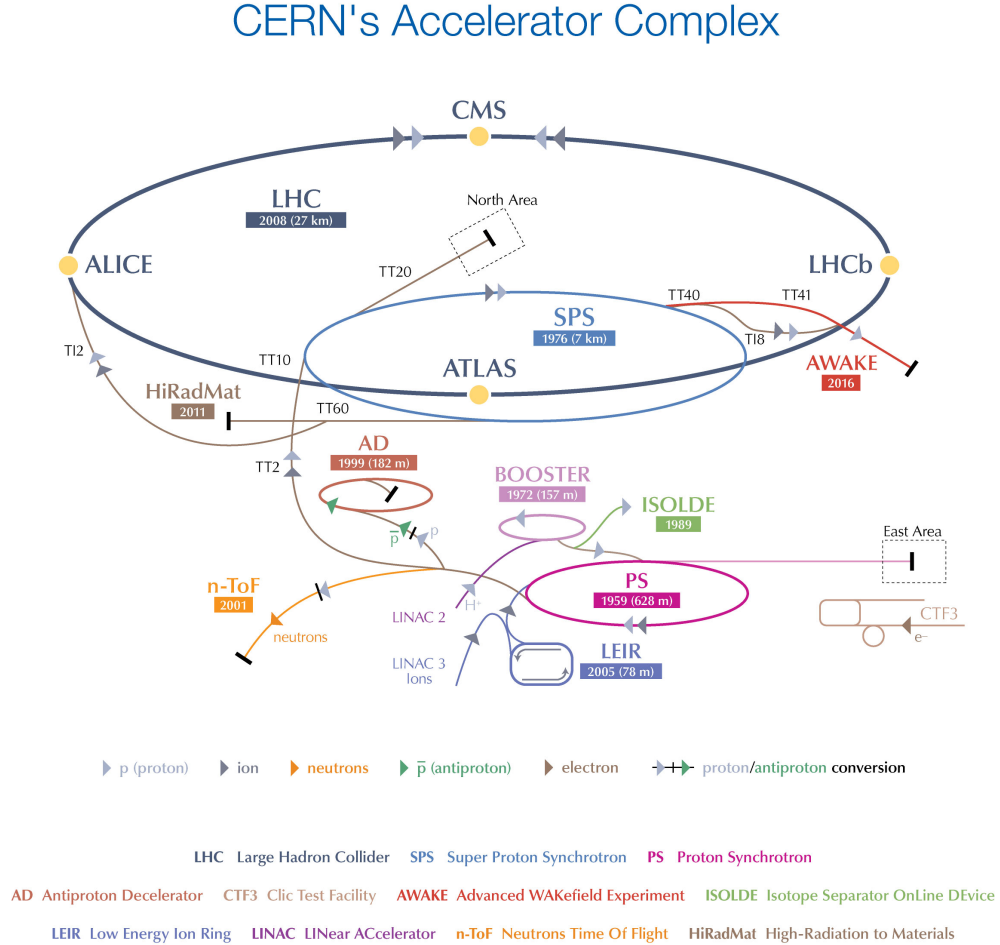


Figure 3.1: Schematic representation of the LHC's injection chain composed of multiple smaller accelerators [31].

3.1.1 LHC structure

The LHC consists of two rings with counter-rotating proton beams. These protons are produced from ionized hydrogen extracted from a tank of hydrogen gas, which are passed to the linear accelerator Linac 2. The proton beams originating from the Linac 2 are accelerated up to 50 MeV in energy. Afterwards, the beams are accelerated with the Proton Synchrotron Booster (PSB) to an energy of 1.4 GeV. Subsequently, the proton's energy is increased in the Proton Synchrotron (PS) to 25 GeV. After the PS, the protons are passed to the Super Proton Synchrotron (SPS), where they reach an energy of 400 GeV prior to injection into the main LHC ring. The beams are further accelerated to their final collision energies of about 7 TeV, provided by the radio frequency (RF) cavities operating at 400 MHz. The circulating beams consist of 2808 bunches of protons, which are brought to cross in intervals of 25 ns. The bunch length is about 53 mm with a transverse width of $15\text{ }\mu\text{m}$ and each bunch contains more than 10^{11} protons. An overview of the LHC proton injection chain is shown schematically in Figure 3.1. The high beam energy requires strong magnetic fields in order to guide the beams and keep the particles on track. The LHC ring contains 1232 superconducting Niobium Titanium (NbTi) dipole magnets, which are cooled down to about 1.9 K using superfluid Helium and create a magnetic field of 8.3 T. Quadrupole magnets squeeze the beams at the collision points. At the LHC, proton beams cross each other at a rate of 40 MHz, producing 20 proton-proton interactions on average. A summary of the principal LHC parameter values is given in Table 3.1.

3.2 The Compact Muon Solenoid detector

The Compact Muon Solenoid (CMS) is a multi-purpose detector designed to study pp collisions produced by the LHC. The CMS detector is situated about 100 m beneath the village of Cessy, France. The detector is 15 m in diameter, 22 m in length, weighs over 12,000 tons, and is constructed in concentric-layers of sub-detectors. An overview of the CMS detector schematic is shown in Figure 3.2.

Table 3.1: LHC design parameters for pp collisions in 2012 [32].

Parameters	Design value
Center-of-mass energy	8 TeV
Number of protons per bunch	1.1×10^{11}
Number of bunches	2808
Designed luminosity	$10^{34} \text{ cm}^{-2} \text{ s}^{-1}$
Luminosity duration	10 Hours
Bunch Length	53 mm
Beam radius at interaction point (IP)	$15 \mu\text{m}$
Time between collisions	25 ns
Bunch crossing rate	40 MHz
Circumference	27 km
Dipole field	8.3 Tesla

The tracking system is closest to the beam pipe, consisting of silicon pixel and strip detectors. The silicon tracking system is surrounded by the lead tungstate electromagnetic (ECAL) and brass/scintillator hadron (HCAL) calorimeters, all contained within a compact superconducting solenoid magnet that provides a 3.8 T magnetic field along the beam-line. A muon system, which is embedded in the return yoke, composed of steel lies outside the magnetic solenoid. The muon system provides a means of identifying muons and an additional momentum measurement. The following sections provide a brief description of the various sub-detectors components, though a more detail description is found in References [33, 34, 35].

3.2.1 Detector coordinates

The CMS collaboration uses by convention a right-handed coordinate system, with the origin at the center of the detector at the nominal pp interaction point. The x -axis points radially inward towards the center of the LHC ring and the y -axis points vertically up. The z -axis points along the counter-clockwise positive beam direction. The CMS detector has cylindrical symmetry around the beam-line. The azimuthal angle ϕ is measured in the xy -plane, while the polar angle θ is measured relative to the positive z -axis. The radial distance r is measured from the z -axis (i.e. beam axis). The ϕ angle is a convenient

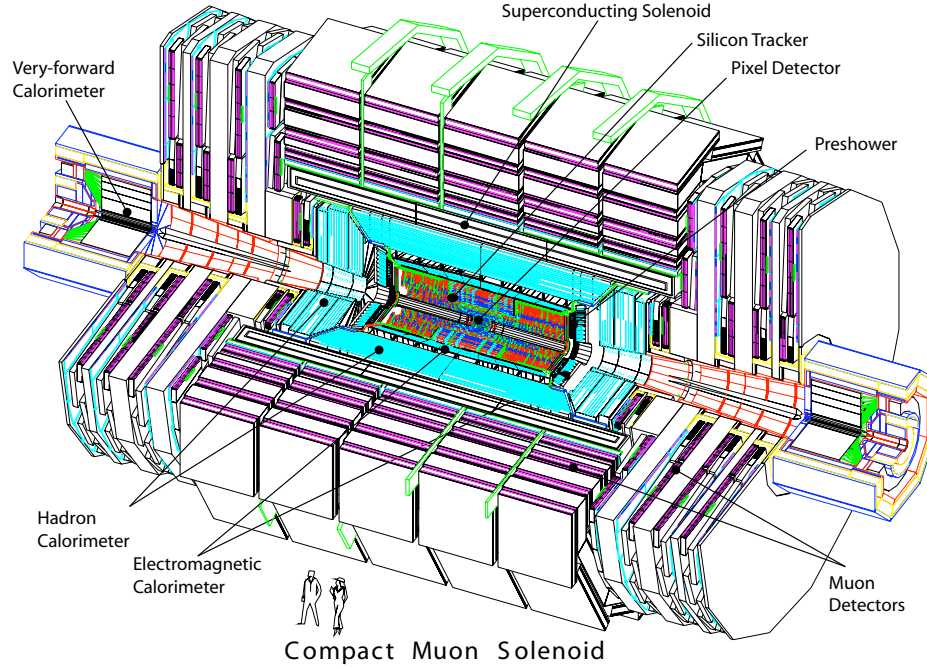


Figure 3.2: A schematic view of the CMS detector [35].

choice for scattering processes, assuming azimuthal symmetry, while θ is not invariant under Lorentz boosts along the z -direction. Instead, the rapidity y is used, since the difference between the rapidities of two particles is Lorentz invariant, and is defined by,

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \quad (3.1)$$

where E and p_z are the energy and momentum component of a particle in the z -direction, respectively. The rapidity is hard to measure for highly energetic (relativistic) particles, since it is necessary to calculate both the energy and total moment of the particle along the z -direction. However, at relativistic energies (or large mass limit) a good numerical approximation of a particle's rapidity is their pseudorapidity (η), which is much easier and quicker to estimate compared to the rapidity. The pseudorapidity, η , quantity is Lorentz

invariant under the longitudinal boosts, and is defined as,

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right). \quad (3.2)$$

It follows from the definition that η is zero for $\theta = 0$ and asymptotically approaches infinity at $\theta = \pi/2$, along the beam-line direction. In hadron colliders, such as the LHC, the composite nature of proton collisions results in interactions that rarely have the center of the mass frame coincide with the CMS detector rest frame, and for highly relativistic particle collisions, $y \simeq \eta$, means that η is the more convenient property to measure.

3.2.2 Superconducting Solenoid Magnet

A central feature of the CMS detector is the superconducting solenoid that encloses the inner tracker system. It has a length of 12.5 m and inner diameter of 6.3 m, which weighs roughly 220 tons. The cylindrical magnet consists of four layers of Niobium-Titanium (NbTi) superconductors, with a current of about 18 kA generating a magnetic field of 3.8 T, cooled by liquid Helium down to a temperature of 4.6 K. The strong magnetic field provided by the magnet ensures a high momentum resolution measurement ($\Delta p/p \approx 10\%$) for highly energetic charged particles ($p = 1$ TeV). The superconducting magnetic coil is installed behind the electromagnetic and hadron calorimeter since any additional material in front of calorimeter will affect the resolution energy measurement and particle identification. The magnetic flux is returned by the 1.5 m thick heavy iron return yoke that is fully saturated.

3.2.3 Tracking system

The purpose of the CMS tracking system [36] is to precisely measure the trajectories of charge particles emitted from collisions, which traverse the tracker with high efficiency. The inner silicon track detector surrounds the collision point and has a length of 5.8 m and a diameter of 2.5 m. The tracking system is composed of silicon based sensors that detect the

ionization produced by the passing of charged particles through the material. The trajectory of charged particles are reconstructed from the linked hits detected in the various layers of the detector. The transverse momentum measurement of the charged particles is performed using the bending radius of the trajectories, due to the magnetic field. In addition, the tracker system is used in the reconstruction of vertices, which is discussed in Section 5.3. The inner most component of the silicon tracker system consists of the inner pixel detector and subsequently the silicon strip tracker.

Based on particle flux and occupancy the tracker consists of three regions. The pixel detector installed at $r \approx 10$ cm, closest to the interaction point, with a high particle flux of about 10^7 s^{-1} and corresponding to an occupancy of roughly 10^{-4} per bunch crossing. Next is the intermediate region, $20 < r < 55$ cm, where there is lower particle flux which allows the use of silicon micro-strip detectors with a smaller cell size of about $10 \text{ cm} \times 80 \mu\text{m}$ and an average occupancy of $2 - 3\%$. Lastly, the outer most region, $r > 55$ cm, of the inner tracker system, where the particle flux drops significantly, uses larger pitch silicon micro-strips with a cell size of about $25 \text{ cm} \times 180 \mu\text{m}$, which leads to an average occupancy of about 1% . An overall schematic cross section of the tracker system is shown in Figure 3.3. The tracker was designed such as to minimize its material budget with the purpose of reducing multiple scatterings, photon conversions, bremsstrahlung, and nuclear interaction of the particles traversing the material.

Pixel detector

The CMS pixel detector has higher granularity than the strip detector and guarantees a three dimensional position resolution of about $15 - 20 \mu\text{m}$ [38, 35, 34]. The detector is specialized in the identification of primary and displaced vertices, which is important in identifying b-quark jets. The pixel detector is part of the tracker system closest to the beam-line (interaction point), consisting of 1440 pixel modules with a total of 66 million pixels distributed over three barrel layers at radii of 4.4 cm, 7.3 cm, and 10.2 cm, each of

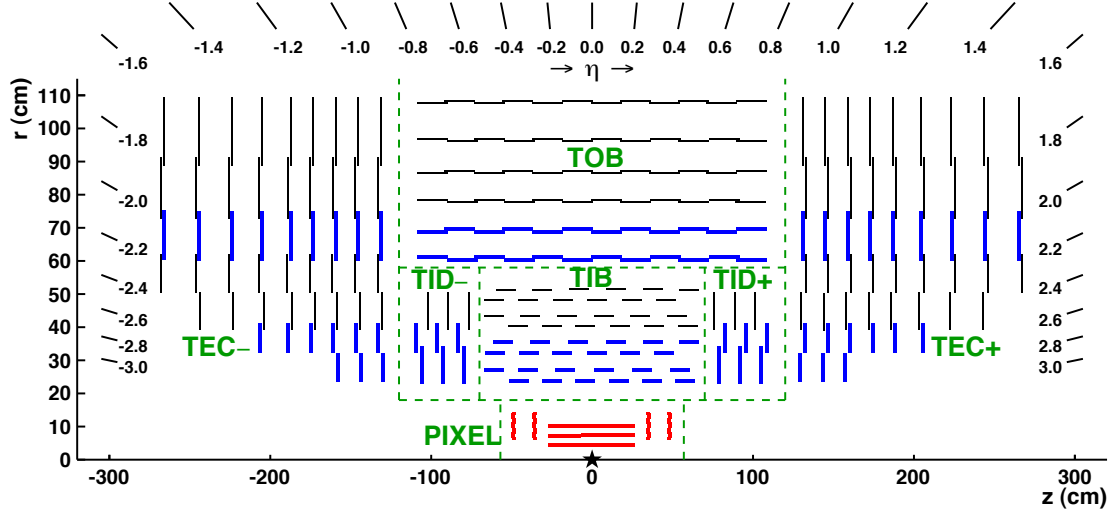


Figure 3.3: Schematic cross section through the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules which deliver stereo hits [37]. Disposition of the different detectors in the silicon tracker. PIXEL (red) refers to silicon pixel detectors while TIB, TID, TOB and TEC (blue) all refer to silicon strip detectors [37].

length 53 cm and two forward disks in the endcap. The endcap disks are on each side, located at $|z| = 34.5$ cm and $|z| = 46.5$ cm and extended in the radial direction from 6 to 15 m, as shown in Figure 3.4. The pixel detector provides the necessary precision of the impact parameter resolution essential for the reconstruction of secondary vertices and covers a range of $|\eta| < 2.5$. The size of the individual pixels are approximately $100 \times 150 \mu\text{m}^2$ in the $r - \phi$ and z -direction, respectively.

Silicon strip detector

The silicon strip tracker lies beyond the pixel detector. At a larger radial distance the occupancy reduces sufficiently that silicon strip detectors with large surfaces can be used. The silicon strip tracker consists of the four sub-systems, specifically the tracker inner barrel (TIB), tracker outer barrel (TOB), tracker inner disks (TID), and tracker endcaps (TEC) [38, 35, 34]. The TIB is composed of 4 layers extended up to $|z| = 65$ cm relying on two silicon sensors with a strip pitch between $80 - 120 \mu\text{m}$. The silicon strips are, in general, parallel to the z -axis, providing measurements in the r - and z -direction. The inner layers

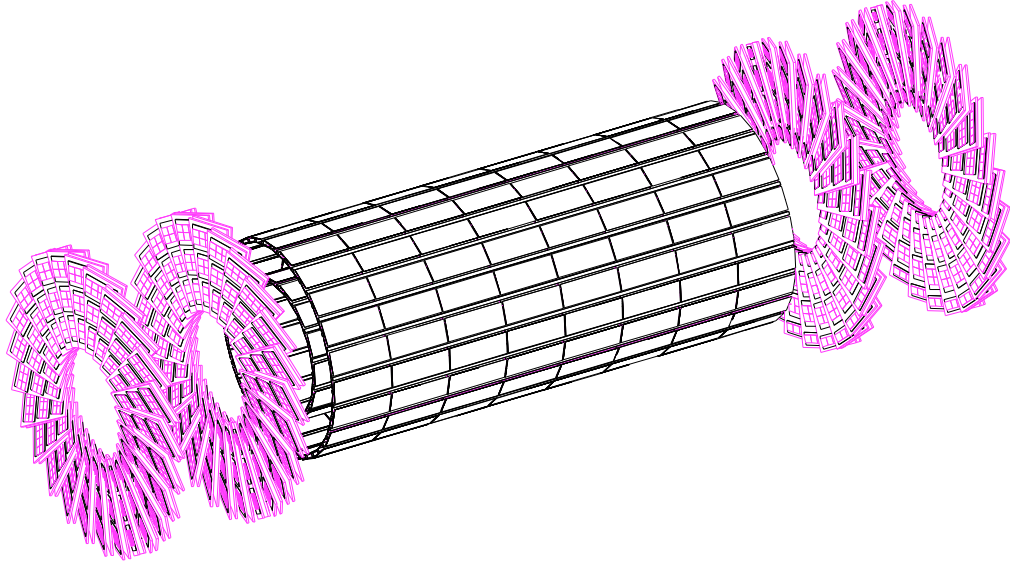


Figure 3.4: Layout of the pixel detectors in the CMS tracker showing the pixel barrel (TPB) and pixel endcaps (TPE), in black and pink, respectively [38].

contain the *stereo* modules, which consists of two strip modules with an angle of 100 mrad (i.e. milliradians) between the two strips, providing a measurement in the ϕ -direction. The TOB is comprised of six layers and extends to $|z| = 118$ cm, which uses strip pitches between $120 - 180 \mu\text{m}$. The TID consists of three disks, which fill the gap between the TIB and TEC. The TEC consists of nine disks in the region $120 < |z| < 280$ cm. The TID uses trip pitch between $100 \mu\text{m}$ to $141 \mu\text{m}$.

3.2.4 Calorimetry system

The tracker system measures precisely the momentum of electrically charged particles, but in the case of neutral particles (e.g. photons) that do not produce hits in the tracker it provides no information for them. Therefore, a calorimetry system, which is designed such that incident particles deposit most of their energies as they pass through the calorimeters, measures the energy of both charged and neutral particles. The calorimeter system consists of two sub-detectors, specifically, the electromagnetic (ECAL) and hadron (HCAL) calorimeters. These calorimeters are installed within the magnetic coil. The ECAL is designed to

absorb light electromagnetic particles, such as electrons and photons, produced in pp collisions and to measure their energies. While the HCAL, with much thicker layers, aims to absorb hadrons, by nuclear interactions, and measures the energies of charged and neutral hadrons. The interaction length for hadrons are much larger compared to the radiation length for photons and electrons requiring certain thicknesses for the materials. Further details on both calorimeters are provided in the following sections.

Electromagnetic calorimeter

The CMS detector features a homogenous and hermetic electromagnetic calorimeter (ECAL) system [35, 34] with the purpose of measuring the energy of photons and electrons with a precise resolution, and any partial energy deposited by charged particles, such as charged hadrons and muons, which are not fully absorbed within the material. Combined with the tracker information photons and electrons can be distinguished, since the former does not leave behind track hits. The ECAL system, which is placed inside the solenoid, consists of a cylindrical barrel and two endcaps concentric with the silicon tracker. The main components that comprise the ECAL system are the scintillating crystal system, photodecators, and preshower detectors. Figure 3.5 shows the general layout of the ECAL system. The largest active volume is composed of lead tungstate crystals. There are 61200 lead tungstate (PbWO_4) scintillating crystals mounted in the ECAL barrel (EB), $|\eta| < 1.479$, that is closed at each end by ECAL endcaps (EE), $1.479 < |\eta| < 3.0$, each containing 7324 crystals. Lead tungstate has a high density of $\rho = 8.28 \text{ g/cm}^3$, a short radiation length of $X_0 = 0.89 \text{ cm}$, and a small Molière radius of $R_M = 2.2 \text{ cm}$, which permitted the construction of a compact calorimeter with fine granularity. Moreover, the ECAL crystals are radiation tolerant and, being fast scintillators, emit about 80% of the light in 25 ns which is the designed time between two bunch crossings.

The advantage of the small radiation length of lead tungstate is that it allows for moderate size ECAL crystals to be used. The EE crystals have a length of 23 cm, while the EB

crystals have a length of 22 cm, corresponding to radiation lengths of $25.8 X_0$ and $24.7 X_0$, respectively. The front surface cross section of the EE crystal is $2.2 \times 2.2 \text{ cm}^2$ (i.e. the square of the Molière radius), corresponding to 0.0174×0.0174 in the $\eta - \phi$ plane, while the rear surface cross section is $2.6 \times 2.6 \text{ cm}^2$. The ECAL crystals are arranged in a rectangular $x - y$ grid in the endcaps, such that the front surface cross section is $2.862 \times 2.862 \text{ cm}^2$, while the rear surface cross section is $3.0 \times 3.0 \text{ cm}^2$.

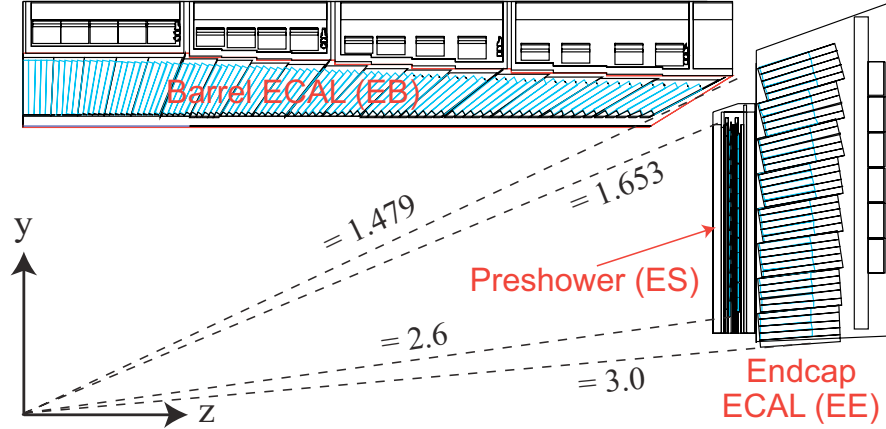


Figure 3.5: Geometric view of one quarter of the ECAL system showing the locations of the electromagnetic barrel (EB), electromagnetic endcaps (EB), and electromagnetic preshower (ES) [35].

The light emitted by the scintillating crystals, excited by ionizing radiation, is measured with the photodetectors, specifically by avalanche photodiodes (APD) and vacuum phototriodes (VPT) in the barrel and endcap regions, respectively. The APD are able to operate in the strong magnetic fields of the CMS detector, while the VPT are less sensitive due to the higher radiation in the endcaps regions. An electromagnetic preshower (ES) detector in the range of $1.653 < |\eta| < 2.6$ is installed in front of each endcap and consists of lead absorbers equipped with silicon strip sensors. The ES detector is a sampling calorimeter that enhances photon identification capabilities, improves measurements of electron and photon positions, helps discriminate electrons against minimum ionizing particles, assists in the identification of neutral pions, and provides additional background rejection in the

forward region. The EE and EB crystals and ES detectors are shown in Figure 3.6.

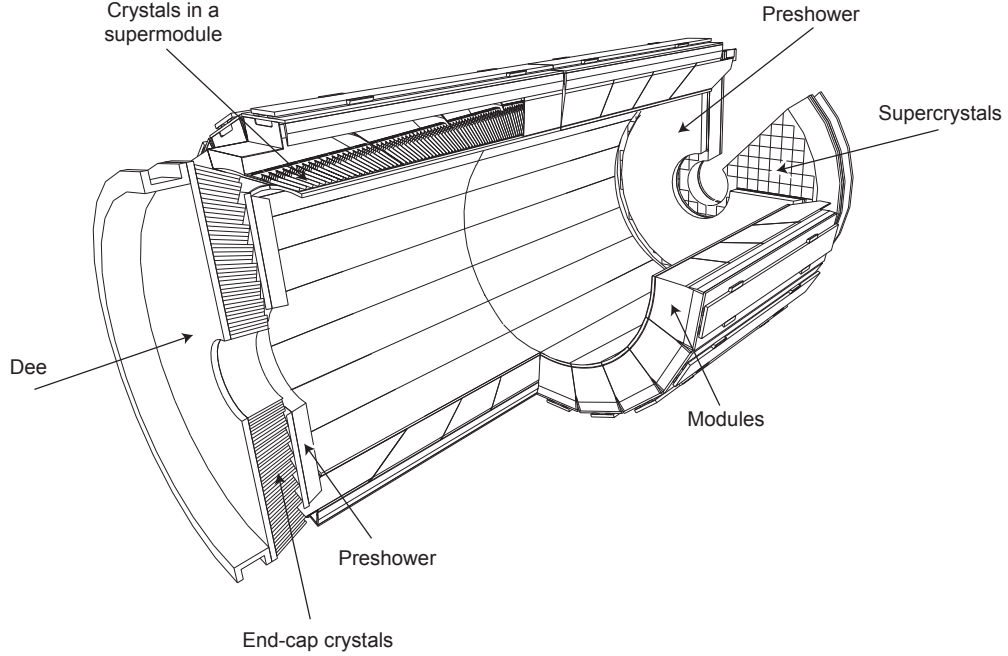


Figure 3.6: A transverse section of the ECAL system showing the arrangement of crystal modules, supermodules, preshower, and endcaps [34].

The energy resolution of the ECAL is fitted with a Gaussian and parameterized by the following expression,

$$\frac{\sigma_E}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C, \quad (3.3)$$

where the stochastic term is given by $S = 2.8\%$, the noise term by $N = 12\%$ and the constant term by $C = 0.3\%$ [39]. The stochastic term reflects the event-to-event fluctuations in the shower containment, photostatic effects, and fluctuations in the gain process. The noise term describes the intrinsic noise from ECAL electronics, digitization process, and pileup. The constant term includes contributions from the non-uniformity of the longitudinal light collection, energy leakage from the back of the calorimeter, and intercalibration errors.

Hadron calorimeter

The hadron calorimeter (HCAL) system is a hermetic sampling calorimeter made of alternating layers of plastic scintillators and brass absorbers [35, 34, 40]. The purpose of the HCAL system is to measure the energy of hadrons with enough accuracy as to allow for the determination of any missing transverse energy. The design of the CMS detector strongly constrains the size of the HCAL system since the calorimeter needs to be small enough to fit inside the volume of solenoid magnetic and at the same time have sufficient stopping power to contain hadronic showering. A flat “cartridge brass” alloy was chosen to satisfy this requirement, which has the added advantage that it is also non-magnetic and composed of relatively low-Z elements, such as zinc and copper, meaning that it will not significantly degrade the muon measurement. The sampling calorimeter consists of about 5 cm thick passive layers made of brass with plastic scintillator tiles interlayed that connect to wavelength shifting fibers. The HCAL detector is comprised of several components, specifically, the hadron barrel (HB), hadron endcap (HE), hadron outer (HO), and hadron forward (HF) calorimeters. The location of the different components of the HCAL system within the CMS detector are shown in Figure 3.7. The chemical and physical properties of the HB brass absorber consists of cartridge brass C26000, with a density of $\rho = 8.53 \text{ g/cm}^3$, radiation length $X_0 = 1.49 \text{ cm}$, and interaction length $\lambda_I = 16.42 \text{ cm}$ [34].

The HB covers a pseudorapidity region of $|\eta| < 1.4$ and has a radius of $1.77 - 2.95 \text{ m}$ that corresponds to the outer radius of the ECAL and the inner radius of the magnet, respectively [41]. It is made of 36 azimuthal wedges that are arranged into two groups of 18 identical 20° wedges in ϕ to form the two halves of the barrel. The wedges are constructed of 16 layers of absorbers made from either brass or steel followed by a plastic scintillating tile. The absorbers have a total thickness ranging from $5.82\lambda_I$ at $\eta = 0$ to $10.6\lambda_I$ at $|\eta| = 1.3$. The plastic scintillators are divided into 16 segments in η and 4 in ϕ , so-called HCAL “towers”, resulting in a segmentation of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$.

The HE extends the pseudorapidity range of the HCAL system, such that $1.3 < |\eta| < 3.0$,

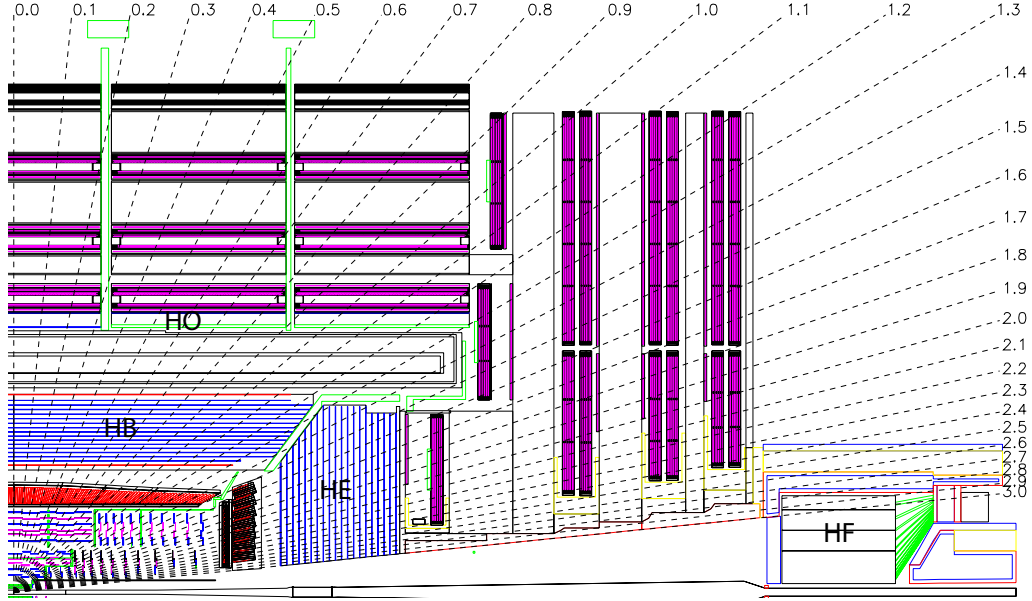


Figure 3.7: A longitudinal view of the CMS HCAL detector showing the locations of the hadron barrel (HB), hadron endcap (HE), hadron outer (HO), and hadron forward (HF) calorimeters [34].

and has a small overlap with the HB to ensure hermeticity [42]. Similar to the interlayered construction of the HB, the HE uses a 79 mm thick brass absorbers followed by plastic scintillators. The granularity of the calorimeters is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ for $|\eta| < 1.6$ and $\Delta\eta \times \Delta\phi \approx 0.17 \times 0.17$ for $|\eta| \geq 1.6$.

The amount of absorber material in the HB is limited by the amount of space between the ECAL and the superconducting solenoid magnet and, as a result, does not provide sufficient containment for hadronic showering. The HCAL system has the HO installed outside the solenoid, which serves as a “tail-catcher”, in order to improve the shower containment in the $|\eta| < 1.26$ region. In addition, it provides measurements for late showering particles that punch-through the magnet [43].

The HF is a cylindrical steel structure located at a distance of 11.2 m from the interaction point. The purpose of the HF is to measure particles in a region of $3.0 < |\eta| < 5.0$ and to increasing the hermeticity of the missing transverse energy measurement [44]. It uses

5 mm thick steel absorber layers with grooves containing quartz fibers. In this region of the detector very high radiation is expected over long running periods and for this reason the quartz fibers were chosen as the medium to collect the Cherenkov radiation emitted by these relativistic particles [44].

The hadronic energy resolution of the ECAL+HCAL combination is fitted with a Gaussian and parametrized with the following expression,

$$\frac{\sigma_E}{E} = \frac{S}{\sqrt{E}} \oplus C, \quad (3.4)$$

where S refers to the stochastic term and C is the constant term that accounts for the non-uniformity of the HCAL system and mis-calibrations. An estimation of the energy resolution is performed using charged pions of 20 – 300 GeV [35]. The HB has values of $S = 84.7\%$ and $C = 7.4\%$ [41], while the HE has values of $S = 153\%$ and $C = 6.3\%$ [42]. The stochastic term for the HB energy resolution remains the same with or without the HO, while the constant terms improves to 6.6% with the HO [43]. Similar results are obtained for the HF, where $S = 279.9\%$ and $C = 11.4\%$ [44]. The measurement of the HCAL energy resolution has also been performed with electrons and muons. For example, the HF has a value of $S = 197.9\%$ and $C = 8.6\%$ for its electromagnetic shower energy resolution [44].

3.2.5 Muon system

Muons are able to pass through the calorimeters without being absorbed. Therefore, the muon system [35, 34] is at the outer most part of the CMS detector. It lies passed the magnetic coil, embedded in the iron flux return yokes, which is used as as the absorber material, and is further subdivided in a barrel and two endcap regions. The system is designed to precisely measure the p_T of muons with $\sigma(p_T)/p_T < 10\%$ resolution over a large p_T range. Besides improvement of the inner tracking momentum measurement, the muon detector aims at identifying muons and provide information to the trigger system.

The muon system relies on three different gaseous ionization detectors, which are embedded through the iron return yoke in the barrel and endcap regions. They comprise of the drift tubes (DT) in the central region ($|\eta| < 1.2$), cathode strip chambers (CSC) in the forward region ($0.8 < |\eta| < 2.4$), and resistive plate chambers (RPC) cover the remaining regions ($|\eta| > 2.1$). A schematic cross section of the muon system is shown in Figure 3.8.

The DT consists of four gap stations, where the magnetic field is uniform, that are arranged in chambers labeled MB1 to MB4, and installed between the layers of the iron flux return yoke at radii of 4.0, 4.9, 5.9, and 7.0 m respectively, see in Figure 3.8. The DTs are used in the barrel region, where the strong magnetic field is uniform. The muon rate is relatively low in this region and the magnetic field is for the most part contained in the iron return yoke. Twelve chambers are integrated in the first three layers, where eight measure the $r - \phi$ coordinate and the other four determine the z -direction. Furthermore, the Level-1 DT muon trigger is subdivided into the DT local trigger and DT regional trigger, where hits in the DT muon detector are arranged in segments by the DT local trigger. Tracks segments consistent with muon trajectories originating from the interaction point are linked by the DT track finder (DTTF), which assigns a momentum, a charge, and a quality index. The DT sorts and provides the best four muons in the central detector to the global muon trigger (GMT) system.

The CSC detector is mounted perpendicular to the beam direction and is comprised of four stations. Arranged in six-layer modules, for each CSC layer the cathode strips run radially outward, allowing for precise measurements in the $r - \phi$ plane. The CSCs are multi-wire proportional chambers and used at the endcaps to handle the intense magnetic field and higher particle multiplicity.

The RPC are parallel-plate gas chambers, serving as a complementary trigger system, which are added in the barrel and endcap ensuring fast triggering. The RPC are arranged in six barrel and four endcap stations with a total of 612 chambers. The RPC system is capable of identifying muons with high efficiency and provide a very fast triggering system.

The information provided by these detectors is analyzed by the L1 Muon Trigger System and the output delivered to the GMT.

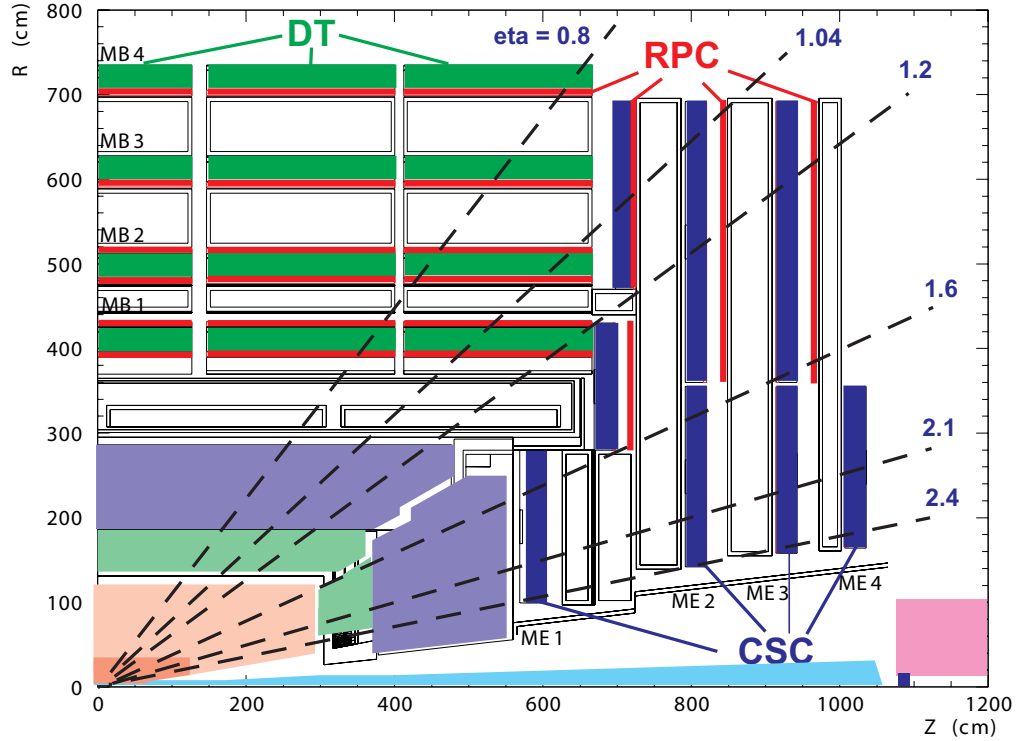


Figure 3.8: A layout of one quarter of the CMS muon system showing the location of the drift tube (DT), resistive plate chamber (RPC), and cathode strip chambers (CSC) systems [35].

3.3 Trigger system

At the LHC's nominal design luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ there will be roughly 10^9 collisions per second in the CMS detector, resulting in about twenty pp collisions. The amount of data produced during the collisions at the LHC exceeds the current computational capability for data processing and storage required to manage the high event rate of the CMS detector. Therefore, the rates must be significantly reduced, while retaining most of the potentially interesting events. It is necessary to reduce the event rate by a factor of 10^6 . In principle, the trigger is designed to select events likely coming from physics processes of interest to the CMS collaboration, out of the millions of uninteresting events.

In the CMS detector, the rate is reduced via a two-fold trigger system, known as the Level-1 (L1) and the High-Level Trigger (HLT) [45, 46]. The L1 trigger is hardware-based, consisting of custom-designed electronics, in order to provide a fast response. The purpose of the L1 trigger system is to reduce the data event rate of 40 MHz down to about 100 kHz, which is achieved by combining information from the calorimeter and muon systems to select only the events that are of interest. The L1 trigger system logic scheme is shown in Figure 3.9. The decision of whether to accept or discard the event needs to be made within $3.2\,\mu\text{s}$ (i.e. latency) for each bunch crossing and store the event information in so-called pipelines. The decision to accept or reject the event is based on the presence of Trigger Primitive Objects (TPO), such as electrons, muons, photons, and jets above a certain p_{T} threshold. The L1 hardware-based trigger uses p_{T} thresholds and selects on global event variables, such as the sum of E_{T} and $E_{\text{T}}^{\text{miss}}$ to reduce the event rates to a manageable level. The events rejected by the L1 trigger are discarded, while those that pass the L1 triggers are passed to the HLT for further processing.

The HLT consists of a software-based trigger system, implemented in a filter farm with roughly one thousand commercial computers, which perform a partial reconstruction of the event and further reduce the rate to about 300 Hz. The HLT event selection is based on predefined sets of trigger paths. These events containing physics objects, such as electrons, muons, photons, and jets above a certain threshold energy are stored for later reconstruction and analysis. The frequency of firing such triggers increases as the p_{T} threshold decreases. Therefore, pre-scale factors are applied to limit these rates. For example, a pre-scale of 1000 means the trigger will fire at every 1000th event. Information provided by the trigger system is used to store subsets of the entire datasets, categorized based on the different physics objects and kinematics.

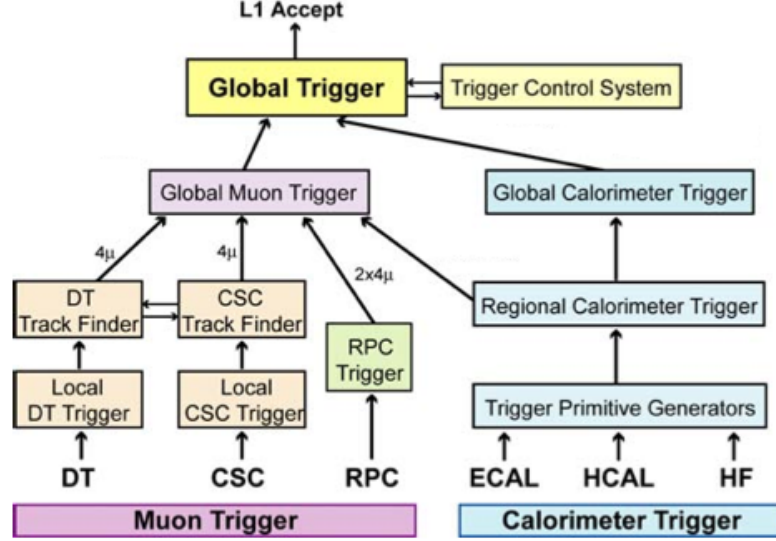


Figure 3.9: Level-1 trigger decision flow of the CMS detector before data is being transferred to the DAQ [34].

3.4 Luminosity measurement

In addition to the beam energy another important quantity is the luminosity. Physics processes are characterized by either their production or decay cross section, therefore, an accurate measurement of the luminosity is necessary to properly normalize simulated events. Recorded data are in separate runs with finer granularity stored in so-called “luminosity sections” (LS), corresponding to 23s time intervals. The instantaneous luminosity $\mathcal{L}$ is measured with the CMS pixel detector and cross checked with the HF calorimeter measurement [47]. The pixel technique for measuring the luminosity has a smaller dependence on beam conditions, such as multiple interactions. The luminosity from proton collisions is measured based on the average number of pixel hits in zero-biased triggered events, which require only occurrences of bunch crossing.

The rate of interaction generated for a given process is given by,

$$\frac{dN}{dt} = L\sigma_{\text{prod}} \quad (3.5)$$

The luminosity L can be related to the rate of a process with,

$$N = \sigma_{\text{prod}} \int \mathcal{L} dt, \quad (3.6)$$

where σ_{prod} is a given production cross section for a specific process. The integrated luminosity L is given by the time integral of $\mathcal{L}$ over the full run time of the collider. In order to have an enormous number of events in a given time period a very high luminosity is needed, since the luminosity measurement is related with the production rate of the process.

The luminosity depends on several beam parameters, and is given by,

$$\mathcal{L} = \frac{\gamma f_{\text{rev}} N_p^2 n_b^2}{4\pi \epsilon_n \beta^*} F, \quad (3.7)$$

where γ corresponds to the Lorentz factor, f_{rev} is the beam revolution frequency, N_p is the number of protons per bunch, n_b is the number of bunches, ϵ_n is the normalized transverse emittance, β^* is the beta function at the interaction point. The emittance factor ϵ_n characterizes the focusing of the area of the beam in position-momentum phase space and is constant along the beam axis. The geometric reduction factor F is introduced as result of the crossing angle of the beams at the interaction point.

The luminosity measurement is determined from the number of pixel clusters per event for a given dataset luminosity section. The sum of all the luminosity sections gives the total integrated luminosity L with an associated uncertainty of about 2.5%. The delivered (blue) and recorded (yellow) integrated luminosity as a function of time by the LHC and CMS, respectively, is shown in Figure 3.10.

The processes we are interested in are rare, therefore, we need high intensity to insure a measurable rate to observe these events. The luminosity is high enough to provide sufficient rate of hard interactions for interesting physics to occur. Additionally, the high beam intensity result in multiple proton-proton interaction per bunch crossing, referred to as “pileup”. The experiments must find ways to identify and account for these additional interactions.

The high bunch crossing rate leads to large amount of data such that experiments design analysis to select out the most interesting process.

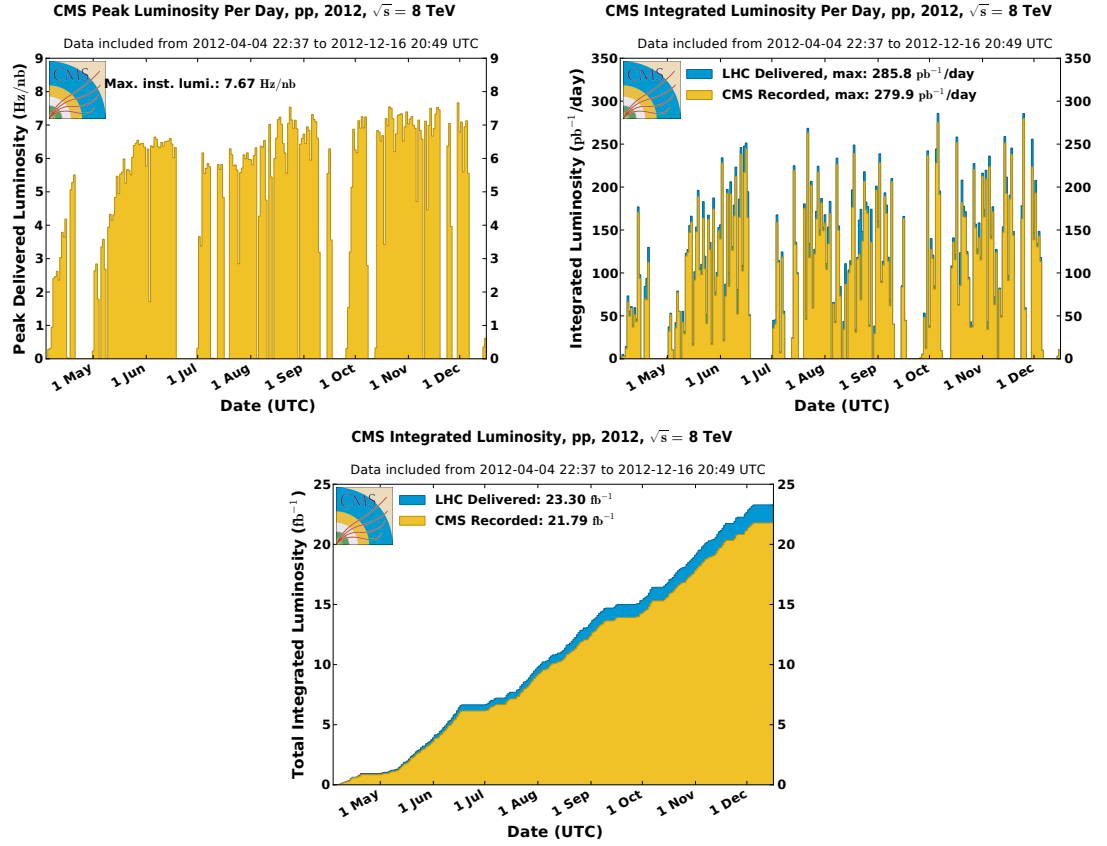


Figure 3.10: Peak luminosity (top left), integrated luminosity (top right), and total integrated luminosity (bottom) per day, delivered (blue) to and recorded by the CMS detector (yellow), for pp collisions at 8 TeV center-of-mass energy in 2012 [48].

Chapter 4

Collision Data and Simulations

In this chapter, we give an overview of the collision datasets used in the analysis, as well as the simulated background and signal samples.

4.1 Collision data

In principle, the purpose of this analysis is to compare the measurements obtained from the CMS detector to the expectation given by the theoretical prediction. The expected backgrounds from SM processes are determined using data and Monte Carlo simulations. The analysis is based on a total integrated luminosity of 19.5 fb^{-1} of proton-proton collisions collected by the CMS detector. The LHC experiment ran at a center-of-mass energy of $\sqrt{s} = 8 \text{ TeV}$. The data samples consists of only *good quality* data recorded by the CMS collaboration. The Data Quality Monitoring (DQM) system established by CMS is responsible for producing quality certification during event reconstruction. The data is defined as good quality, if all sub-detectors, trigger and physics objects, such as tracks, jets, and leptons, pass strict quality selection criteria. The following primary datasets that are used to perform the analysis include **MuEG**, **DoubleMu**, and **DoubleElectron**. The CMS collision events are stored into these and other various datasets categorized based on the corresponding set of triggers fired, as described in Section 3.2.3. Double-lepton triggers require the presence of a charged electron or muon above a certain p_T threshold. The **MuEG** dataset corresponds to the collections of events satisfying trigger criteria requiring at least one muon candidate, and either an electron or photon candidate. Events which fire a trigger that requires at least two muon candidates are stored in the **DoubleMu** dataset. Lastly, the

DoubleElectron dataset is a collection of event satisfying a trigger which requires at least two electron candidates.

There are two categories for data reconstruction, namely “prompt reconstruction”, for events reconstructed immediately after being recorded, and “re-reconstructed data”, for data that have been reprocessed due to, for example, updates in detector alignment conditions. The collision data is separated into different runs and luminosity sections, such that “good” runs and luminosity blocks are selected depending on whether they pass the quality criteria consistent with DQM and physics validation. Accordingly, only data certified as “good” is used for analysis. Care must be taken when using all the dilepton datasets in order to avoid double counting events that appear in multiple datasets. Information on the different run ranges and their respective integrated luminosities for the primary datasets used in the analysis are shown in Table 4.1.

Table 4.1: Dataset names from the pp collisions.

Primary dataset name	Run range	Luminosity $\mathcal{L}$ (fb ⁻¹)
/MuEG/Run2012A-recover-06Aug2012-v1	190782 – 190949	0.082
/MuEG/Run2012A-13Jul2012-v1	190456 – 193621	0.808
/MuEG/Run2012B-13Jul2012-v1	193834 – 196531	4.428
/MuEG/Run2012C-24Aug2012-v1	198049 – 198522	0.495
/MuEG/Run2012C-PromptReco-v2	98941 – 203742	6.401
/MuEG/Run2012D-PromptReco-v1	203777 – 209151	7.273
/DoubleMu/Run2012A-recover-06Aug2012-v1	190782 – 190949	0.082
/DoubleMu/Run2012A-13Jul2012-v1	190456 – 193621	0.808
/DoubleMu/Run2012B-13Jul2012-v4	193834 – 196531	4.428
/DoubleMu/Run2012C-24Aug2012-v1	198049 – 198522	0.495
/DoubleMu/Run2012C-PromptReco-v2	98941 – 203742	6.401
/DoubleMu/Run2012D-PromptReco-v1	203777 – 209151	7.273
/DoubleElectron/Run2012A-recover-06Aug2012-v1	190782 – 190949	0.082
/DoubleElectron/Run2012A-13Jul2012-v1	190456 – 193621	0.808
/DoubleElectron/Run2012B-13Jul2012-v1	193834 – 196531	4.428
/DoubleElectron/Run2012C-24Aug2012-v1	198049 – 198522	0.495
/DoubleElectron/Run2012C-PromptReco-v2	98941 – 203742	6.401
/DoubleElectron/Run2012D-PromptReco-v1	203777 – 209151	7.273

We use an unbiased H_T triggered dataset to perform trigger efficiency studies, as discussed in Section 6.2, where H_T is a measure of the hadronic activity in an event. Besides

the double lepton dataset we use the `SingleMu` dataset for the identification and isolation efficiency measurements, as discussed in Section 6.3. We select events from each dataset that satisfy the trigger paths listed in Appendix B, where the collected events are reconstructed following the methods detailed in Chapter 5. For example, following naming conventions established by the CMS collaboration, the trigger path `HLT_El17-El8` selects events with at least two electrons and applies a p_T threshold of 17 GeV on the first electron and 8 GeV on the second.

4.2 Simulated samples

In high energy experimental particle physics, simulations are used to model collisions, to predict SM background processes, and to search for new physics. We describe in the following sections the general principles, as well as the steps, involved in the production of simulated events. The procedure of generating events begins with the hard scattering process, which factorizes into two parts, namely, the determination of which particles can initiate the process defined by parton distribution functions and the matrix element. Further details follow in subsequent sections.

4.2.1 Monte Carlo simulations

The Monte Carlo (MC) method is a numerical integration approach based on pseudorandom number generators responsible for the production, interaction, and decay of particles. Approximation methods are utilized to calculate event rates and event kinematics for the large number of particles originating from hard collision processes since they cannot be determined analytically. The main MC generators involved in the calculation of background and signal samples are the `MADGRAPH5 v1.3.30` [49, 50], `POWHEG` [51], and `PYTHIA6 v4.20` [52, 53] programs.

Parton distribution functions and matrix elements

The parton distribution function (PDF) of a proton gives the probability of finding a parton of a particular flavor, either a quark or gluon, carrying a certain fraction of the longitudinal momentum of the proton at a specified energy scale of the hard interaction. Cross sections are evaluated by convoluting the parton level cross section with the PDFs. The PDFs are determined by a fit to Deep Inelastic Scattering (DIS) data, since QCD does not predict the parton content of the proton, from various experimental observables in the processes. The general procedure for determining PDFs starts with the parameterization of the non-perturbative PDFs at low energy scales, either by fitting to various sets of experimental data or by making a simplified assumption of their analytical form. For the generated simulation samples, the CTEQ6L1 leading order PDFs are used [54, 55].

The parton level cross sections are determined using the matrix element (ME) of the hard process, also referred to as the scattering matrix, which relates the initial state and the final state of a physical system undergoing a scattering process. Incoming particles are given random momentum based on the PDFs of the proton while the outgoing particles are given random momentum based on the available kinematic distribution of the phase space.

Parton showering, fragmentation, and hadronization

Part of a scattering process involves the cascade of radiation produced via QCD processes, referred to as parton showering, and is simulated with event generators such as PYTHIA. The parton shower models the evolution of partons until the strong coupling constant reaches the order of $\alpha_S \approx 1$ using the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) equations. Partons also undergo fragmentation, where there is a probability that a parton fragments into a hadron carrying a fraction of the original parton’s momentum. After the parton showering comes the non-perturbative process of hadronization. In this phase, the partons are converted into colorless hadrons. The hadronization process is described in PYTHIA using the Lund string model, where the color potential between two partons increases linearly with

distance. When the energy within the string field become energetic enough to create quark–anti-quark pairs the string is broken into two separate color singlet states. The process is repeated multiple times until in the end the color-connected partons form on-shell hadrons.

The underlying event

Additional partons, from the colliding protons besides the two undergoing the primary hard scattering process, will likewise hadronize and produce particles. The underlying event (UE) describes the effects due to these partons interacting with one another. Multiple parton interaction, pileup, beam-beam remnants, including initial state (ISR) and finale state radiation (FSR) constitute the UE. In general, the UE is difficult to calculate since processes are soft and non-perturbative. Therefore, models are used in event generators, which rely on several “tuning” parameters, to simulate these events. The UE components are adjusted in these models to better match data. For this analysis, simulated samples are produced with the tune Z2* of PYTHIA to model the underlying event in the CMS detector.

Jet matching procedure

The parton shower procedure is only valid in the limit of soft and collinear emissions. For the hard initial and final state QCD radiation with large emission angles this description breaks down and perturbative matrix element calculations have to be used. The ME method, however, diverges in the limit of soft and collinear emission. The perturbative approach is no longer valid in the regime where the strong coupling constant reaches $\alpha_S \approx 1$. Therefore, simulated samples with additional partons at the matrix element level are subject to a parton-jets matching procedure, in order to avoid the overlapping of the phase-space descriptions given by MADGRAPH and the parton showering produced by PYTHIA. We require a minimum distance in the phase space between partons at the matrix-element level, specified by the parameter xQ_{cut} , and a maximum distance between a jet and a parton to be matched with each other, specified by the matching scale parameter Q_{cut} , to avoid

this overlapping between the perturbative and non-perturbative regimes. The jet matching procedure is applied to simulated samples using the MLM prescription, which uses an event rejection scheme as opposed to an event reweighting method.

Detector simulation

The CMS detector response, simulated with the GEANT4 program [56], proceeds the matrix element calculation, parton showering, and hadronization processes. The detector simulation describes the modeling of the particles passing through several detector components and interaction with the detector material, including the simulation of the signals arising from the CMS detector electronic readout, and the simulation of the trigger system. The fast simulation software developed by CMS [57] reduces the computational time over the full detector simulation and is based on a more simplified model of the detector. The simulated samples are stored in the same format as collision data along with additional generator level information, such as parent particle decay sequence.

Monte Carlo scaling

Simulation samples must to be scaled with respect to the integrated luminosity of the collision data in order to compare expected yields, therefore, simulated events are weighted by the factor,

$$w = \frac{\epsilon_{\text{filtering}} \cdot \sigma \cdot \mathcal{B} \cdot \mathcal{L}_{\text{int}}}{N_{\text{gen}}}, \quad (4.1)$$

where $\epsilon_{\text{filtering}}$ is a filtering efficiency, if there were any preselections performed on the sample, σ is the production cross section of the simulation, $\mathcal{B}$ corresponds to the branching fraction, if only certain decay modes were allowed in the sample, and $\mathcal{L}_{\text{int}}$ is the integrated luminosity of the collision dataset to which the simulation is compared, while N_{gen} is the number of generated events.

Pileup re-weighting

The number of additional interactions (pileup) in data is estimated by using the measured instantaneous luminosity for each bunch-crossing. The number of interactions from bunch-bunch collisions is obtained by multiplying the instantaneous luminosity by the total inelastic cross section. This distribution changes with increasing luminosity from the recorded data. While simulated events are generated with a default pileup scenario using PYTHIA and are superimposed on the hard collision. This results in the distribution of reconstructed primary vertices that differs from what is observed in data. In order to avoid potential bias of simulation, which is based on the number of expected pileup interactions in data, simulated processes must be adjusted to the pileup distribution found in data. The correction is performed by using a weighting procedure of the simulated process on an event-by-event basis, prior to applying any selection requirements. The pileup weight factor is obtained from the ratio of the number of events in data to the number of expected number of events for a given number of pileup interactions. All MC simulated events in this analysis are reweighted to represent the pileup conditions observed in data. Full simulations take into account both in-time and out-of-time pileup, where particles are originating from additional proton-proton interactions either from the same or neighboring bunch crossings, respectively.

4.2.2 Background samples

The major SM backgrounds for this analysis are from $t\bar{t}$, WZ, and ZZ productions with minor backgrounds arising from rare SM processes. We use simulated events to predict background contributions that can not be obtained with data-derived methods. Most samples have up to two or three additional partons at the matrix element level where jet matching has been applied. The simulation of SM events relies on the event generators MADGRAPH5 v1.3.30 [49, 50], POWHEG [51], or PYTHIA6 v4.20 [52, 53] with leading-order CTEQ6L1 [54, 55] parton distribution functions. Parton showering, hadronization, and the

decay of SM particles for all simulated events are processed using PYTHIA. The multiplicity of pileup interactions observed in data are taken into account to adjust the simulated events. The detector response is described by the full simulation GEANT4 package [56].

The cross sections are normalized to next-to-leading (NLO) or next-to-next-to-leading (NNLO) when available [58, 59, 60, 61, 62, 63, 64, 65]. These SM samples are used to compare collision data against estimates from background contributions. Tables 4.2 and 4.3 list all the simulated samples and the MC production information, respectively.

Table 4.2: List of simulation samples for background estimations. All samples are from the Summer12_DR53X-PU_S10_START53_V7A-v1 MC production campaign with the exception of the fully leptonic $t\bar{t}$, semi-leptonic $t\bar{t}$, and $t\bar{t}Z$ samples which are from the v2, ext-v1, and V7C-v1 versions, respectively.

Simulation names
/TTJets_FullLeptMGDecays_8TeV-madgraph
/TTJets_SemiLeptMGDecays_8TeV-madgraph
/TTWJets_8TeV-madgraph
/TTZJets_8TeV-madgraph_v2
/TBZToLL_4F_TuneZ2star_8TeV-madgraph-tauola
/ZZJetsTo4L_TuneZ2star_8TeV-madgraph-tauola
/WZJetsTo3LNu_TuneZ2_8TeV-madgraph-tauola
/GluGluToHTToTauTau_M-125_8TeV-powheg-pythia6
/GluGluToHTToWWTo2LAndTau2Nu_M-125_8TeV-powheg-pythia6
/GluGluToHTToZZTo4L_M-125_8TeV-powheg-pythia6
/VBF_HToTauTau_M-125_8TeV-powheg-pythia6
/VBF_HToWWTo2LAndTau2Nu_M-125_8TeV-powheg-pythia6
/VBF_HToZZTo4L_M-125_8TeV-powheg-pythia6
/WH_ZH_TTH_HToTauTau_M-125_8TeV-pythia6-tauola
/WH_ZH_TTH_HToWW_M-125_8TeV-pythia6

4.2.3 Signal samples

Monte Carlo simulations are also used in the production of signal events in order to determine the signal acceptance for the different search regions. The model parameters for the possible new physics scenarios, such as mass spectra and branching fractions, are generated with SuSpect v2.4.1 [66] according to the SUSY Les Houches Accord (slHA) [67] standards whenever appropriate. We use both the MADGRAPH and PYTHIA generators to perform

Table 4.3: Number of generated events and cross sections for background simulations given in Table 4.2.

SM background	Number of generated events	Cross section σ (pb)
$t\bar{t}$ + jets (fully leptonic decay)	12,119,013	23.08
$t\bar{t}$ + jets (semi-leptonic decay)	25,423,514	97.97
$t\bar{t}W$ + jets	196,046	0.232
$t\bar{t}Z$ + jets	209,677	0.208
$t\bar{b}Z(\rightarrow \ell^+\ell^-)$	148,504	0.0114
$ZZ(\rightarrow 4\ell)$ + jets	4,804,781	0.1769
$WZ(\rightarrow 3\ell\nu)$ + jets	2,016,678	1.0575
$gg \rightarrow H \rightarrow \tau\tau$	967,566	1.2466
$gg \rightarrow H \rightarrow WW^* \rightarrow 2\ell 2\nu$	299,975	0.4437
$gg \rightarrow H \rightarrow ZZ \rightarrow 4\ell$	995,117	0.0053
$WW/ZZ \rightarrow H \rightarrow \tau\tau$	1,000,000	0.0992
$WW/ZZ \rightarrow H \rightarrow WW^* \rightarrow 2\ell 2\nu$	299,687	0.0282
$WW/ZZ \rightarrow H \rightarrow ZZ^* \rightarrow 4\ell$	49,876	0.000423
$WH/ZH/t\bar{t}H(\rightarrow \tau\tau)$	200,000	0.0778
$WH/ZH/t\bar{t}H(\rightarrow WW^*)$	200,408	0.254

matrix element calculations and produce event level information in accordance with the Les Houches Event (LHE) [68] convention. Full simulation of new physics signals, for centrally produced samples by the CMS collaboration, are performed using GEANT4 for the detector response, while privately produced samples used the CMS fast simulation program [57] in place of GEANT4 to reduce computational resources.

Lastly, reconstruction of the simulated event is accomplished with the same software involved in processing collision data to facilitate the analysis and comparison between data and MC simulations. All signal processes are normalized to cross section calculations at NLO with PROSPINO2 or RESUMMINO. The PROSPINO2 software determines k -factors, which is the ratio of σ_{NLO} to σ_{LO} , to adjust leading order cross sections from MADGRAPH or PYTHIA to next-to-leading order cross sections. All signal cross sections have theoretical uncertainties associated with the PDFs, renormalization, and factorization scales used. Further information regarding the signal production of the various SUSY scenarios investigated in this analysis are provided in the following sections.

Production of the natural Higgsino NLSP scenario

The signal samples for the natural higgsino NLSP scenario are generated using MADGRAPH, where the SUSY decays are performed with the BRIDGE v2.24 program [69]. Parton showering, hadronization, and the decay of SM particles, are processed with PYTHIA. The model is generated in the limit of no mixing between higgsinos and electroweak gauginos, and as a consequence the light neutralinos and charginos become degenerate [70]. Therefore, the $\tilde{\chi}_1^0$ and $\tilde{\chi}_2^0$ higgsinos are assigned masses of 5 GeV below and above the mass of the $\tilde{\chi}_1^\pm$ higgsino, respectively, such that the mass splitting is representative of the proximity to this limit. For this reason, $\tilde{\chi}_1^0$ and $\tilde{\chi}_2^0$ are referred to as “higgsino-like”. The Goldstino component of the massless and non-interacting gravitino is assumed to be massless. The cross sections for both strong and electroweak production processes are calculated at leading-order (LO) with PYTHIA and afterwards scaled to next-to-leading-order (NLO) using k -factors from PROSPINO2. We assign a flat uncertainty of 20% on the theoretical cross sections to account for the uncertainties associated with the PDF and with the renormalization and factorization scales. Details about the underlying physical concepts of the model are found in Section 10.2.1.

Production of the wino NLSP with a Higgs boson scenario

The signal samples for the wino NLSP with a Higgs boson scenario are generated using MADGRAPH including up to two additional partons at the matrix element level. The decays of the SUSY particles are performed using the BRIDGE software. Parton showering, hadronization, and the decay of SM particles, are processed with PYTHIA. The signal model is generated in the limit of no mixing between winos, binos, and higgsinos, and as a consequence the light neutralinos and charginos become degenerate [70]. For this reason, $\tilde{\chi}_1^\pm$ is referred to as “wino-like” and $\tilde{\chi}_2^0$ as “bino-like”. The Goldstino component of the massless and non-interacting gravitino is assumed to have a mass of 1 GeV. The cross sections are calculated at NLO plus next-to-leading-logarithm (NLL) using RESUMMINO [71, 72, 73, 74,

75]. The theoretical uncertainties on the cross sections vary between 3% and 14%. Details about the underlying physical concepts of the model are found in Section 10.2.2.

Production of the slepton co-NLSP and stau-(N)NLSP scenario

The signal samples for the slepton co-NLSP and stau-(N)NLSP models are generated using PYTHIA. The cross sections for both models are calculated at LO with PYTHIA and afterwards scaled to NLO using k -factors from PROSPINO2 [76]. We assign a flat uncertainty of 30% on the theoretical cross sections for both models. Details about the underlying physical concepts of the models are found in Sections 10.2.3 and 10.2.4.

Production of the third-generation SMS T1tttt and T6ttWW scenarios

The signal samples for the T1tttt and T6ttWW scenarios are generated using MADGRAPH with parton showering, hadronization, and decay of SM particles, performed by PYTHIA. The cross sections for both models are calculated at NLO+NLL [77, 78, 79, 80, 81]. The theoretical uncertainties on the cross sections for the T1tttt scenario vary between 23% and 27%, while those for the T6ttWW scenario are assigned a flat 30% [73]. Details about the underlying physical concepts of the models are found in Sections 10.2.5 and 10.2.6.

Chapter 5

Event Reconstruction

Different types of particles leave signatures in various parts of the detector, corresponding to their underlying interaction with matter. The final-state particles reconstructed for this analysis include muons, electrons, hadronic τ -leptons, and jets. A detailed description of the reconstruction steps for the various physics objects is found in the following sections.

5.1 Particle-Flow event reconstruction

Events in the CMS detector are reconstructed with the particle-flow (PF) algorithm [82, 83, 84]. The resulting list of reconstructed stable particles, produced in proton-proton collisions, such as electrons, muons, photons, and charged and neutral hadrons, provides a global (i.e. full) event description by combining information from all CMS sub-detectors. The silicon tracker is used to reconstruct particle tracks with a high precision and low fake rate, even for low p_T particles. The ECAL, with its high granularity, allows it to distinguish photons from electrons. The HCAL has a much lower granularity in comparison to the ECAL and, as a result, charged and neutral hadrons end up in the same energy cluster. However, when the information from the pixel and strip tracker is combined with the energy deposit measured in the HCAL, neutral and charged hadrons can be distinguished.

The PF algorithm consists of first acquiring information from the different sub-detectors, specifically the tracks from the central tracker, clusters of energy deposits in the calorimeters, and track segments from the muon system, afterwards a linking algorithm combines all the information from the sub-detectors together into *PFBlocks*. Lastly, these blocks

are directed to the particle identification and reconstruction step, which produces the final list of particles in the event. An iterative track finding and cluster algorithm is applied to ensure that there is a high efficiency for the building blocks and a low misidentification rate. These PF candidates can later be used to construct jets and to measure the missing transverse energy of an event.

5.2 Track reconstruction

The CMS silicon tracker is used to detect and measure trajectories of charged particles produced in the pp collisions. Charged particles traversing the detector interact with the material of the tracking system and leave so-called “hits”. This hit information is then linked together in order to reconstruct the trajectory, referred to as a “track”, of the particle. The path of charged particles is curved due to the 3.8 T magnetic field, and this curvature is proportional to their p_T , hence the measurement of the tracks of the charged particles are essential for the precise measurement of their momentum.

In the CMS collaboration tracks are used for the determination of primary interaction vertices and in the reconstruction of many particles such as muons, electrons, τ -leptons, jets, charged and neutral hadrons. The reconstruction of tracks is performed with the combinatorial track finder (CTF) algorithm by using information from both the silicon strip tracker and pixel detector [36]. The CTF algorithm finds and reconstructs track candidates starting with pairs and triplets of hits in the tracker, assuming they originate from the beam spot or known vertices. These initial track estimates with associated uncertainties, start the “seed” propagation. The seed parameters are estimated at the interaction point (beam spot) using a helix fit to the hit points, and the seed tracks are propagated outward by the Kalman filter (KF) algorithm [85]. Depending on the seed parameter, the KF proceeds iteratively through the layers of the tracking system successively updating the track parameters. In principle, the KF algorithm searches for compatible tracker hits based on the predicted trajectory from the equations of motion in a constant magnetic field,

including energy losses. The KF algorithm relies on the information from the current state of the particle trajectory, the statistical noise from the detector, and the underlying process of the particle interaction. Hits associated to tracks after each iteration are removed from the tracker hits collection. This results in a smaller set of hits to be used in the proceeding iteration, thereby removing reconstructed tracks after each iteration that are more likely to be fake tracks.

The CTF iterative tracking algorithm procedure uses six iterations, which are as follows:

- Find tracks reconstructed with relatively high p_T and hence straight tracks. Any hits associated with tracks are removed from the track collection.
- Iteration 0: Designed to detect prompt tracks with $p_T > 0.8 \text{ GeV}$, which originate close to the interaction vertex, and contain at least three pixel hits.
- Iteration 1: Designed to recover prompt tracks with exactly two hits and slightly lower p_T in the pixel detector.
- Iteration 2: Looks for low- p_T tracks.
- Iteration 3 – 5: Designed to find non-tracks originating away from the primary interaction point and any tracks missed in the previous iteration [37].

Furthermore, each iteration of the CTF algorithm consists of an additional four steps. The first step is the seed generation which provides an initial track candidate with two or three hits. The next step is the track-finding. Afterwards a track-fitting is performed, which provides the best estimate for the track parameters based on a Kalman filter. Lastly, tracks not satisfying certain quality selection criteria, including χ^2 per degree of freedom of the fitted tracks (distance from primary vertex), are rejected, and those that pass are considered “high purity”.

5.3 Primary vertex reconstruction

Vertex identification plays a central role in event reconstruction since we are interested in particles originating from the hard collision of two protons. There are several challenges in vertex reconstruction, such as particle interactions in the tracker volume, effects due to multiple proton interactions within the same bunch crossing leading to increasing pileup at higher instantaneous luminosity and center-of-mass energy, including multiple overlapping events containing high track density. The vertex reconstruction [86] begins with the selection of prompt tracks if they pass quality requirements discussed in the previous section. These selected tracks are grouped using the deterministic annealing (DA) clustering algorithm [87].

Track reconstruction can be used to determine the location of vertices. Vertex reconstruction algorithms measure the coordinates and associated uncertainties for each vertex produced in the proton-proton collisions, including those arising from pileup interactions. Vertex fitting is used to determine the vertex position formed by a given set of tracks, whereby the goodness of a fit metric, such as the probability $\mathcal{P}(\chi^2)$, is applied to either accept or reject a vertex candidate, thereby providing the best vertex estimate. For each track a point of closest approach in the z -direction to the beam-line, denoted by z_i , is determined with an associated uncertainty σ_i . The χ^2 metric is used to measure the performance and is giving by,

$$\chi^2 = \sum_{ij} w_{ij} \cdot \frac{(z_i - z_j)^2}{\sigma_i^2}, \quad (5.1)$$

where w_{ij} is interpreted as a probability.

The DA algorithm, rather than determining directly the total minimized χ^2 , from the pair z_j and w_{ij} , finds the most likely distribution of w_{ij} given $\langle \chi^2 \rangle$ and proceeds to decrease it until a local minima is found [87]. Vertex candidates are defined by clustered tracks within $d_z \leq 1$ cm with respect to their nearest neighbor, where d_z corresponds to the longitudinal position of the extrapolated impact parameter (point of closet approach to the

beam-line). After tracks are assigned to various vertices, a vertex fit is performed with the *Adaptive Vertex Fitter* (AVF) algorithm [88]. The vertex candidates that have at least two tracks are fitted with the AVF, which is a re-weighted least-squares fit algorithm, which rather than rejecting outlier tracks it weights them down. These weights depend on the track compatibility, based on χ^2 , with a specific vertex. The tracks associated with the vertex will receive a weight close to unity, while those not consistent with a vertex are given a weight close to 0. These weighted tracks are fit to the vertex positioning. After every iteration these weights are adapted to the re-calculated vertex position. The number of degrees of freedom of the reconstructed vertex is defined as $n_{\text{dof}} = 2 \cdot \sum w_i - 3$, where w_i corresponds to the weight of the i^{th} track. This sum of weights effectively gives the total number of tracks accepted by the AVF algorithm.

Most events contain more than one vertex, therefore the reconstructed vertices are sorted based on the weighted scalar $\sum p_{\text{T}}^2$ of tracks found in the track cluster. The vertex with the largest sum associated with the interaction vertex is selected as the “primary vertex”. While vertices originating from the decay of long-lived particles are considered “secondary vertices”. Secondary vertices provide important information in identifying long-lived particles from either τ -leptons and heavy flavor hadrons. Charged particles arising from additional inelastic proton-proton collisions (i.e pileup) are rejected. Tracks which are considered to originate from the primary interaction are accepted if they satisfy quality criteria, such as the number of hits associated with tracks and the normalized χ^2 of the Kalman track fit.

5.4 Photon reconstruction

Photons and electrons are reconstructed from so-called “superclusters” (SC), which are energy deposits in the ECAL crystals [89, 90]. In the barrel region, the superclusters are comprised of 1×5 ECAL crystals (i.e. 1 in the η -direction and 5 in the ϕ -direction), while in the endcap region, the superclusters are formed from 5×5 ECAL crystals. The supercluster energy is measured by summing the energy contributions in the array of crystals for their

respective detector regions. At most 35 ECAL crystals can be used to form superclusters. The superclusters reconstruct almost the full energy emitted by photons when they convert by way of $\gamma \rightarrow e^+e^-$, and where the electrons are subsequently bent by the magnetic fields in the CMS detector producing bremsstrahlung radiation.

An observable called R_9 is used to help distinguish converted photons from unconverted photons. The quantity R_9 is defined as the ratio of the energy contained within the 3×3 array of crystals centered on the most energetic crystal in the supercluster to the total energy of the supercluster. Converted photon candidates, in the barrel region with $R_9 < 0.94$ and in the endcap region with $R_9 < 0.95$, are given an energy equal to the supercluster energy. While unconverted photon candidates, in the barrel region with $R_9 > 0.94$ and in the endcap region with $R_9 > 0.95$, are assigned the energy within the 5×5 ECAL crystals around the highest energy crystal. The endcap region has larger crystals in comparison to the barrel region, and as a result the R_9 variable has a larger threshold in the endcap region. For example, neutral pions have lower R_9 values compared to unconverted isolated photon. The showers from photons that convert before reaching the ECAL will have wider transverse profiles and, therefore, lower R_9 values than those of unconverted photons. The superclusters have several energy corrections applied to them to take into account the effects due to the interactions with the material in front of the ECAL and shower containment. The reconstructed photons are given a momentum based on the location of the reconstructed primary vertex. The photon isolation is obtained by summing the p_T of charged hadrons, neutral hadrons, and photons inside an isolation cone of size $\Delta R = 0.5$.

5.5 Lepton reconstruction

Leptons emerging from the collision point produce hits in the inner tracker, deposit energy in the calorimeter and muon systems, and from these signals the reconstruction and identification of the particles is possible. The signatures found in the detector vary from particle to particle. Electrons and muons are reconstructed within the geometrical acceptance of

the tracker system. The basic reconstruction algorithms for leptons are discussed in the following sections. A more detailed discussion on the isolation and identification selection criteria applied for this analysis is found in Chapter 6.

5.5.1 Muon reconstruction

Muons are minimum ionizing particles (MIP) for a large energy range, leaving only a small amount of their energy in the calorimeters and, as a result, are able to traverse all components of the CMS detector. Information from both the inner silicon tracker and outer muon system are used in the muon reconstruction, where trajectories found in the tracker are matched to tracks in the muon detector. Tracks reconstructed using the inner tracker system are so-called “tracker tracks”, while those that use the muon system are referred to as “stand-alone muon tracks”. The stand-alone muon tracks are reconstructed with a Kalman filter algorithm in order to fit the hits in the muon stations. The CMS collaboration relies on two main algorithms to reconstruct PF muons, which combines information from both types of track reconstructions [91, 92].

The first algorithm consists of an outside-in approach where the reconstruction matches standalone muon tracks with tracker tracks and links hit information to construct a global muon track by performing a fit. When the lowest χ^2 value of the global fit between the two types of tracks is found the tracks are linked. These PF muon candidates are referred to as “global muons”. Only global muons with the lowest χ^2 are retained to suppress muons arising from hadronic punch-through into the muon system and from muon decay in-flight. Requirements are imposed that the $\chi^2/n_{\text{dof}} < 10$ for the global muon track fit. Additionally, the fit is required to contain muon hit segments in at least two of the muon stations, at least one pixel hit, and more than five tracker layers with hits.

The second algorithm consists of an inside-out approach where the reconstruction starts with a reconstructed track in the tracking system that is extrapolated to the muon system. This is accomplished by matching tracker tracks to at least one muon hit segment in the

muon system. PF muons candidates reconstructed in this manner are referred to as “tracker muons”. The candidates with the lowest χ^2 value is selected when there is ambiguity concerning track reconstruction or matching.

Another category of muon candidates exists called “stand-alone muons”, which are meant to gather stand-alone muons tracks that are rejected by the global and tracker muon reconstruction algorithms. These types for muons candidates account for about 1% of the muons reconstructed in the detector and are usually due to cosmic radiation. For low p_T values the tracker muon reconstruction has a better efficiency than the global muon reconstruction since it only requires one muon hit segment in the muon system, while the latter only becomes more efficient with at least two muon hit segments.

5.5.2 Electron reconstruction

Electrons are the lightest and most stable of the charged leptons. A combination of a stand-alone approach [93] and a global PF algorithm [94, 83, 95] are used to optimize the electron reconstruction, which associates a track reconstructed in the silicon tracking system with a cluster of energy in the ECAL system. Electrons transversing through the tracking system curve due to the magnetic field of the CMS detector and, therefore, undergo bremsstrahlung radiation, whereby photons are emitted, and as a consequence the electrons lose some of their energy. The trajectory of the electron is reconstructed using a modeling of the electron energy loss from bremsstrahlung radiation and fitted with a Gaussian sum filter (GSF), where a Gaussian mixture is used rather than a single Gaussian distribution, as is the case for the Kalman filter. Almost all of the energy of an electron is deposited onto ECAL crystals once they have passed the tracker volume.

PF electrons are reconstructed using a tracker-driven seeding algorithm for the track reconstruction process. The energy deposited in the ECAL crystal by the photons arising from the bremsstrahlung radiation are reconstructed as ECAL clusters by extrapolating a straight line tangent to the electron track as it traverses through each layer of the tracker.

The PF electrons momentum is determined by combining the information from the ECAL clusters and track. In addition to the PF algorithm for electron reconstruction, a complementary stand-alone method is also implemented. An ECAL-driven seeding algorithm for the track reconstruction process is used instead. This approach uses ECAL superclusters seeds that are required to be matched to tracks, which are composed of hits in the inner track layers. Electrons reconstructed in this manner are referred to as GSF electrons and their final energy is taken from a combination of the supercluster energy and GSF track. The PF algorithm for electron reconstruction performs better for electrons that have low p_T or that are inside jets, while stand-alone approach is optimized for isolated electrons.

5.5.3 Hadronic τ -lepton reconstruction

The τ -lepton has several decay modes, where approximately 35.08% of the time it decays, within the primary vertex of the event, to a light-lepton and a pair of neutrinos [17]. These decay modes are included in the electron and muon selections. The remaining 64.92% of the time the τ -lepton decays to various hadronic final states. The CMS collaboration developed a “decay mode” based hadronic τ -lepton, τ_h , identification method called the hadron plus strips (HPS) algorithm [96]. The method takes into account pileup and has a relatively high efficiency with a low misidentification rate.

Hadron plus strips algorithm

The HPS algorithm for the reconstruction of τ_h -leptons uses PF jets reconstructed with the anti- k_T clustering algorithm [97, 94]. The PF jets use a distance parameter of $\Delta R = 0.5$. The algorithm then proceeds to reconstruct the photons, which are the decay products of the π^0 originating from the τ_h -lepton decay, into “strips”. The strips are centered on the most energetic electromagnetic particles within the PF jets and have a window size of $\Delta\eta = 0.05$ and $\Delta\phi = 0.20$, to take into account the broadening of calorimeter energy deposits from photon conversions. All electromagnetic particles associated with the strip,

as well as the charged hadrons are used in the calculation of the momentum and energy of the τ_h -lepton for a given decay mode. The individual decay modes are reconstructed by combining the charged hadron and all strips with a minimum p_T of 1 GeV within a cone size of $\Delta R = 2.8 \text{ GeV}/p_T^{\tau_h}$, where $p_T^{\tau_h}$ is the transverse momentum of the τ_h -lepton. The invariant mass of strips and charged hadrons should be consistent with the intermediate resonances π^0 , ρ , or $a_1(1260)$ [97].

The HPS algorithm considers three different decay topologies, separately, depending on the strip signatures the neutral pions produce, specifically one charged hadron with up to two neutral hadrons and three charged hadrons with up to one neutral hadron. The *single hadron* decay mode corresponds to the $h^-\bar{\nu}_\tau$ and $h^-\pi^0\bar{\nu}_\tau$ decays, where the low energetic neutral pions fail to be reconstructed as strips. The *single hadron plus one strip* decay mode corresponds to the $h^-\pi^0\bar{\nu}_\tau$ decay, where the emerging photons from the $\pi^0 \rightarrow \gamma\gamma$ decay are too close together and will, as a result, be reconstructed as one strip. The *single hadron plus two strips* decay mode corresponds to the $h^-\pi^0\bar{\nu}_\tau$ decay, where photons from the $\pi^0 \rightarrow \gamma\gamma$ decay are reconstructed as two strips since they are well separated. Lastly, the *triple hadron* decay mode corresponds to the $h^-h^+h^-\bar{\nu}_\tau$ decay, where the three charged hadrons are required to originate from the same secondary vertex. These decay topologies are used to reconstruct the $h^-\pi^0\pi^0$ and $h^-h^+h^-\pi^0\bar{\nu}_\tau$ decay modes as well.

An isolation criteria is applied to distinguish PF τ_h -lepton from normal jets. We require that the $\sum p_T$ of charged and neutral hadrons, excluding the decay products of the τ_h -lepton, within an isolation cone of size $\Delta R = 0.5$ around the τ_h -lepton be less than a given threshold. A $\Delta\beta$ -correction is applied to the τ_h -lepton isolation measurement to remove any contributions from charged and neutral hadrons arising from pileup interaction. The HPS PF τ_h -lepton reconstruction and identification efficiencies are 70%, 53% and 33% for the “loose”, “medium”, and “tight” working points with a misidentification rate of 1%, 0.4% and 0.2%, respectively [97].

5.6 Jet reconstruction

Quarks and gluons are produced in high rates in the hard process of proton-proton collisions or from gluon radiation in strong interactions due to colliding quarks. In addition, these partons are not directly detected, however, due to the nature of the strong force, they hadronize into sprays of collimated particles that are reconstructed as jets. The purpose of jet reconstruction is to measure the momentum of partons which began the hadronization process. The reconstruction is performed by grouping together final state particles which are visible in the detector, through the selection of an appropriate jet algorithm. This algorithm is applied to reconstructed objects and generator-level particles [94]. These reconstructed jets are later matched to their corresponding generator jet. At generator level, the list of final state particles produced in the hadronization process constitute the list of objects that will be clustered [98]. At the reconstruction level a list of particle candidates is passed to the algorithm, produced by the particle-flow technique.

5.6.1 Jet algorithm

Jets are by nature composite objects that require an algorithm to define them. The jet clustering algorithms combines particles that result from the hadronization process of partons. Jet algorithms are required to be both collinear- and infrared-safe in order to provide finite theoretical predictions to all orders of QCD perturbation theory. Collinear safety refers to when a particle with a certain momentum is replaced by two collinear particles with half the original momentum and the result of the clustering sequencing remains unchanged. On the other hand, infrared safety involves infinitely soft gluons added to set of particles which have to be clustered and the results of the clustering sequence also remains unaffected. In essence, infrared-safe refers to the robustness against the addition of low energetic particles arising from long distance interaction, while collinear-safe implies the stability of the algorithm when energetic particles are split into two or more soft collinear partons.

A sequential recombination jet cluster algorithm is implemented instead of “fixed” or

“iterative” cone algorithms. In principle, the algorithm clusters pseudo-particles based on a recombination distance metric between two candidate particles. A common clustering algorithm is characterized by defining the distance d_{ij} between two inputs pseudo-particles “ i ” and “ j ” as,

$$d_{ij} = \min(k_{T,i}^{2n}, k_{T,j}^{2n}) \cdot \frac{\Delta R_{ij}^2}{R^2} \quad (5.2)$$

where $k_{T,i}$ and $k_{T,j}$ are the transverse momenta of the two pseudo-particles, respectively, such that $\Delta R_{ij} = \sqrt{\Delta\eta_{ij}^2 + \Delta\phi_{ij}^2}$ is the euclidean distance between them in the $\eta - \phi$ plane, while R is the algorithm’s radius parameter. Another distance parameter is defined as well between the beam-line and “ i ” pseudo-particle and is given by,

$$d_{iB} = k_{T,i}^{2n} \quad (5.3)$$

The sequential clustering algorithm proceeds by determining the lesser of the two values d_{ij} and d_{iB} . If $d_{ij} < d_{iB}$ then it recombines particles “ i ” and “ j ” by summing their 4-momentum into a new pseudo-particle. Otherwise, if $d_{iB} > d_{ij}$ then it removes particle “ i ” from the list of pseudo-particle and promotes the pseudo-particle to a jet candidate. The algorithms iterates until only jet candidates remain.

The type of jet algorithm is specified by the value of n . For $n = -1$ results in the anti- k_T jet algorithm [99], $n = 0$ results in the Cambridge-Aachen algorithm [100], and $n = 1$ results in the k_T algorithm [101]. Moreover, these jet algorithm requires the cutoff parameter R to be specified, defining a characteristic size of the jets, where pairs of jet candidates satisfy $\Delta R_{ij} > R$. The jet algorithm definition does not allow jets to contain particles at distances greater than R from their central axis. All jets used in this analysis are defined through the anti- k_T algorithm with spatial separation $R = 0.5$.

Jet energy corrections, which are functions of the jet p_T and ϕ , are applied according to Reference [102] to account for the nonlinear and nonuniform detector response. Additionally,

contributions from overlapping pp interactions from pileup to the jet p_T values are removed using the jet area method described in Reference [103].

5.6.2 b-tagged jets

Flavor tagging algorithms, such as those for b-tagged jets [104], that rely on identifying jets arising from the hadronization of heavy flavored b quarks play an important role in many physics analyses. They provide a method by which to reduce large background contributions from processes involving light quark fragmentation, jets from gluon, and from lighter c-quark hadronization. These b jets are identified (i.e. tagged) by using a combined secondary vertex (CSV) b-tagging algorithm [104], which uses a track-based lifetime and simple secondary vertex (SSV) [105] information associated with jets to compute a likelihood-based discriminate in order to distinguish jets originating from b-quark hadronization from those coming from gluons, or charm and light flavor quarks.

The main property used in identifying b quarks come from the large lifetimes of B-hadrons, of about 1.5 ps (decay length of about 450 μm) [106]. The CSV algorithm returns a b-tagging discriminate value between 0 and 1, with higher values indicating a higher probability of the jet to originate from a b quark, while lower values indicate that the jet is more likely to have originated from light-flavor quarks or gluons. A b-tagging discriminant threshold value of 0.679 is selected, giving a tagging efficiency of about 70% with a misidentification rate of about 1% for light-flavor jets. The efficiency for b-tagging jets is measured in both data and simulation. In order to account for any observed differences between data and simulations corrections are applied to simulated events, depending on the p_T and η of the b jet. Various working points are available for the b-tagging jet algorithms, indicating a misidentification rate of light quark jets of 10% (loose), 1% (medium), or 0.1% (tight). Further information on b-tagging efficiencies is provided in Section 6.4.

5.7 E_T^{miss} reconstruction

In the CMS experiment, proton-proton collisions may produce particles that escape detection, such as neutrinos, which only interact weakly and, therefore, do not leave tracks in the inner tracker or muon system, nor do they deposit energy in the electromagnetic or hadronic calorimeters. Several BSM theories predict the production of new exotic particles that may also go undetected. For this reason, the reconstruction of E_T^{miss} is of great importance for many physics analyses. We detected these particles indirectly by measuring the amount of momentum imbalance, left behind in the CMS detector, in the plane transverse to the proton beam. Hence, the missing transverse momentum is defined, using momentum conservation, as the negative of the vector sum of the transverse momentum of all stable final-state particles reconstructed in the detector, and is expressed by,

$$\vec{E}_T = - \sum_i \vec{p}_T^i, \quad (5.4)$$

where the sum is over the final-state particles. The magnitude of the missing transverse momentum vector $\vec{E}_T$ gives the missing transverse energy E_T^{miss} . The reconstruction of E_T^{miss} is affected by many factors including any mis-measurement of particle momenta, large shower fluctuations, non-linear calorimeter response, instrumental noise or poorly instrumented regions of the detector, beam-halo particles, and cosmic-ray particles.

The CMS collaboration developed three different methods for the E_T^{miss} reconstruction, specifically the PF E_T^{miss} , Calo E_T^{miss} , and TC E_T^{miss} algorithms [107, 108, 109, 110]. The PF E_T^{miss} is calculated from visible reconstructed PF candidates, the Calo E_T^{miss} is calculated using calorimeter energies and the calorimeter tower geometry, while the TC E_T^{miss} is based on the Calo E_T^{miss} but uses tracks reconstructed in the inner tracker to improve the E_T^{miss} response and resolution. We use PF E_T^{miss} for the purposes of this analysis.

The E_T^{miss} response and resolution is studied in $Z \rightarrow \mu^+ \mu^-$ events since there is no intrinsic missing transverse energy. The transverse momentum of a vector boson and hadronic

activity should be balanced in an event, which we expressed as,

$$p_{\text{T}}^{\text{Z}} + \vec{u}_{\text{T}} + \vec{\cancel{E}}_{\text{T}} = 0 \quad (5.5)$$

where p_{T}^{Z} is the transverse momentum of the Z boson and $\vec{u}_{\text{T}}$ is vector sum of the transverse momenta of all particles except the Z boson, referred to as the hadronic recoil. The energy correction to the $E_{\text{T}}^{\text{miss}}$ response is obtained from the projection of the hadronic recoil onto an axis parallel to the direction of the vector boson momentum, $u_{\parallel}$. While the correction to the $E_{\text{T}}^{\text{miss}}$ resolution is determined from the standard deviation of the projection of the hadronic recoil $\vec{u}_{\text{T}}$ onto an axis perpendicular to the direction of the vector boson momentum, $\sigma(u_{\perp})$.

Due to a systematic difference of the reconstructed $E_{\text{T}}^{\text{miss}}$ with respect to the amount of energy carried away by invisible particles, corrections are applied in order to account for this effect. There are several types of corrections that can be applied to the $E_{\text{T}}^{\text{miss}}$ response, such as the so-called “type-0” correction, which reduce the effects of pileup by subtracting charged hadrons and compensating for the remaining imbalance from neutral hadrons, and “type-I” corrections, which propagate the jet energy scale corrections to the $E_{\text{T}}^{\text{miss}}$.

Chapter 6

Physics Object and Event Selections

The object reconstruction algorithms developed by the CMS collaboration were discussed in the previous chapter. This chapter introduces the object identification and selection requirements, as well as the final event selection, used in the analysis.

6.1 Object identification and selection

Particles are identified by the characteristic signatures they leave behind in detectors. Electrons and muons are reconstructed with the particle flow (PF) algorithm, as previously mentioned, which uses measured quantities from the tracker system, electromagnetic calorimeter, hadron calorimeter, and muon system. The tracking system determines whether these particles are charged. The neutral particles are detected in the calorimeters, where their energies are measured. The calorimeters also help determine whether the particles interacted electromagnetically or hadronically. Muons do not generally produce electromagnetic showers, therefore, they are identified by their presence in the most outer parts of the detector, since other particles are absorbed in the calorimeter system. Candidate particles that match tracks are required to satisfy quality selection criteria, as well as spatially match with the energy deposits in either the electromagnetic or hadronic calorimeter, or with tracks found in the muon system. The purpose of the additional analysis level selection criteria required on the physics objects is to ensure that the number of “high” quality reconstructed objects is maximal and to reduce the amount of background contributions arising from fake objects.

Genuine lepton background sources can arise near or inside jets, hadronic showers with large electromagnetic showers, photon conversions, and hadrons misidentified as leptons from punch-throughs into the muon system. Certain lepton selection criteria such as isolation requirements can greatly reduce background from misidentified leptons, because most of these leptons occur inside jets. A kinematic variable called relative isolation, I_{rel} , is defined as the ratio of the sum p_T of all particle flow candidates in a cone $\Delta R < 0.3$, around the lepton, divided by the p_T of the lepton. We require electrons, muons, and isolated tracks, to have $I_{\text{rel}} < 0.15$.

Once the isolation selection is applied, the most significant source of background leptons originate from heavy B-meson decays, where the leptons are characteristically more isolated due to the high p_T with respect to the jet direction. These types of backgrounds are reduced by requiring leptons to come from within one centimeter from the primary vertex in the z direction with a small impact parameter, between the event vertex and track in the plane transverse to the beam direction, of $d_{xy} < 0.02$ cm.

In the following sections, we list the selection requirements for muons, electrons, τ -leptons, photons, jets, b jets, and E_T^{miss} .

6.1.1 Muon selections

Muons, as previously described in Section 5.5.1, are reconstructed from the particle flow (PF) algorithm [84] in three steps using information from the tracker, calorimeter, and muon system. The first step is the reconstruction of the “tracker muon” using information from the tracker system. The next step is the reconstruction of the “stand alone muon” based only on information from the muon chamber. In the last step, we link information from both the tracker and muon chamber together to obtain a combined fit for the reconstruction of the “global muon”. A summary of all the selection criteria muons must satisfy for the purposes of the analysis is given in Table 6.1. We briefly describe the meaning of the different muon selection criteria below.

- We require PF muons with $p_T > 10 \text{ GeV}$ and $|\eta| < 2.4$.
- Global χ^2/n_{dof} is the χ^2 per degree of freedom of the fit to the tracks left behind by the muon as it traverses the detector.
- ***TrackerLayersWithMeasurement***: is the number of tracker layers with hits.

Table 6.1: Muon selection requirements.

Muon observable	Selection
Identification (ID)	Particle Flow and Global
Transverse momentum p_T	$> 10 \text{ GeV}$
Pseudo rapidity $ \eta $	< 2.4
Global χ^2/n_{dof} of the fit	< 10
Number of valid pixel hits	> 0
TrackerLayersWithMeasurement	> 5
Number of valid hits in muon chamber	> 0
Number of muon stations with muon segments	> 1
Transverse impact parameter $ d_{xy}(PV) $	$< 0.02 \text{ cm}$
Longitudinal impact parameter $ d_z(PV) $	$< 0.5 \text{ cm}$
Relative PF isolation I_{rel} within $\Delta R < 0.3$, with beta corrections for pileup	< 0.15

6.1.2 Electron selections

Electrons, similarly to muons, are also reconstructed with the PF algorithm [95], as described in Section 5.5.2, by combining information from the tracker and ECAL system. A summary of all the selection criteria electrons must satisfy is given in Table 6.2. We briefly describe the meaning of the different electron selection criteria below.

- We require PF electrons with $p_T > 10 \text{ GeV}$ and $|\eta| < 2.4$, but with the additional requirement that electrons in the gap region $1.4442 < |\eta| < 1.566$ are rejected.
- $\Delta\eta_{\text{In}}$ is the pseudorapidity difference between SuperCluster position and track direction at vertex extrapolated to ECAL assuming no radiation.
- $\Delta\Phi_{\text{In}}$ is the azimuthal difference between SuperCluster position and track direction at vertex extrapolated to ECAL assuming no radiation.

- $\sigma_{i\eta i\eta}$ is the supercluster η width taken from the covariance matrix using logarithmic weights, where $i\eta$ represents the i^{th} ECAL crystal in the η direction. In other words, it is the second moment of the ECAL energy cluster distribution in the η direction.
- Hadronic leakage variable H/E is the ratio between the energy deposit recorded in the HCAL tower, just behind the ECAL supercluster seed, and the ECAL supercluster energy associated with the electron.
- $|1/E - 1/p|$ is the absolute value of the difference between the reciprocal of the electron energy and the reciprocal of the magnitude of the electron momentum.
- Conversion rejection cut allows for rejecting electrons identified as originating from the conversion of a photon. Further information on photon conversion is found in Section 7.3.4.
- Electrons are required to be $\Delta R > 0.1$ away from selected muons.

Table 6.2: Electron selection requirements. Several electron ID criteria are different for the barrel ($|\eta| < 1.44$) and endcap ($1.56 < |\eta| < 2.4$) regions. Electrons in the gap region $1.4442 < |\eta| < 1.566$ are rejected.

Electron observable	Selection	
	Barrel	Endcap
Identification (ID)	Particle Flow	Particle Flow
Transverse momentum p_T	$> 10 \text{ GeV}$	$> 10 \text{ GeV}$
Pseudo rapidity $ \eta $	< 2.4	< 2.4
Spatial (η) matching between track and supercluster $\Delta\eta_{\text{In}}$	< 0.007	< 0.009
Spatial (Φ) matching between track and supercluster $\Delta\Phi_{\text{In}}$	< 0.15	< 0.10
Transverse shape of the electromagnetic cluster $\sigma_{i\eta i\eta}$	< 0.01	< 0.03
Hadronic leakage variable H/E	< 0.12	< 0.10
Transverse impact parameter $ d_{xy}(PV) $	< 0.02	$< 0.02 \text{ cm}$
Longitudinal impact parameter $ d_z(PV) $	< 0.1	$< 0.2 \text{ cm}$
$ 1/E - 1/p $	< 0.05	< 0.05
Relative PF isolation I_{rel} within $\Delta R < 0.3$	< 0.15	< 0.15
Conversion rejection cut	0	0
Number of missing expected inner tracker layer hits	< 2	< 2
ΔR to nearest selected muon	> 0.1	> 0.1

6.1.3 τ_h -lepton selections

Tau leptons can decay leptonically, τ_ℓ , to electrons or muons, as well as hadronically, τ_h , to pions or kaons. Hadronically decaying τ -leptons can produce either a single charged track (one-prong) or three charged tracks (three-prong), with or without the presence of additional electromagnetic energy due to the decay of neutral pions. Reconstruction of the τ_h -lepton is performed with the hadron plus strips (HPS) algorithm using charged hadrons and photons to construct the various hadronic decay modes. Candidate particles that appear to be consistent with mis-reconstructed light-leptons are rejected. The τ_h -leptons are required to have a p_T greater than 20 GeV and $|\eta| < 2.3$. We require that τ_h -lepton candidates satisfy tau discriminate selections, recommended by the CMS collaboration, and which have an inherent isolation threshold, in order to identify τ_h -leptons for use in the analysis. A summary of all the τ_h -lepton selections is given in Table 6.3. We briefly describe the meaning of the different τ_h -lepton selection criteria below.

- We require HPS PF Taus with $p_T > 20 \text{ GeV}$ and $|\eta| < 2.3$.
- ***ByDecayModeFinding***: A discriminant value that determines whether the HPS algorithm is able to reconstruct one and three-prong decay modes for τ_h -lepton candidates.
- ***AgainstElectronLoose***: A discriminant value that rejects misidentified electrons, since they are prone to be reconstructed as one-prong τ_h -leptons, by requiring that the electron pion multivariate analysis (MVA) value be less than 0.6.
- ***AgainstMuonLoose***: A discriminant that rejects misidentified muons by requiring that there is no ΔR matching between the leading track of the τ_h -lepton candidate and chamber hits left by a muon.

- ***ByLooseCombinedIsolationDBSumPtCorr***: An isolation discriminant value corresponding to the sum p_T of all charged and neutral candidates in the isolation annulus, taking into account the effects of pileup, with p_T greater than 0.5 GeV and $\Delta R = 0.5$.
- τ_h -leptons are required to be $\Delta R > 0.1$ from selected light-leptons.

Table 6.3: τ_h -lepton selection requirements.

Hadronic τ -lepton observable	Selection
Identification (ID)	HPS Particle Flow
Transverse momentum p_T	$> 20 \text{ GeV}$
Pseudo rapidity $ \eta $	< 2.3
ByDecayModeFinding	1
AgainstElectronLoose	1
AgainstMuonLoose	1
ByLooseCombinedIsolationDeltaBetaCorr	1
PF isolation E_τ^{iso}	$< 2.0 \text{ GeV}$
ΔR to nearest selected light-leptons	> 0.1

6.1.4 Photon selections

Photon reconstruction is based on clustering energy deposits in the ECAL system, so-called superclusters (SC) [89], as previously described in Section 5.4. We require the supercluster not to match pixel hits consistent with a track from the interaction region. The photon candidates are discriminated against background processes with the use of isolation variables that require a cone size of $\Delta R = 0.3$ for all isolation sums. A summary of all the photon selections is given in Table 6.4.

6.1.5 Jet selections

The jet reconstruction is performed using the anti- k_T algorithm [99] with cone radius of $\Delta R = 0.5$ and a list of PF candidates as input, as previously described in Section 5.6. The transverse momenta requirements for jets are chosen to obtain reliable jet energy reconstruction, due to the large uncertainties from low energy jets. Furthermore, the jet

Table 6.4: Photon selection requirements. Several photon ID criteria are different for the barrel ($|\eta| < 1.44$) and endcap ($1.56 < |\eta| < 2.4$) regions.

Photon observable	Selection	
	Barrel	Endcap
Transverse momentum p_T	$> 10 \text{ GeV}$	$> 10 \text{ GeV}$
Pseudo rapidity $ \eta $	< 2.4	< 2.4
Conversion safe electron veto	1	1
Single tower H/E	< 0.06	< 0.05
Transverse shape of the electromagnetic cluster $\sigma_{i\eta i\eta}$	< 0.011	< 0.034
Rho corrected relative PF charged hadron isolation	< 0.06	< 0.05
Rho corrected relative PF neutral hadron isolation	< 0.16	< 0.10
Rho corrected PF photon isolation	< 0.08	< 0.12

energy corrections are obtained from data and applied to simulations to account for any non-linearity responses of the detector calorimeter, and pileup effects due to multiple particle interactions. In order to account for the CMS detector acceptance jets are considered only in the range of $|\eta| < 2.5$. Moreover, following the recommendations of the CMS collaboration we use the loose jet identification criteria. A summary of all the jet selections is given in Table 6.4. We briefly describe the meaning of the different jet selection criteria below.

- We require PF jets with $p_T > 30 \text{ GeV}$ and $|\eta| < 2.5$.
- **Number of constituents:** A jet must consist of more than one constituent in its reconstruction.
- **Residual corrections:** For simulation samples we apply L1FastL2L3 corrections to the PF jets, while for data we apply L1FastL2L3residual corrections. The p_T and η dependence between simulation and data differ, therefore, a correction is applied to jets in data in order to remove the observed difference.
- **Neutral hadron fraction:** The fraction of energy deposited in the hadron calorimeter from neutral particles is required to be less than 0.99.
- **Neutral EM fraction:** The fraction of energy deposited in the electromagnetic

calorimeter from neutral particles is required to be less than 0.99.

- ***Charged hadron fraction:*** The fraction of energy deposited in the hadron calorimeter from charged particles must be greater than zero for $|\eta| < 2.4$.
- ***Charged EM fraction:*** The fraction of energy deposited in the electromagnetic calorimeter from charged particles must be smaller than 0.99 for $|\eta| < 2.4$.

Table 6.5: Jet selection requirements.

Jet observable	Selection
Transverse momentum p_T	$> 30 \text{ GeV}$
Pseudo rapidity $ \eta $	< 2.5
Number of constituents in jet	> 1
Neutral hadron fraction of total jet energy	< 0.99
Neutral EM fraction of total jet energy	< 0.99
Charged hadron fraction ($ \eta < 2.4$)	> 0
Charged EM fraction ($ \eta < 2.4$)	< 0.99
Number of tracks ($ \eta < 2.4$)	> 0

***b*-jet identification**

Events are additionally classified based on the number of identified *b* jets, those jets consistent with originating from the bottom quark hadronization process. The secondary vertex (SV) technique is used to identify jets containing products from B-mesons. The SV reconstruction is performed with an inclusive vertex search from a given list of tracks [105]. These SV candidates are required to share less than 65% of their tracks with the primary vertex (PV) and their radial distance to the PV must exceed vertices coming from long lived particles. The combined secondary vertex (CSV) algorithm [104] is used for the identification of *b* jets. This CSV algorithm combines information from track impact parameters and secondary vertices in a jet to form a likelihood discriminant that distinguishes between *b* jets and other light-flavor jets. We use a “medium” working point for the analysis, which has a 70% *b*-tagging efficiency and a 1% mis-tagging rate for light-flavor jets.

6.1.6 E_T^{miss} selection

As previously described in Section 5.7, the missing transverse energy in an event is defined by the negative vectorial sum of the transverse momentum of all PF candidates. The main source of E_T^{miss} in an event arises due to particles, such as neutrinos, that carry away momentum which cannot be directly measured with the CMS detector. However, the limited acceptance range of the detector, defective calorimeter cells, and momentum miss-measurements of particles can all contribute to the E_T^{miss} measurement.

The quality selection requirements for E_T^{miss} are as follows:

- We use PF E_T^{miss} physics objects.
- ***Primary vertex filter***: This is an event filter that is applied and which requires that at least one “good” primary vertex be reconstructed in the event. A PV is identified as good if it has a number of degree of freedom $n_{\text{dof}} > 4$ and has a position of $|z| < 24$ cm and $\rho < 2$ cm, to reject noisy events, due to pileup, and ensure good collision candidates.
- The following filters, recommended by the CMS collaboration, are applied to veto events that could contain an increased amount of fake E_T^{miss} :
 - ***CSC tight beam halo filter***: Secondary particle showers resulting from beam-gas collisions in the vacuum chambers induce beam halo noise in the detector. Therefore, this filter is used in order to identify events containing large beam backgrounds.
 - ***HBHE noise filter***: Rejects isolated noise originating from the HCAL barrel and endcap readout electronics, which may be mis-reconstructed as hadronic energy deposit.
 - ***ECAL dead cell trigger primitive filter***: This is applied in order to reject fake E_T^{miss} coming from high energy particles which have deposited their energy in

noisy crystal cells in the ECAL that have been left out of the event reconstruction.

- ***Tracking failure filter***: The tracking algorithms may fail for some of its iterations when too many cluster hits are found. Therefore, a selection requirement based on the ratio of the $\sum p_T$ of all tracks associated with the primary vertex to the H_T in the event is applied to reject such events.
- ***Bad EE Supercrystal filter***: Designed to reject events containing anomalously large energy in the ECAL endcap superclusters.

6.2 Measurement of the lepton trigger efficiency

The LHC experiment produces vast amounts of collision data all of which cannot be stored for subsequent analysis due to our limited processing and storage capacity, as discussed in Section 3.3. The purpose of triggers is to select potentially interesting events to be recorded, while the rest of the data is safely discarded. The physics objects of interest used in a search indicate what triggers are to be chosen. Therefore, in a multilepton search we select triggers that fire on events with at least two lepton candidates. We use data collected by several different double-lepton triggers for our analysis, which use isolation, good quality tracks and vertices, and other requirements in order to select leptons. A complete list of the un-prescaled dilepton triggers used in the analysis that have the lowest available muon and electron p_T thresholds is given in Appendix B.

We measure the lepton trigger efficiencies in data in order to correct the MC simulations since we chose not to model the triggers in the samples. The average efficiency for a sample, which satisfies two trigger selections, can be expressed as $\epsilon_{ij} = \epsilon_i \times \epsilon_j$ if they are assumed to be uncorrelated, where ϵ_i and ϵ_j are the efficiencies for the i^{th} and j^{th} trigger, respectively. If we are interested in the ϵ_i trigger efficiency then we must measure the ratio ϵ_{ij}/ϵ_j .

The trigger efficiencies are obtained from an H_T triggered data sample, to ensure an unbiased selection, by using independent jet energy triggers [111, 112, 113]. The kinematic quantity H_T is a measure of the jet activity in an event defined by the sum of transverse

momentum of selected jets. We determine the trigger efficiency for every double lepton trigger by counting the number events that fire the triggers of interest, given in the formula,

$$\epsilon_{\text{trigger}} = \frac{N(2\ell \text{ \& dilepton trigger fired \& } H_T \text{ trigger fired})}{N(2\ell \text{ \& } H_T \text{ trigger fired})}, \quad (6.1)$$

where the numerator corresponds to the number of events with at least two leptons that pass lepton isolation requirements, have $p_T > 10 \text{ GeV}$, and that have fired both dilepton and H_T triggers, while the denominator is the same except that the dilepton triggers are not required to have fired. Additionally, we require $H_T > 550 \text{ GeV}$ and $E_T^{\text{miss}} > 180 \text{ GeV}$, or $H_T > 300 \text{ GeV}$ and $E_T^{\text{miss}} > 70 \text{ GeV}$ to suppress any bias effects from the H_T trigger turn-on curve by being in the plateau region, thereby removing any correlations with respect to the dilepton trigger efficiency. The large number of double-muon triggers demands that the logical “OR” be used in the trigger efficiency calculation, given in Equation 6.1, for all triggers in this category, similarly, for the double-electron and electron-muon triggers.

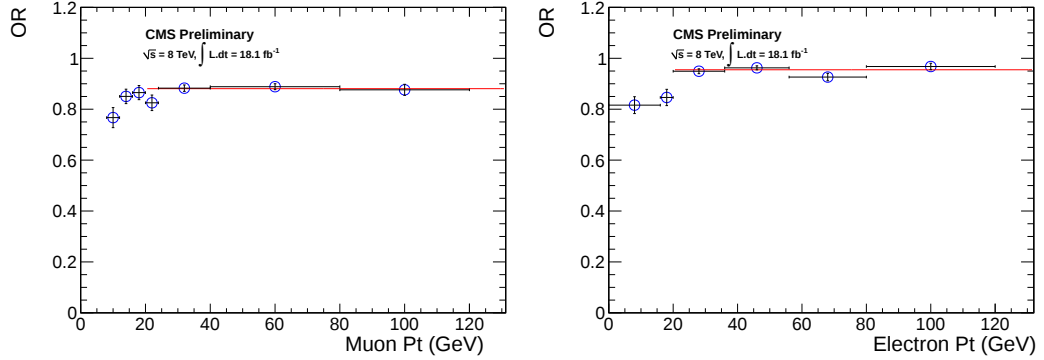


Figure 6.1: Double muon “OR” efficiency (left) and double electron “OR” efficiency (right) as a function of the sub-leading muon and electron p_T , respectively.

Figure 6.1 shows the efficiencies for the double-muon and double-electron triggers with respect to the sub-leading muon or electron p_T , respectively. The double-muon trigger has an efficiency of 90% with no notable dependence on p_T . The double-electron trigger efficiency is 95% when the sub-leading electron has $p_T > 20 \text{ GeV}$, but is only 82% when the sub-leading electron has $p_T < 20 \text{ GeV}$. The double-electron trigger reaches a 95%

efficiency with a sharp turn-on curve. Lastly, the electron-muon trigger efficiency is 93% when the sub-leading lepton has $p_T > 20$ GeV and is 86% when the sub-leading lepton has $p_T < 20$ GeV. The systematic uncertainties associated with the lepton trigger efficiencies give rise to uncertainties in the SM background estimates, as described in Section 8.3. Simulation samples are corrected for measured lepton trigger efficiencies by weighing events according to the likelihood that an event will fire the double-lepton triggers.

Table 6.6: Lepton trigger efficiency for the “OR” of all dielectron and dimuon triggers determined using different H_T triggers paths.

Trigger name	Dilectron “OR” efficiency	Dimuon “OR” Efficiency
HLT_HT650_v7	$95.9\% \pm 1.1\%$	$93\% \pm 6\%$
HLT_HT750_v3	$93.6\% \pm 1.5\%$	$94\% \pm 2\%$
HLT_HT750_v4	$96.7\% \pm 2.4\%$	$91\% \pm 6.4\%$
HLT_HT750_v7	$96\% \pm 0.7\%$	$87\% \pm 1.8\%$

6.3 Measurement of the lepton isolation and identification efficiency

In this section, we describe the efficiency associated with the isolation and identification requirements on leptons at the analysis level, corresponding to the set of selection criteria for muons, electrons, and τ_h -leptons discussed previously. A common technique used for measuring lepton selection efficiencies is the “tag-and-probe” method based on $Z \rightarrow \ell^+\ell^-$ events. In general, the method requires the “tag” lepton to satisfy analysis level selections and the “probe” lepton to pass looser selection requirements. In this case, tag muons and electrons are required to pass the analysis selections listed in Sections 6.1.1 and 6.1.2, respectively, with the exception that the p_T requirement is increased to $p_T > 20$ GeV. We require probe muons to be global muons with $p_T > 5$ GeV and $|\eta| < 2.1$, whereas probe electrons have their selection requirements for $\sigma_{I\eta I\eta}$, H/E , $\Delta\eta$, and $\Delta\phi$ relaxed from those used in the analysis. The isolation and identification efficiency are inversely proportional to jet activity, as well as number of pileup interactions. In order to ensure that the selection efficiencies reflect the event kinematics, they are measured as a function of η , in two different regions of the CMS detector, namely, in the barrel ($|\eta| < 1.5$) and the endcap regions

($1.5 < |\eta| < 2.1$), and as a function of the number of jets and reconstructed vertices.

The isolation and identification efficiencies are measured separately and in several probe lepton p_T ranges. For MC simulations, we normalize the event count in the Z-boson mass region ($80 - 100$ GeV) to the same number of events found in data. We define the lepton isolation efficiency as,

$$\epsilon_{\text{Isolation}} = \frac{N(\text{probe leptons which satisfy isolation \& ID requirements})}{N(\text{probe leptons which satisfy ID requirements})}, \quad (6.2)$$

where the numerator is the number of events within the Z-boson mass range, such that the probe lepton satisfies isolation requirements in addition to having passed the identification selection, while the denominator is the same with the exception that the isolation requirement does not need to be satisfied. Similarly, the identification efficiency is defined as,

$$\epsilon_{\text{ID}} = \frac{N(\text{probe leptons which satisfy ID \& isolation requirements})}{N(\text{probe leptons which satisfy isolation requirements})}, \quad (6.3)$$

where the numerator is the number of events within the Z-boson mass range, such that the probe lepton passes the identification selection and satisfies isolation requirements, while the denominator is the same except that the identification selection does not need to be satisfied. Background contributions are taken into account by simultaneously fitting the invariant mass distribution of $Z \rightarrow \ell^+ \ell^-$ events in both the below Z-boson mass range ($55 - 80$ GeV) and the above Z-boson mass range ($100 - 125$ GeV) with a polynomial function to predict the number of background events within the Z-boson window.

Figure 6.2 shows the muon isolation (top left) and identification (top right) efficiency as a function of probe muon p_T . Similarly, Figure 6.2 shows the electron isolation (middle left) and identification (middle right) efficiency as a function of probe electron p_T . The measurement of the muon identification efficiency for MC simulation is well modeled when compared to collision data. On the other hand, this is not the case for the muon and

electron isolation efficiency when the p_T is less than 25 GeV, and the difference is even more noticeable for the electron identification efficiency.

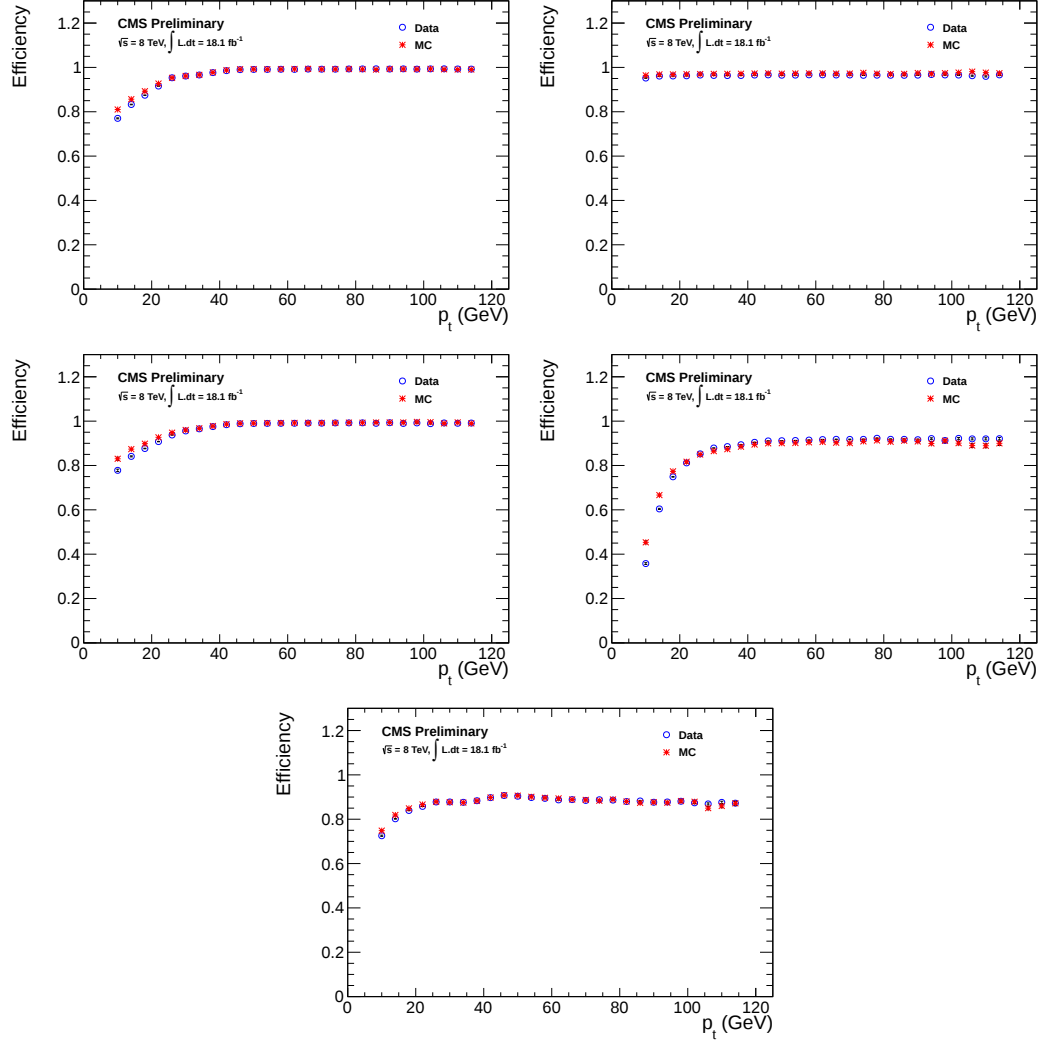


Figure 6.2: The isolation (top left) and identification (top right) efficiency as a function of the probe muon p_T . The isolation (middle left) and identification (middle right) efficiency as a function of the probe electron p_T . The isolation efficiency (bottom) as a function of the probe τ_h -lepton p_T .

We measure the isolation and identification efficiency for leptons in both collision data and MC simulations, however, some discrepancies are observed between them. In order to correct the modeling of the lepton selection efficiency in MC simulations a “scale factor” is calculated, which is subsequently applied to the simulations. The scale factor is defined as the ratio between the selection efficiency measured in data to that measured

in simulation. Simply stated, the ratios between the data and simulation values for each individual distribution in Figure 6.2 represent the scale factors associated with the lepton isolation and identification efficiencies, as shows in Figure 6.3. To have a more accurate measurement of the scale factors, we perform a fitting procedure on their distributions and parameterize them in terms of the their probe lepton p_T , as proposed in Reference [114]. The parametrized function for the lepton selection efficiency scale factor is given by,

$$SF_\ell(p_T) = SF_\ell^{\text{plateau}} \cdot \text{erf}\left(\frac{p_T - p_T^{\text{threshold}}}{\sigma_\ell}\right) + SF_\ell^0 \cdot \left[1 - \text{erf}\left(\frac{p_T - p_T^{\text{threshold}}}{\sigma_\ell}\right)\right], \quad (6.4)$$

where SF_ℓ^{plateau} is the efficiency scale factor in the plateau region of the plot, erf is the error function, $p_T^{\text{threshold}}$ is the specific p_T selection value of 8 GeV for leptons, SF_ℓ^0 is the efficiency scale factor value at $p_T^{\text{threshold}}$, and σ_ℓ determines the rate of change in value as p_T decreases.

The muon and electron isolation and identification efficiency factors are then estimated with Equation 6.4. The resulting fit parameters for the muon isolation efficiency scale factor are,

- $\sigma_\mu = 11.6361$
- $SF_\mu^{\text{plateau}} = 0.9985$
- $SF_\mu^0 = 0.9324$,

while the scale factor for the muon identification efficiency is constant, as can be seen from Figure 6.2 (top right), and is given by,

- $SF_\mu = 0.9925$

The fit parameters for the electron isolation efficiency scale factor are,

- $\sigma_e = 16.4017$
- $SF_e^{\text{plateau}} = 0.9982$

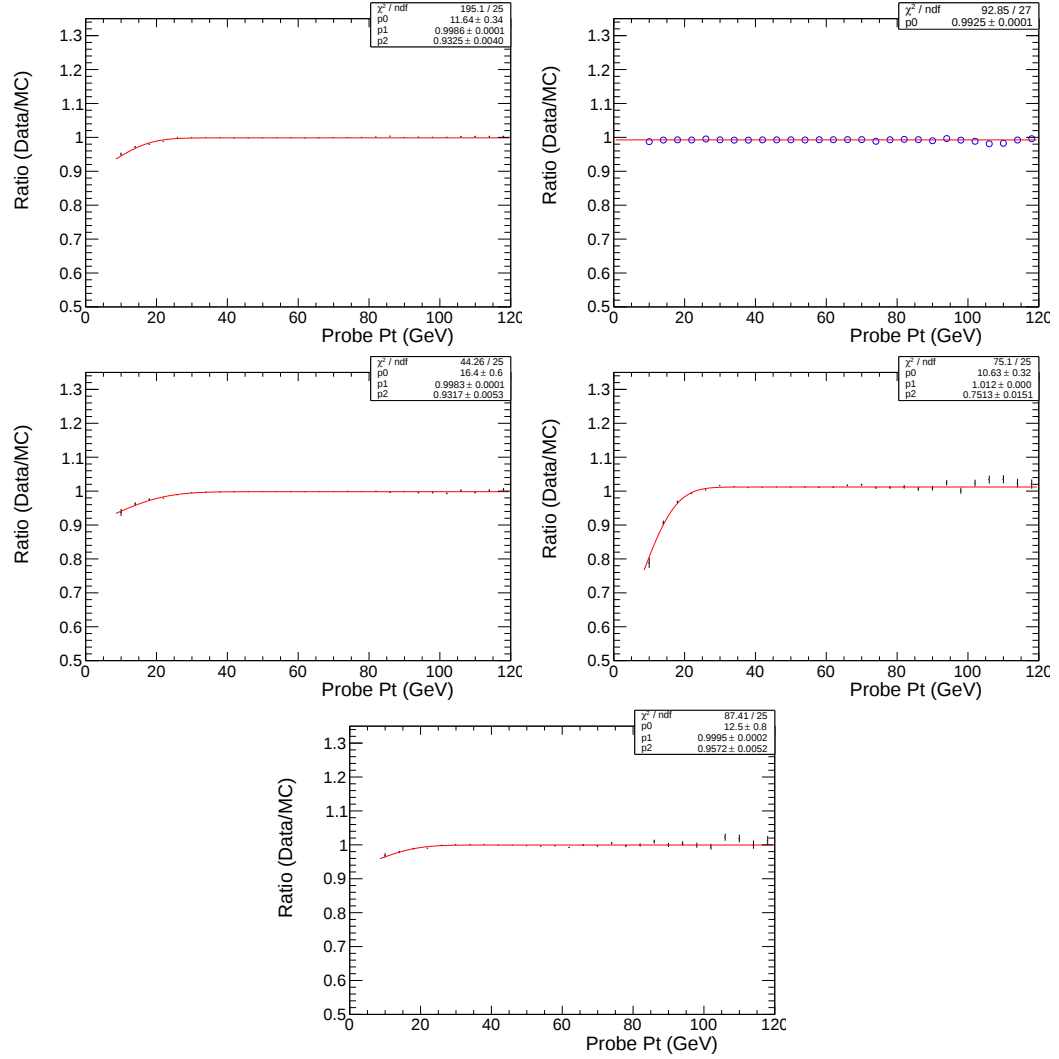


Figure 6.3: The isolation (top left) and identification (top right) efficiency scale factors as a function of the probe muon p_T . The isolation (middle left) and identification (middle right) efficiency as a function of the probe electron p_T . The isolation efficiency scale factor (bottom) as a function of the probe τ_h -lepton p_T .

- $SF_e^0 = 0.9316$.

while the electron identification efficiency scale factor has the following fit parameters,

- $\sigma_e = 10.6258$
- $SF_e^{\text{plateau}} = 1.012$
- $SF_e^0 = 0.7513$

There are inherent difficulties associated with measuring the τ_h -lepton isolation efficiency due to the missing energy that occurs in hadronically decaying τ -leptons, thereby, making the reconstruction of the Z boson mass inaccurate. Furthermore, applying the “tag-and-probe” technique to the $Z \rightarrow \tau_h^+ \tau_h^-$ decay mode is not suitable since the method varies the number of jets and vertices in the events but at the same time these quantities are used as selection criteria to identify τ_h -leptons. For this reason, we use muons as “tag” and “probe” objects, where probe muons are required to match τ_h -lepton candidates to within 0.0001% with respect to p_T , η , and ϕ quantities. The probe muons are then required to pass various tau discriminates in order to determine the isolation efficiency in both data and MC simulation. Figure 6.2 (bottom) shows the τ_h -lepton isolation efficiencies when muons are used. We can then calculate scale factors and perform a fitting procedure using Equation 6.4. The resulting fit parameters for the τ_h -lepton isolation efficiency scale factor are,

- $\sigma_\tau = 12.5016$
- $SF_\tau^{\text{plateau}} = 0.999458$
- $SF_\tau^0 = 0.957157$

Systematic uncertainties on the lepton isolation and identification efficiencies are obtained by varying the number of jets and reconstructed vertices in the event, for example, selecting events with one jet, three jets, five to ten vertices, or fifteen to twenty vertices.

Changes in the fit parameters are taken as systematic uncertainties, as described with greater detail in Section 8.3.

6.4 b-tagging scale factor

Jets that arise from bottom-quark hadronization processes usually contain B hadrons and can be tagged as b jets, as previously discussed in Section 5.6.2. The b-tagging efficiency differs between data and simulation, and as a result scale factors are measured in order to correct for this effect. The b-tagging scale factor SF is defined as the ratio of the b-tagging efficiency measured in data to the measured b-tagging efficiency in simulation and is parametrized as a function of the jet p_T and η values.

We correct the simulation, using the instructions given in Reference [115], on an event by event basis in order to adjust the overall b-tagging efficiency of the sample. The CMS collaboration provides nominal fit functions for the weights and efficiencies of b-quark jets, where a b-quark jet is a reconstructed jet with an associated b quark at the generator level. We proceed to construct a scale factor for each event from the product of the weights of each jet in the event as follows. For b-quark jets that are tagged as a b jet, the scale factor for the event is given by,

$$SF = \prod_{i=1}^{N_{\text{jet}}} w_i^{\text{b-tagged}}, \text{ where } w_i^{\text{b-tagged}} = \frac{\epsilon_i^{\text{Data}}}{\epsilon_i^{\text{MC}}}, \quad (6.5)$$

whereas, for b-quark jets that are not tagged as a b jet, the scale factor is given by,

$$SF = \prod_{i=1}^{N_{\text{jet}}} w_i^{\text{not b-tagged}}, \text{ where } w_i^{\text{not b-tagged}} = \frac{1 - \epsilon_i^{\text{Data}}}{1 - \epsilon_i^{\text{MC}}} = \frac{1 - w_i^{\text{b-tagged}} \cdot \epsilon_i^{\text{MC}}}{1 - \epsilon_i^{\text{MC}}} \quad (6.6)$$

where, $w^{\text{b-tagged}}$ and ϵ^{MC} are supplied by CMS. If an event has both b-tagged and non-b-tagged jets then the overall scale factor is the product of the two expressions above. Systematic uncertainties related to the b-tagging scale factors are described in Section 8.3.

6.5 Event selection

We select events with at least three prompt and isolated leptons, where the threshold requirements mentioned previously are satisfied. Additionally, events must not have more than one τ_h -lepton. All events are also required to have fired at least one of the dilepton triggers listed in Appendix B. In order to eliminate low mass Drell-Yan processes or low mass resonances, such as J/ψ and Υ , we reject events that have an invariant mass $m_{\ell^+\ell^-}$ below 12 GeV, where $\ell = e, \mu$. The decay products of a Z boson may emit final-state-radiation, such that the photons may further undergo internal conversions to leptons. Therefore, to reduce background contributions from these SM processes events with $E_T^{\text{miss}} < 30$ GeV and $H_T < 200$ GeV that do not have a two body invariant mass $m_{\ell^+\ell^-}$ within the Z-boson mass region (i.e. 75 – 105 GeV) but do have a three body invariant mass $m_{\ell^+\ell^-\ell'^\pm}$ or $m_{\ell^+\ell^-\ell^\pm}$ within the Z-boson mass region are rejected, where ℓ' indicates a lepton flavor different from ℓ . Further details regarding photon conversions are found in Section 7.3.4. We define H_T , for the purposes of this analysis, as the scalar sum of the transverse momentum of all selected jets, given by $H_T = \sum_{\text{jets}} p_T^2$.

Chapter 7

Event Classification

In this chapter we present an overview for our inclusive analysis. The first section details the general search strategy of using an exclusive multichannel counting experiment. We categorize multilepton events based on their topological properties, which include the number of leptons and their flavors, lepton charges, the number of tagged b-jets. Additionally, kinematic properties of the event, such as H_T and E_T^{miss} , are also considered. The latter parts of this chapter describe the SM background estimations for the search channels using simulations and data-derived methods.

7.1 Search strategy

We perform a search for physics beyond the standard model without targeting any particular theory beforehand. This is referred to as an inclusive search analysis. The general search strategy of the analysis is to categorize multilepton events, with at least three charged leptons in the final state and where at most one of them is a τ_h candidate, into mutually exclusive search channels in order to increase sensitivity to new physics signals.

The amount of background from SM processes varies considerably across the search channels. In general, categories containing events with three leptons have higher background contributions than the ones containing events with four or more leptons. Therefore, we separate the low-background channels from the high-background channels in the interest of optimizing the overall sensitivity for new physics. For this reason, three lepton events are categorized independently from events with four or more leptons. To retain higher

sensitivity we also consider search channels with τ_h separately, since they have higher backgrounds due to τ_h misidentification. In addition, events with at least one tagged b-jet have higher background contributions from $t\bar{t}$ production, hence, events without tagged b-jets are categorized separately.

Multilepton channels that have events with opposite-sign, same-flavor (OSSF) lepton pairs suffer from larger background contributions than channels without such pairs, due to Drell-Yan production. We therefore gain signal sensitivity by classifying events in terms of the maximum number n of OSSF lepton pairs that can be formed using each lepton candidate once, designated by OSSF n . To illustrate, $\mu^+\mu^+e^-$ candidate events are categorized as OSSF0, while both $\mu^+\mu^-\mu^-$ and $\mu^+\mu^-e^-$ events, and $\mu^+\mu^-e^+e^-$ events are categorized as OSSF1 and OSSF2, respectively.

Additionally, events with any OSSF lepton pairings which have an invariant mass $m_{\ell+\ell^-}$ that lies in the Z-boson mass range, $75 < m_{\ell+\ell^-} < 105$ GeV, are classified as being “on-Z”, and “off-Z” otherwise. Since three lepton events have more sources of SM background than events with four or more leptons, we may again maximize signal sensitivity by defining, for three lepton events, “above-Z” and “below-Z” categories. For above-Z and below-Z events all possible OSSF lepton pairs have $m_{\ell+\ell^-} > 105$ GeV and $m_{\ell+\ell^-} < 75$ GeV, respectively. In the case where there are two possible OSSF lepton pairs, such that one is above-Z and the other below-Z, the pair with $m_{\ell+\ell^-}$ closer to the Z boson mass is used to classify the event. The reason for this separation is that there are more sources SM background below than above the Z-boson mass range.

Production of squarks and gluinos in SUSY events may exhibit a large H_T value in comparison to SM events, which motivates categorizing events on whether they have H_T greater than 200 GeV (high H_T), or less than 200 GeV (low H_T). Moreover, events are further sub-divided into several E_T^{miss} ranges: 0 – 50, 50 – 100, 100 – 150, 150 – 200, and above 200 GeV. These E_T^{miss} ranges are selected with the aim that SM processes should occupy the lower E_T^{miss} regions, while signals will lie in the higher E_T^{miss} spectrum.

7.2 Background estimation from simulation

An essential aspect of any search for new physics signals involves estimating the background contributions from SM processes for each search channel. SM processes with characteristics similar to that of the search signature, if not the exact same final states topology, are referred to as “irreducible backgrounds”. The reason for this is that there are no selection requirements which aid in reducing the background to any degree and improve signal sensitivity relative to the background. For these types of backgrounds the signal will be apparent only as an excess over SM predictions or by having a different shape in kinematic variables. An example of an irreducible background for a three lepton search channel is $WZ + \text{jets}$ production, where both electroweak bosons decay leptonically. Similarly, $ZZ + \text{jets}$ processes are irreducible backgrounds for the four or more lepton search channels since they may produce four prompt and isolated leptons. Irreducible backgrounds cannot be predicted directly from the data, therefore the only means to estimate them is by using simulations. Furthermore, backgrounds that can be reduced to a degree by applying some selection requirement are referred to as “reducible backgrounds”. The production of $t\bar{t} + \text{jets}$ events is an example of a reducible background since any leptons arising from the heavy-flavor jets is reduced by placing a requirement on the impact parameter of the lepton. We estimate reducible backgrounds with data-derived methods with the exception of $t\bar{t} + \text{jet}$ events for which simulation is used. Simulated events are validated in data control regions. The control regions are selected such that a particular SM process is enhanced, where scale factors can be extracted from data in order to correct simulated events. In the following sections we describe the estimates of background contributions from MC simulations.

7.2.1 Backgrounds from WZ and ZZ production

SM processes can produce three prompt and isolated leptons events with significant E_T^{miss} and H_T through $WZ + \text{jets}$ production, where the vector bosons decay leptonically (i.e. $W^\pm \rightarrow \ell^\pm \nu$ and $Z \rightarrow \ell^+ \ell^-$). As previously mentioned, this is an irreducible background for

three lepton search channels, which are estimated using simulation. In general, simulated events are adjusted to account for inaccuracies in their modeling of physical interactions, the detector simulation, or the tails of kinematic distributions. For this reason, we normalize the WZ cross section in a WZ-dominated data control region, by selecting tripleton events which have an OSSF lepton pair that is on-Z, no tagged b-jets, low H_T , and $50 < E_T^{\text{miss}} < 100$ GeV and scaling the transverse mass (M_T) distribution of the simulation to that of data, as shown in Figure 7.1 (left). The transverse mass is defined as $M_T = \sqrt{2 \cdot E_T^{\text{miss}} p_T^\ell (1 - \cos(\theta_{E_T^{\text{miss}} \ell}))}$, where θ is the angle between the E_T^{miss} vector and the lepton not belonging to the OSSF lepton pair [17]. This results in a 6% correction to the WZ production cross section. The corresponding WZ cross section is given in Table 4.3.

In addition, we validate the normalization on the cross section in a second WZ-dominated data control sample, shown in Figure 7.1 (right), by selecting tripleton events with an on-Z OSSF lepton pair, no tagged b-jets, low H_T , and $50 < M_T < 100$ GeV. The background source labeled in the figure by “Misidentified” refers to SM background from Drell–Yan processes, misidentified τ_h decays, and internal photon conversions, each of which will be described in Section 7.3. A comparison of data and simulation, after corrections are applied, shows good agreement.

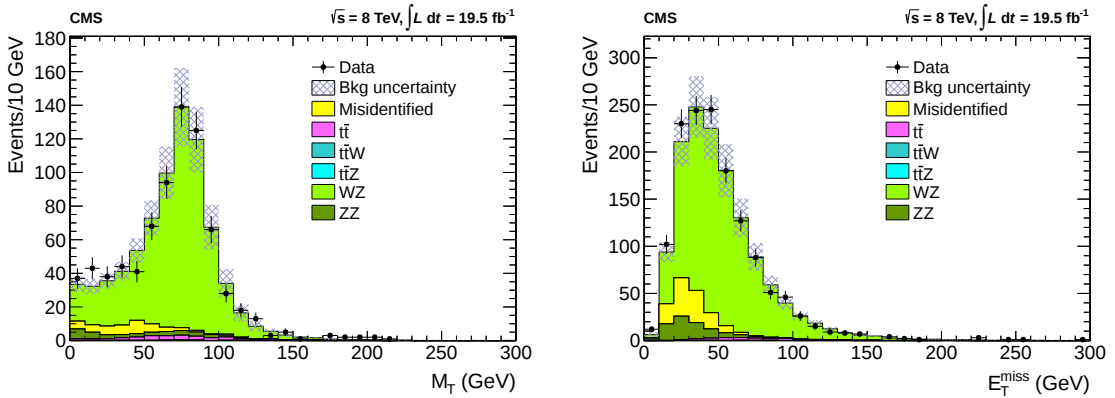


Figure 7.1: Distribution of M_T (left) and E_T^{miss} (right) in a data control region enriched with WZ events. We obtain a normalization scale factor from the M_T distribution and validate it in the E_T^{miss} distribution.

A source of irreducible background for the four or more lepton search channels is $ZZ + \text{jet}$

production. Similar to the WZ simulation, we normalize the ZZ cross section in a ZZ-dominated data control region, by selecting four lepton events which have two OSSF lepton pairs that are on-Z, no tagged b-jets, have low H_T , and $E_T^{\text{miss}} < 50$ GeV. The simulation is scaled to the data in the four lepton invariant mass, $m_{4\ell}$, distribution. This results in a 12% correction to the production cross section. The normalization is validated in a data control region similar to the previously described region with the exception that events are now required to have at least one on-Z OSSF lepton pair, as shown in Figure 7.2. The corresponding ZZ cross section is given in Table 4.3. There is overall good agreement between data and simulation.

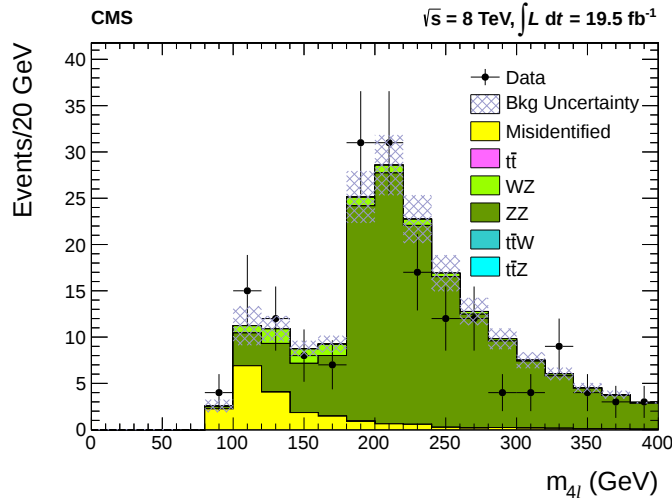


Figure 7.2: Distribution for the four lepton invariant mass in a data control region enriched with ZZ events.

Apart from the corrections to the overall normalization of the WZ and ZZ simulations, the backgrounds are also adjusted by the measured trigger and lepton efficiencies, b-tagging scale factors, jet multiplicity scale factors, as well as differences in the E_T^{miss} resolution. We investigate the E_T^{miss} distribution in a Z-enriched data control sample for several H_T ranges and different number of reconstructed vertices to correct the simulated E_T^{miss} resolution [116]. For a given H_T range and fixed number of vertices the components of E_T^{miss} in the x -direction and y -direction are found to be approximately Gaussian. Jet activity and

pileup both negatively impact the E_T^{miss} resolution. The effect of jet activity leads to systematically larger tails in the E_T^{miss} distribution due to mis-reconstruction, while the effect of pileup is stochastic, and therefore only affects the widths of the Gaussian distributions. We match the E_T^{miss} resolution of the simulation to that of data by applying correction factors to the Gaussian E_T^{miss} widths of the simulated events. These smearing factors can be as large as 25% for certain H_T ranges and number of vertices. The jet multiplicity scale factors are determined in a $WZ + \text{jets}$ enriched data control region with low H_T . They are measured by matching the distribution on the number of jets in data to that of simulation. The MC simulations are corrected on an event by event basis depending on the number of jets.

7.2.2 Background from $t\bar{t}$ production

A primary source of background for the three lepton search channels results from $t\bar{t}$ production. The top quark decays almost exclusively to a bottom quark and a W boson (i.e. $t \rightarrow bW$), which may subsequently decay leptonically. Therefore, $t\bar{t}$ processes may produce two leptons from the leptonic decay of the W bosons with a third lepton arising from the semileptonic decay of the b-jet daughter of one of the two top quarks. A prompt and isolated electron or muon originating from a jet is referred to as a “fake lepton” for the purposes of this analysis. We may reduce the background from $t\bar{t}$ processes by applying an impact parameter selection on the leptons in the event, which helps to eliminate most leptons arising from jets. For this reason, $t\bar{t}$ events are considered as a reducible background and are evaluated using simulation.

Even though the $t\bar{t}$ background involves mainly an electroweak process, and therefore, has a well defined jet flavor composition and p_T spectra, the simulation still needs to be validated against data. We define a $t\bar{t}$ -enriched data control region by selecting events with an opposite sign $e\mu$ lepton pair and at least one tagged b-jet for the purpose of validating the E_T^{miss} and H_T distributions since search channels are categorized based on

these kinematic variables. Figure 7.3 shows the distributions of E_T^{miss} and H_T for the data and simulation in a $t\bar{t}$ -enriched control sample, which have been corrected for trigger and lepton efficiencies, b-tagging scale factor, jet multiplicity scale factor, as well as for the E_T^{miss} resolution measurements. The jet multiplicity scale factors are determined in a similar manner to those obtained for the WZ simulation with the exception that we instead use a $t\bar{t} + \text{jets}$ enriched data control region. There is overall good agreement between data and simulation.

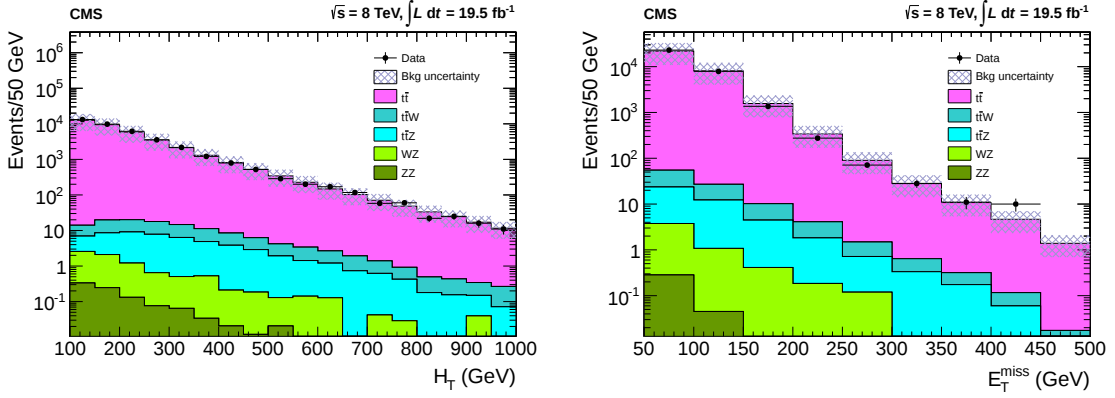


Figure 7.3: Distribution of H_T (left) and E_T^{miss} (right) in a data control region enriched in $t\bar{t}$ events.

In addition, we validate the relative isolation distribution of leptons originating from jets to ensure that the simulation properly models the isolation of fake leptons. We study the relative isolation distribution of non-prompt (i.e. large impact parameter) light-leptons from b jets of the $t\bar{t}$ decay in a single muon data control region, by selecting events containing exactly one prompt and isolated muon with $p_T > 30$ GeV, at least three jets, one of which must be b-tagged, and $S_T > 300$ GeV, where S_T is the sum of all selected lepton p_T , E_T^{miss} , and H_T . The isolation distribution of light-leptons from b jets is assumed to be independent of their impact parameter. Figure 7.4 shows the isolation distributions of non-prompt muons (left) and electrons (right). We extract a correction factor of 1.5 from the relative isolation distribution of non-prompt muons in the range of 0 – 0.15 in order for the simulation to match data. The isolation scale factor is validated in the relative isolation distribution of

non-prompt electrons to ensure that the simulation correctly predicts the rate for these leptons. The correction factor is applied to $t\bar{t}$ simulation when this source of background contributes to a search channel, requiring there to be a fake lepton. The corresponding $t\bar{t}$ cross section is given in Table 4.3.

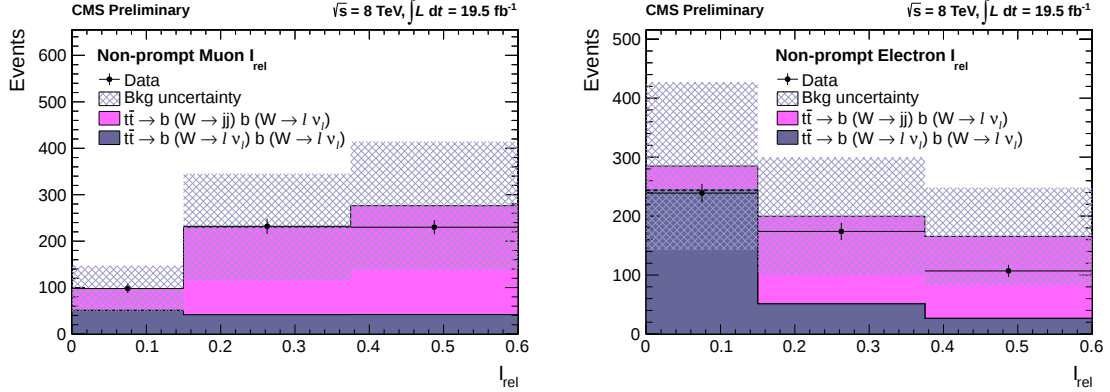


Figure 7.4: Distribution of relative isolation of non-prompt muons (left) and non-prompt electrons (right) in the single muon data control region enriched in $t\bar{t}$ events. A scale factor of 1.5 is applied to both distributions, which is obtained from measuring the discrepancy between simulation and data for the plot on the left in the relative isolation range of 0 – 0.15.

7.2.3 Background from rare standard model processes

Contributions from rare SM processes, such as $t\bar{t}W$, $t\bar{t}Z$, tbZ , and SM Higgs boson production from gluon-gluon fusion, vector boson fusion (VBF), or Higgs boson production in association with a W boson, Z boson, or $t\bar{t}$ pair, are estimated with the use of simulation. These simulations are corrected for trigger and lepton efficiencies, b-tagging scale factor, and E_T^{miss} resolution. The corresponding cross sections for these rare processes are given in Table 4.3.

7.3 Background estimation from data-derived methods

We estimate other sources of reducible background apart from $t\bar{t}$ processes, for example, $Z + \text{jets}$, $W^+W^- + \text{jets}$, and $Z + \gamma$ production, where the electroweak bosons decay leptonically, which contribute to multilepton events. The third lepton in these types of background

originate either from jets or photon conversions. Simulations do not always correctly model physics processes, as in the case, of leptons arising from jets, or kinematic properties of events, such as the tail of the E_T^{miss} distributions from $Z + \text{jet}$ production. For this reason, we rely on the use of data-derived methods to estimate SM background contributions, for the different search channels, from sources that are not properly modeled by simulations.

7.3.1 General principles of data-derived methods

We first discuss the general principles that the data-derived methods are based on followed by how they are applied to estimate background contributions. Collision data is used to determine the amount of SM background from reducible backgrounds, such as $Z + \text{jets}$ and $W^+W^- + \text{jets}$. We are mainly concerned with estimating the rate at which lepton candidates that originate from jets or photon conversions, or misidentified hadronic taus, are produced. These type of lepton candidates that pass selection requirements, e.g. promptness and isolation, are referred to as “fake leptons”. The general idea is that the amount of fake leptons in a given data sample can be determined by measuring the amount of other types of objects in the event, which are produced through similar physical processes and with similar kinematic properties. The reason for using these other physics objects is that they are produced in higher abundance than fake leptons. Additionally, the systematic uncertainties are simpler to evaluate since correlations are reduced. We refer to these other physics objects, which pass our selection requirements, with the exception of the identification selection, as “proxy objects”.

The main measurement of interest for the data-derived methods is the lepton fake rate, which is the ratio between the production rates of fake leptons to proxy objects. The fake rates are measured in data control regions and applied to search channels. We adjust the lepton fake rate to account for the discrepancies introduced when applying it in an environment different from where the measurement was performed. Therefore, the lepton fake rates are parametrized with respect to a different set of proxy objects in the event

that help to measure how the physical environment, such as jet p_T spectra, or jet flavor composition, differs in the search region in comparison to the data control region. We estimate the SM background contribution for the three lepton search channels from the product of the lepton fake rate and the number of events with two leptons and one proxy object, referred to as “seed events”. Similarly, to determine the background contributions for the channels with at least four leptons, we multiply the number of seed events that have two leptons and two proxy objects by the product of lepton fake rates, corresponding to the different types of proxy objects used. In general, we subtract background contributions estimated from simulations (i.e. $t\bar{t}$, WZ, ZZ, and rare SM production) from the total number of seed events before multiplying the result with the fake rate. We implement this in order to prevent double counting of background contributions between those obtained using simulation and data-derived methods.

The following sections discuss the measurement of lepton fake rates for electrons and muons originating from jets, jets misidentified as τ_h leptons, and electrons and muons arising from asymmetric internal photon conversions.

7.3.2 Backgrounds from misidentified electrons and muons from jets

An important source of background that involves fake light-leptons are electrons and muons arising from jets that undergo heavy-flavor quark decays, which are reduced by isolation and vertex requirements. The character of these background differs significantly from that of $t\bar{t}$ + jet events because they typically have a softer jet p_T spectrum. Moreover, the largest contributions for trilepton events are from Z + jets processes with low E_T^{miss} and low H_T , where the Z boson decays leptonically and the third lepton candidate is a real lepton from heavy-flavor decays in jets, or to a lesser extent a misidentified hadron. Another important source of background for trilepton events are from W^+W^- + jets processes. We assume that misidentifying three leptons from QCD processes and backgrounds from cosmic muons are negligible, while backgrounds from beam-halo muons are included in our background

estimations. There are many factors that affect the rate at which jets can produce lepton candidates, such as the jet p_T spectrum, jet flavor, jet shape, and form factors. Simulations do not properly model this background. Therefore, we rely on data-derived methods to estimate the background contributions due to fake light-leptons originating from jets.

A suitable proxy object for fake light-leptons are isolated tracks from pions and kaons since jets produce tracks at higher rates, thereby providing a larger statistical data sample. The tracks are required to pass the same selections as the light-leptons with the exception of lepton identification. The light-lepton fake rate is therefore defined as the ratio of number of prompt and isolated leptons to the number of prompt and isolated tracks [111], given by,

$$f_\ell = \frac{N_\ell^{\text{Iso}}}{N_{\text{track}}^{\text{Iso}}}, \quad (7.1)$$

where ℓ represents electrons and muons.

Using the isolation efficiencies $\epsilon_\ell^{\text{Iso}} = N_\ell^{\text{Iso}}/N_\ell^{\text{Non-iso}}$ and $\epsilon_{\text{track}}^{\text{Iso}} = N_{\text{track}}^{\text{Iso}}/N_{\text{track}}^{\text{Non-iso}}$, we reformulate the light-lepton fake rate as,

$$f_\ell = \frac{N_\ell^{\text{Non-iso}}}{N_{\text{track}}^{\text{Non-iso}}} \cdot \frac{\epsilon_\ell^{\text{Iso}}}{\epsilon_{\text{track}}^{\text{Iso}}}, \quad (7.2)$$

where $N_\ell^{\text{Non-iso}}$ corresponds to the number of prompt and non-isolated leptons and $N_{\text{track}}^{\text{Non-iso}}$ corresponds to the number of prompt and non-isolated tracks.

The proportionality factor f_ℓ between isolated leptons and tracks from jets depends on the heavy-flavor content of the data sample. In order to understand the changes in the jet composition, we use the fact that tracks from jets with heavy-flavor quark decays have larger impact parameters than tracks produced from light-flavor jets due to displaced vertices. For this reason, the average impact parameter of non-isolated tracks provides a way to characterize the heavy-flavor content of a data sample. Accordingly, the parameter R_{dxy} is defined as the ratio of the number of non-isolated tracks $N_{\text{track}}^{\text{Non-prompt}}$ with large impact parameter, $|d_{xy}(PV)| > 0.02 \text{ cm}$, to the number of non-isolated tracks $N_{\text{track}}^{\text{Prompt}}$ with

small impact parameter, $|d_{xy}(PV)| < 0.02 \text{ cm}$, which is given in Equation 7.4. We measure R_{dxy} using tracks with $p_T > 8 \text{ GeV}$, $|\eta| < 2.4$, that are not within ΔR of 0.1 of selected electrons and muons, and ΔR of 0.3 of selected τ_h -leptons. Additionally, tracks originating from the leading p_T jet in an event are rejected. Collision data which has mostly light-flavor jets have R_{dxy} between 2 – 3%, while a data sample with mainly tagged b-jets has R_{dxy} around 20 – 30%.

We parameterize the ratio of isolation efficiencies in terms of R_{dxy} using an analytical expression given by Equation 7.15, the derivation of which is described in the proceeding section. Therefore, the light-lepton fake rate is defined by,

$$f_\ell = \frac{N_\ell^{\text{Non-iso}}}{N_{\text{track}}^{\text{Non-iso}}} \cdot \epsilon_{\text{ratio}}^{\text{Iso}}(R_{\text{dxy}}), \quad (7.3)$$

where $N_\ell^{\text{Non-iso}}$, $N_{\text{track}}^{\text{Non-iso}}$, $N_{\text{track}}^{\text{Non-prompt}}$, and $N_{\text{track}}^{\text{Non-prompt}}$ all serve as additional proxy objects for this data-derived method. Moreover, background contributions from simulations are subtracted from these quantities to avoid double counting of background predictions.

The background estimate for a three light-lepton search channel due to fake light-leptons is calculated from the product of the fake rate with the number of seed events consisting of two isolated light-leptons and an isolated track. The fake rate for the three lepton search region is estimated using Equations 7.3 and 7.15. The number of prompt and non-isolated leptons and tracks, as well as the value of R_{dxy} are measured in the trilepton search region.

Derivation of the efficiency ratio as a function of R_{dxy}

In this section, we present the derivation of the analytical expression that relates the isolation efficiency ratio, ϵ_{ratio} , to the parameter R_{dxy} . We define, for a pure data sample, the parameter R_{dxy} as the ratio between the fraction of non-prompt and non-isolated tracks to

the fraction of prompt and non-isolated tracks, as,

$$R_{\text{dxy}} \equiv \frac{\frac{N_{\text{track}}^{\text{Non-prompt}}}{N_{\text{track}}^{\text{Non-prompt}} + N_{\text{track}}^{\text{Prompt}}}}{\frac{N_{\text{track}}^{\text{Prompt}}}{N_{\text{track}}^{\text{Non-prompt}} + N_{\text{track}}^{\text{Prompt}}}}, \quad (7.4)$$

where $N_{\text{track}}^{\text{Non-prompt}}$ refers to the number of non-isolated tracks with $d_{\text{xy}} > 0.02 \text{ cm}$ and $N_{\text{track}}^{\text{Prompt}}$ refers to the number of non-isolated tracks with $d_{\text{xy}} < 0.02 \text{ cm}$.

In the case of a mixed data sample, we define R_{dxy} by combining two different data samples in terms of a free parameter α , which may vary between 0 – 1. We denote two different pure data samples with the superscripts “a” and “b”.

$$R_{\text{dxy}}(\alpha) \equiv \frac{\alpha \cdot \frac{N_{\text{track}}^{\text{Non-prompt, a}}}{N_{\text{track}}^{\text{Non-prompt, a}} + N_{\text{track}}^{\text{Prompt, a}}} + (1 - \alpha) \cdot \frac{N_{\text{track}}^{\text{Non-prompt, b}}}{N_{\text{track}}^{\text{Non-prompt, b}} + N_{\text{track}}^{\text{Prompt, b}}}}{\alpha \cdot \frac{N_{\text{track}}^{\text{Prompt, a}}}{N_{\text{track}}^{\text{Non-prompt, a}} + N_{\text{track}}^{\text{Prompt, a}}} + (1 - \alpha) \cdot \frac{N_{\text{track}}^{\text{Prompt, b}}}{N_{\text{track}}^{\text{Non-prompt, b}} + N_{\text{track}}^{\text{Prompt, b}}}}. \quad (7.5)$$

Solving for the free parameter α gives,

$$\alpha(R_{\text{dxy}}) = \frac{1}{1 - \frac{R_{\text{dxy}} - R_{\text{dxy}}^{\text{a}}}{R_{\text{dxy}} - R_{\text{dxy}}^{\text{b}}} \cdot \frac{1 + R_{\text{dxy}}^{\text{b}}}{1 + R_{\text{dxy}}^{\text{a}}}}. \quad (7.6)$$

This relationship will be used later in the derivation of the efficiency ratio.

The isolation efficiency for leptons in a pure data sample can be measured using the following relation,

$$\epsilon_{\ell}^{\text{Iso}} \equiv \frac{\frac{N_{\ell}^{\text{Iso}}}{N_{\ell}^{\text{Non-iso}} + N_{\ell}^{\text{Iso}}}}{\frac{N_{\ell}^{\text{Non-iso}}}{N_{\ell}^{\text{Non-iso}} + N_{\ell}^{\text{Iso}}}}, \quad (7.7)$$

where N_{ℓ}^{Iso} is the number of prompt and isolated leptons and $N_{\ell}^{\text{Non-iso}}$ is the number of prompt and non-isolated leptons. Similarly for tracks,

$$\epsilon_{\text{track}}^{\text{Iso}} \equiv \frac{\frac{N_{\text{track}}^{\text{Iso}}}{N_{\text{track}}^{\text{Non-iso}} + N_{\text{track}}^{\text{Iso}}}}{\frac{N_{\text{track}}^{\text{Non-iso}}}{N_{\text{track}}^{\text{Non-iso}} + N_{\text{track}}^{\text{Iso}}}}, \quad (7.8)$$

where $N_{\text{track}}^{\text{Iso}}$ is the number of prompt and isolated tracks and $N_{\text{track}}^{\text{Non-iso}}$ is the number of prompt and non-isolated tracks.

We define the lepton isolation efficiency, for a mixed data sample, similar to how R_{dxy} was constructed,

$$\epsilon_{\ell}^{\text{Iso}}(\alpha) \equiv \frac{\alpha \cdot \frac{N_{\ell}^{\text{Iso, a}}}{N_{\ell}^{\text{Non-iso, a}} + N_{\ell}^{\text{Iso, a}}} + (1 - \alpha) \cdot \frac{N_{\ell}^{\text{Iso, b}}}{N_{\ell}^{\text{Non-iso, b}} + N_{\ell}^{\text{Iso, b}}}}{\alpha \cdot \frac{N_{\ell}^{\text{Non-iso, a}}}{N_{\ell}^{\text{Non-iso, a}} + N_{\ell}^{\text{Iso, a}}} + (1 - \alpha) \cdot \frac{N_{\ell}^{\text{Non-iso, b}}}{N_{\ell}^{\text{Non-iso, b}} + N_{\ell}^{\text{Iso, b}}}}, \quad (7.9)$$

$$= \frac{\epsilon_{\ell}^{\text{Iso, b}} \cdot (1 + \epsilon_{\ell}^{\text{Iso, a}}) + \alpha \cdot (\epsilon_{\ell}^{\text{Iso, a}} - \epsilon_{\ell}^{\text{Iso, b}})}{(1 + \epsilon_{\ell}^{\text{Iso, a}}) - \alpha \cdot (\epsilon_{\ell}^{\text{Iso, a}} - \epsilon_{\ell}^{\text{Iso, b}})}, \quad (7.10)$$

where we have re-expressed $\epsilon_{\ell}^{\text{Iso}}$ in terms of the relative isolation efficiencies, $\epsilon_{\ell}^{\text{Iso, a}}$ and $\epsilon_{\ell}^{\text{Iso, b}}$. Similarly for tracks, we have,

$$\epsilon_{\text{track}}^{\text{Iso}}(\alpha) = \frac{\alpha \cdot \frac{N_{\text{track}}^{\text{Iso, a}}}{N_{\text{track}}^{\text{Non-iso, a}} + N_{\text{track}}^{\text{Iso, a}}} + (1 - \alpha) \cdot \frac{N_{\text{track}}^{\text{Iso, b}}}{N_{\text{track}}^{\text{Non-iso, b}} + N_{\text{track}}^{\text{Iso, b}}}}{\alpha \cdot \frac{N_{\text{track}}^{\text{Non-iso, a}}}{N_{\text{track}}^{\text{Non-iso, a}} + N_{\text{track}}^{\text{Iso, a}}} + (1 - \alpha) \cdot \frac{N_{\text{track}}^{\text{Non-iso, b}}}{N_{\text{track}}^{\text{Non-iso, b}} + N_{\text{track}}^{\text{Iso, b}}}}, \quad (7.11)$$

$$= \frac{\epsilon_{\text{track}}^{\text{Iso, b}} \cdot (1 + \epsilon_{\text{track}}^{\text{Iso, a}}) + \alpha \cdot (\epsilon_{\text{track}}^{\text{Iso, a}} - \epsilon_{\text{track}}^{\text{Iso, b}})}{(1 + \epsilon_{\text{track}}^{\text{Iso, a}}) - \alpha \cdot (\epsilon_{\text{track}}^{\text{Iso, a}} - \epsilon_{\text{track}}^{\text{Iso, b}})}. \quad (7.12)$$

The isolation efficiency ratio is the ratio between the lepton isolation efficiency and the track isolation efficiency, and is given by,

$$\epsilon_{\text{ratio}} \equiv \frac{\epsilon_{\ell}^{\text{Iso}}}{\epsilon_{\text{track}}^{\text{Iso}}}. \quad (7.13)$$

Therefore, the isolation efficiency ratio as a function of α is given by,

$$\epsilon_{\text{ratio}}(\alpha) = \frac{\frac{\alpha^{-1} \cdot \epsilon_{\ell}^{\text{Iso, b}} \cdot (1 + \epsilon_{\ell}^{\text{Iso, a}}) + (\epsilon_{\ell}^{\text{Iso, a}} - \epsilon_{\ell}^{\text{Iso, b}})}{\alpha^{-1} \cdot (1 + \epsilon_{\ell}^{\text{Iso, a}}) - (\epsilon_{\ell}^{\text{Iso, a}} - \epsilon_{\ell}^{\text{Iso, b}})}}{\frac{\alpha^{-1} \cdot \epsilon_{\text{track}}^{\text{Iso, b}} \cdot (1 + \epsilon_{\text{track}}^{\text{Iso, a}}) + (\epsilon_{\text{track}}^{\text{Iso, a}} - \epsilon_{\text{track}}^{\text{Iso, b}})}{\alpha^{-1} \cdot (1 + \epsilon_{\text{track}}^{\text{Iso, a}}) - (\epsilon_{\text{track}}^{\text{Iso, a}} - \epsilon_{\text{track}}^{\text{Iso, b}})}}. \quad (7.14)$$

Lastly, replacing α with $\alpha(R_{\text{dxy}})$ results in the analytical expression of the efficiency

ratio as a function of R_{dxy} , and is given by,

$$\epsilon_{\text{ratio}}(R_{\text{dxy}}) = \frac{\left(1 - \frac{R_{\text{dxy}} - R_{\text{dxy}}^{\text{a}}}{R_{\text{dxy}} - R_{\text{dxy}}^{\text{b}}} \cdot \frac{1 + R_{\text{dxy}}^{\text{b}}}{1 + R_{\text{dxy}}^{\text{a}}}\right) \cdot \epsilon_{\ell}^{\text{Iso, b}} \cdot (1 + \epsilon_{\ell}^{\text{Iso, a}}) + (\epsilon_{\ell}^{\text{Iso, a}} - \epsilon_{\ell}^{\text{Iso, b}})}{\left(1 - \frac{R_{\text{dxy}} - R_{\text{dxy}}^{\text{a}}}{R_{\text{dxy}} - R_{\text{dxy}}^{\text{b}}} \cdot \frac{1 + R_{\text{dxy}}^{\text{b}}}{1 + R_{\text{dxy}}^{\text{a}}}\right) \cdot (1 + \epsilon_{\ell}^{\text{Iso, a}}) - (\epsilon_{\ell}^{\text{Iso, a}} - \epsilon_{\ell}^{\text{Iso, b}})} \cdot \frac{\left(1 - \frac{R_{\text{dxy}} - R_{\text{dxy}}^{\text{a}}}{R_{\text{dxy}} - R_{\text{dxy}}^{\text{b}}} \cdot \frac{1 + R_{\text{dxy}}^{\text{b}}}{1 + R_{\text{dxy}}^{\text{a}}}\right) \cdot \epsilon_{\text{track}}^{\text{Iso, b}} \cdot (1 + \epsilon_{\text{track}}^{\text{Iso, a}}) + (\epsilon_{\text{track}}^{\text{Iso, a}} - \epsilon_{\text{track}}^{\text{Iso, b}})}{\left(1 - \frac{R_{\text{dxy}} - R_{\text{dxy}}^{\text{a}}}{R_{\text{dxy}} - R_{\text{dxy}}^{\text{b}}} \cdot \frac{1 + R_{\text{dxy}}^{\text{b}}}{1 + R_{\text{dxy}}^{\text{a}}}\right) \cdot (1 + \epsilon_{\text{track}}^{\text{Iso, a}}) - (\epsilon_{\text{track}}^{\text{Iso, a}} - \epsilon_{\text{track}}^{\text{Iso, b}})} \quad (7.15)$$

For the purposes of this analysis, we use a $Z + \text{jets}$ enriched (i.e. no tagged b-jets) data control region as our “a” data sample and a $t\bar{t}$ enriched (i.e. at least one tagged b-jet) data control region as our “b” data sample [117, 118, 119].

Figure 7.5 shows the efficiency ratio as a function of R_{dxy} . The left end of the curve represents a $Z + \text{jets}$ enriched data region while the right end of the curve represents a $t\bar{t}$ enriched data region.

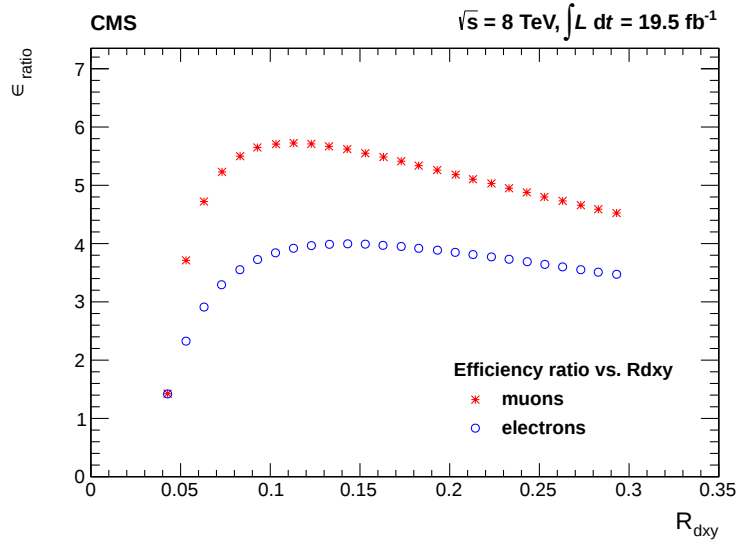


Figure 7.5: Efficiency ratio vs R_{dxy} .

In a data sample dominated by $Z(\rightarrow e^+e^-) + \text{jets}$ events, we determine f_{μ} to be 0.6% for the background from misidentified muon candidates. Similarly, in a data sample of $Z(\rightarrow \mu^+\mu^-) + \text{jets}$ events, we find f_e to be 0.7% for the background from misidentified electron candidates. Equation 7.15 is validated in a $Z + \text{b-jets}$ data control sample, corresponding

to the region of the plot where the slope of the curve turns. A 30% systematic uncertainty is associated with the light-lepton fake rate on the number of background events.

7.3.3 Background from jets misidentified as τ_h leptons

The τ -lepton, as the heaviest lepton, decays approximately a third of the time to a light-lepton and a pair of neutrinos and two-thirds of the time to various hadronic final states. Since hadronically decaying τ -leptons cannot be accurately identified, in contrast to electrons and muons, without the use of isolation requirements, search channels with τ_h candidates experience larger background contributions. Therefore, the primary source of τ_h backgrounds are from misidentified jets since sufficiently isolated jets tend to resemble hadronically decaying τ -leptons. We rely on a data-derived technique to measure the background due to misidentified jets.

A tight-loose isolation method is used to estimate the background from fake τ_h -leptons in data [111]. We define three isolation regions, specifically the “isolation region” (Iso) with $E_\tau^{\text{iso}} < 2.0$ GeV, the “sideband region” (SB) with $6.0 < E_\tau^{\text{iso}} < 15.0$ GeV, and the “others region”, where $E_\tau^{\text{iso}} > 15.0$ GeV. The hadronic tau fake rate is given by the ratio between the number of isolated τ_h candidates to the number of sideband τ_h candidates, more explicitly,

$$f_\tau = \frac{N_\tau^{\text{Iso}}}{N_\tau^{\text{SB}}}, \quad (7.16)$$

where sideband τ_h candidates are used as proxy objects in the data-derived method.

The isolation distribution changes drastically with respect to jet multiplicity and jet p_T spectrum and, therefore, the fake rate f_τ can vary greatly from one data sample to another. Consequently, we define an additional quantity that parametrizes the amount of jet activity

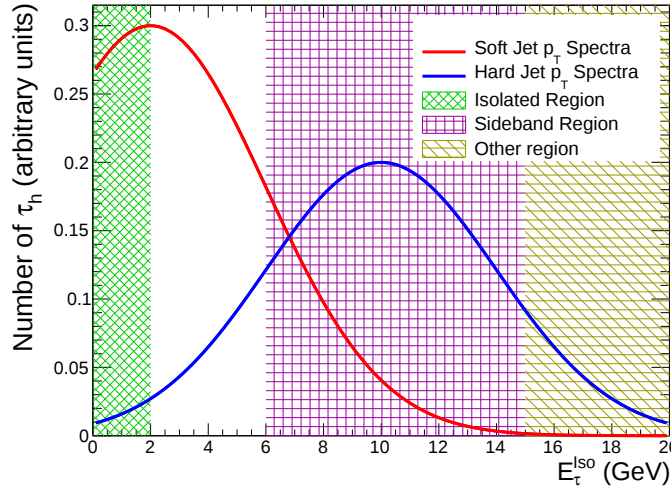


Figure 7.6: Illustration of isolation distribution for τ_h candidates in two different jet p_T spectra.

in data based on an isolation sideband method,

$$f_{\text{SB}} = \frac{N_{\tau}^{\text{SB}}}{N_{\tau}^{\text{SB}} + N_{\tau}^{\text{Other}}}, \quad (7.17)$$

where the ratio represents the number of sideband τ_h candidates to the total number of non-isolated τ_h candidates. For a data sample with a soft jet p_T spectra both f_{τ} and f_{SB} will have very large values in comparison to a data sample with a hard jet p_T spectrum. Figure 7.6 shows an illustration of the isolation distribution for τ_h candidates in two different jet p_T spectra, while Figure 7.7 shows the isolation distribution of τ_h candidates for various $\sum p_T^{\text{track}}$ ranges in data.

We measure the hadronic tau fake rate, f_{τ} , and the parameter f_{SB} in a data control region where there are no genuine τ_h leptons. Accordingly, a data sample enriched in $Z + \text{jet}$ events, such that $Z \rightarrow e^+e^-$ or $\mu^+\mu^-$, is used by selecting events with exactly one on-Z OSSF lepton pair, low E_T^{miss} , and low H_T , defined by $E_T^{\text{miss}} < 50 \text{ GeV}$ and $H_T < 200 \text{ GeV}$. The control region is further subdivided into different $\sum p_T^{\text{track}}$ ranges, that is the sum of all track p_T associated with the primary vertex of the event. In each subregion we measure f_{τ} and f_{SB} and plot their distribution, given in Figure 7.8. We perform a fit on the distribution

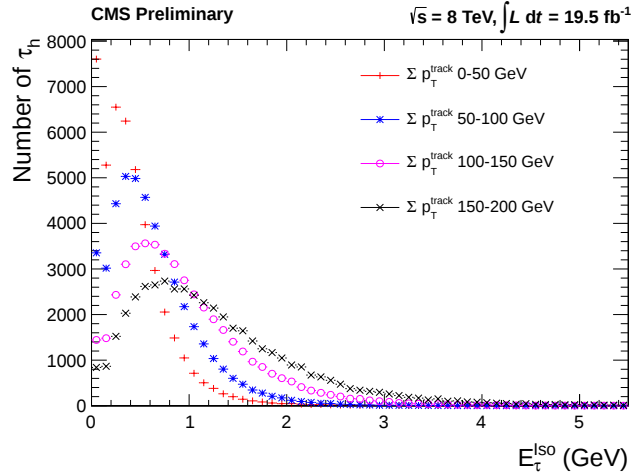


Figure 7.7: Distribution of isolation for τ_h candidates for different ranges of $\sum p_T^{\text{track}}$.

in Figure 7.8 to determine an empirical relationship between f_τ and f_{SB} since there are no analytical expressions that can be derived to relate them. We measure the fake rate f_τ to be 20%.

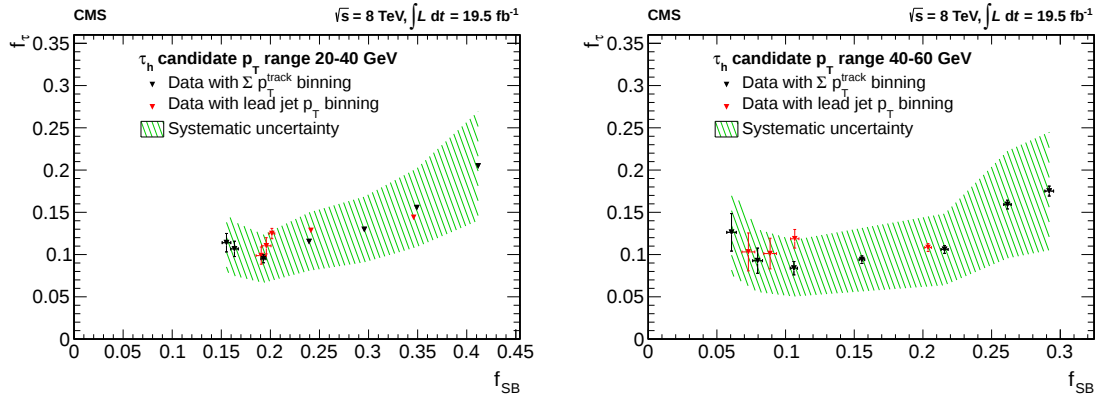


Figure 7.8: Distribution of f_τ versus f_{SB} for τ_h candidates with p_T between 20 – 40 GeV (left) and 40 – 60 GeV (right) in a data control region, with exactly one on-Z OSSF lepton pair, low E_T^{miss} , and low H_T .

The background estimate for a three lepton search channel with a τ_h from fake τ_h backgrounds is determined from multiplying the fake rate by the number of seed events that have two light-leptons and a sideband τ_h candidate. The fake rate that is used in the estimate is calculated by evaluating the fit function of f_τ at f_{SB} , where f_{SB} is measured by counting the number of sideband τ_h and non-isolated τ_h in the three lepton search region.

A 30% systematic uncertainty is associated with the hadronic tau fake rate on the number of background events. This follows a similar procedure as the background estimations from electrons and muons arising from jets described in Section 7.3.2.

7.3.4 Background from asymmetric internal photon conversions

A possible source of background from $W\gamma^*$ production, where $\gamma^* \rightarrow \ell^\pm(\ell^\mp)$ represents an off-shell photon conversion, was not initially considered in Higgs boson searches for the $H \rightarrow WW^* \rightarrow \ell\nu\ell'\nu$ channel [120]. The parenthesis used in the decay notation indicates that the off-shell photon produces a pair of leptons with a large p_T difference, which is referred to as asymmetric conversion. Background sources which produce photon conversions affect multilepton searches as well. There are two different types of photon conversions, namely “external” and “internal”. In external conversions an on-shell photon radiates in the presence of the external magnetic field of the detector or by interacting with the material in the detector producing an $\ell^+\ell^-$ pair. Such conversions produce mainly e^+e^- pairs in comparison to $\mu^+\mu^-$ pairs since electrons are much lighter than muons. The rate of external conversions to e^+e^- pairs is several orders of magnitude larger than to $\mu^+\mu^-$ pairs, but due to the electron identification requirements the background from external conversions is greatly reduced. In internal conversions of an off-shell photon γ^* undergoes conversion in vacuum to an $\ell^+\ell^-$ pair, where electron or muon pairs are produced in equal amounts.

We are interested in estimating backgrounds from two different cases of asymmetric internal photon conversion that contribute to the multilepton search channels. The first case, which is accurately modeled by simulations, concerns photon conversions where both leptons from the decay are reconstructed and pass lepton selection requirements. In the second case, the simulation fails to properly model the event rate of photon conversions where one of the leptons is not reconstructed or does not pass lepton selections. There is a minimum lepton p_T threshold set in simulations that prevents production of leptons with low enough p_T to fail reconstruction or lepton selections. Therefore, we rely on data-derived

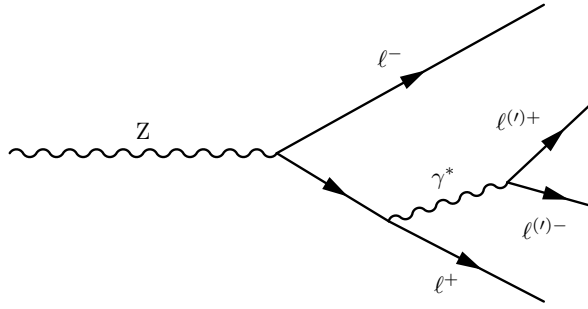


Figure 7.9: A Feynman diagram of a Z boson decaying to an OSSF lepton pair, where one of the leptons undergoes asymmetric internal photon conversion into a pair of leptons such that one of them fails to be properly reconstructed or does not pass lepton selection requirements. The label ℓ' indicates that the leptons from the photon conversion can be of the same flavor, but not necessarily, as those leptons from the Z boson decay.

methods to estimate the background from these types of asymmetric conversions due to the absence of any simulation that could adequately model them. An example of this process is shown by the Feynman digram in Figure 7.9.

The production rate of off-shell photons, which yield asymmetric conversions, is assumed to be proportional to the rate for the production of on-shell photons. The asymmetric internal photon conversion fake rate is therefore measured in a data control region, by selecting events with low E_T^{miss} and low H_T , defined by $E_T^{\text{miss}} < 30 \text{ GeV}$ and $H_T < 200 \text{ GeV}$ [111]. The fake rate is defined by,

$$C_\ell = \frac{N_{\ell^+\ell^-\ell'^{\pm}}}{N_{\ell^+\ell^-\gamma}}, \quad (7.18)$$

where $N_{\ell^+\ell^-\ell'^{\pm}}$ is the number of events with $|m_{\ell^+\ell^-\ell'^{\pm}} - m_Z| < 15 \text{ GeV}$ and $N_{\ell^+\ell^-\gamma}$ those with $|m_{\ell^+\ell^-\gamma} - m_Z| < 15 \text{ GeV}$. The label ℓ' indicates that the leptons from the photon conversion can be of the same flavor, but not necessarily, as those leptons from the Z boson decay. All possible OSSF lepton pairs are required to have an invariant mass that is off-Z. We use on-shell photons as proxy objects in the data-derived method to estimate the background contributions for our objects of interest, which are leptons originating from

photon conversions.

The fake rate for electrons C_e is measured to be $(2.0 \pm 0.3)\%$ and for muons C_μ is $(0.7 \pm 0.1)\%$, where the uncertainty is only statistical. We multiply these fake rates by the number of seed events with $\ell^+\ell^-\gamma$ to determine the amount of background events in the different search channels, with a systematic uncertainty of 50%. Figure 7.10 show the three-body μ invariant mass distribution for a tri- μ control region, similar to the one previously mentioned. The off-Z mass ranges serve to validate the fake rate.

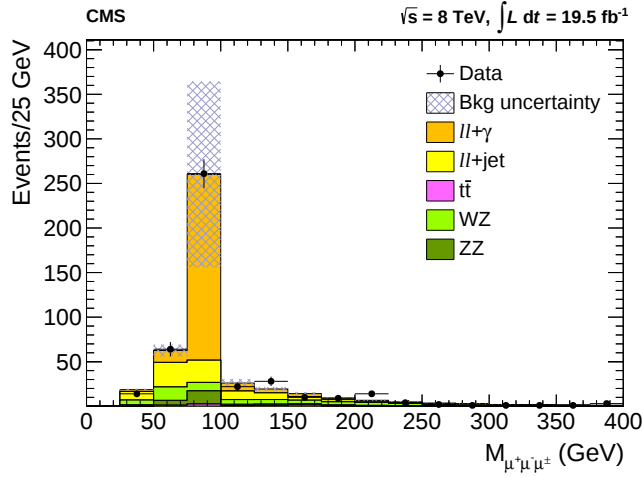


Figure 7.10: Distribution for a three-body μ invariant mass in a tri- μ control region.

In addition, in order to reduce backgrounds from $Z \rightarrow \ell^\pm(\ell^\mp \rightarrow \ell^\mp \gamma^*)$ processes we exclude events with three leptons that have one OSSF lepton pair, low E_T^{miss} , and low H_T , such that all possible OSSF lepton pairs are off-Z but the three-body invariant mass is on-Z. We still have background contributions from other sources of asymmetric internal photon conversions, such as from $W^+W^-\gamma^*$ production.

Chapter 8

Experimental Uncertainties

In this chapter we present several sources of uncertainties that affect the analysis. We discuss in the following sections the systematic uncertainties that are taken into account before the limit setting procedure can be performed. The effects of these uncertainties on background and signal efficiencies vary from less than a few percent to as high as 30% in certain regions of phase space.

8.1 An overview on experimental uncertainties

The results from experiments depend on both the accuracy and precision of the measurements, therefore, experimental uncertainties are by their nature inexact. The lower the accuracy and precision of a measurement the larger its uncertainty. Uncertainties are categorized as either statistical or systematic. Statistical uncertainties result from stochastic fluctuations arising from measurements based on a finite set of observations. These uncertainties may follow a Poisson distribution with a mean λ and a standard deviation σ given by $\sigma = \sqrt{N}$, where N correspond to the number of events for a counting experiment. Besides statistical uncertainties there are also systematic uncertainties that need considerations to properly determine the uncertainties associated with the measurement. These systematic uncertainties arise from biases in measurements or inaccurate modeling of the observables used in the measurements. The influence of these types of uncertainties on the analysis can be studied by shifting a quantity within its uncertainty and, afterwards, determine their correlation and effects on the final results.

Signal models and background simulations are subject to various sources of uncertainties which include the integrated luminosity measurement uncertainty, cross section uncertainty, trigger and lepton efficiency uncertainties, E_T^{miss} resolution uncertainty, jet energy scale uncertainty, and b-tagging scale factor uncertainty. In addition, signal efficiencies are affected by uncertainties due to ISR modeling [121]. Moreover, there are systematic uncertainties identified with the light-lepton, hadronic tau, and asymmetric internal photon conversion fake rates that effect background estimations from data-derived methods. In the following sections we discuss how these systematic uncertainties, which are relevant for this analysis, are taken into account.

8.2 Sources of systematic uncertainties from data-derived methods

In this section, we discuss the evaluation of the systematic uncertainties associated with the various data-derived methods applied in the background estimations. In general, the uncertainties on the background predictions, obtained using light-lepton, hadronic tau, and asymmetric internal photon conversion fake rates, depend on the accuracy of proxy object approach.

We measure the muon fake rate f_μ using a $Z(\rightarrow e^+e^-) + \text{jets}$ data control region from the $Z + \text{jets}$ sample described in Section 7.3.2, such that the muon has a large probability of originating from a jet. The fake rate is measured again in a $Z(\rightarrow \mu^+\mu^-) + \text{jets}$ sample. In principle, the muon fake rate should not depend on the decay of the Z boson, and consequently the difference between the two measurements is taken as a systematic uncertainty on the fake rate. Accordingly, a systematic uncertainty of 30% is assigned to the background estimates from misidentified muon candidates. The assumption that the average quantity $\left\langle \frac{N_\ell^{\text{Non-iso}}}{N_{\text{track}}^{\text{Non-iso}}} \cdot \frac{\epsilon_\ell^{\text{Iso}}}{\epsilon_{\text{track}}^{\text{Iso}}} \right\rangle = \left\langle \frac{N_\ell^{\text{Non-iso}}}{N_{\text{track}}^{\text{Non-iso}}} \right\rangle \cdot \left\langle \frac{\epsilon_\ell^{\text{Iso}}}{\epsilon_{\text{track}}^{\text{Iso}}} \right\rangle$ is also taken into account and included in the systematic uncertainty. Similarly, the systematic uncertainty on background estimates from misidentified electron candidates is evaluated to be 30% as well.

The evaluation of the systematic uncertainty on the background estimates from fake

τ_h candidates relies on an alternative subdivision of the $Z + \text{jets}$ data control region than the $\sum p_T^{\text{track}}$ ranges described in Section 7.3.3. We divide the same control region into different $p_T^{\text{lead jet}}$ ranges, and repeat the fitting procedure on the f_τ versus f_{SB} distribution. A systematic uncertainty of 30% is assigned to the background estimates from misidentified τ_h candidates to cover the variation between the two hadronic tau fake rate fits, as shown by the green bands in Figure 7.8.

In order to evaluate the systematic uncertainty on the background estimates from asymmetric internal photon conversions, we subdivide the Z -enriched sample described in Section 7.3.4. We measure the muon fake rate C_μ in a $Z(\rightarrow e^+e^-) + \mu^\pm$ and $Z(\rightarrow \mu^+\mu^-) + \mu^\pm$ data control region. The difference between the two fake rates is assigned as a systematic uncertainty, which also includes the uncertainty associated with our underlying assumption of the proportionality between events with off-shell and on-shell photons. The systematic uncertainty on the electron fake rate C_e is determined in a similar manner. We assign an uncertainty of 50% on the background estimates due to misidentified muon and electron candidates from asymmetric internal photon conversions.

8.3 Sources of systematic uncertainties for simulations

In general, there are sources of systematic uncertainties that mainly affect data-derived methods, while others affect estimates from signal and background simulations. In this section, we deal with those involving the latter, such as integrated luminosity measurement uncertainty, cross section uncertainty, lepton trigger, identification, and isolation efficiency uncertainties, E_T^{miss} resolution uncertainty, jet energy scale uncertainty, and b-tagging scale factor uncertainty. A source of systematic uncertainty that specifically affects signal estimates is due to ISR modeling, described in a subsequent section.

Integrated luminosity measurement uncertainty

The luminosity is measured based on a pixel cluster technique, where the number of hits in the pixel detector is used as a measure to be proportional to the luminosity. A more detailed discussion of the luminosity measurement is found in Section 3.4. The absolute scale is calibrated using the Van Der Meer Scan method with systematic uncertainty measurements of 2.2% for 7 TeV and 2.5% for 8 TeV. A summary of the sources of uncertainties related the integrated luminosity measurement and their contribution can be found in Table 8.1 [47]. These sources include beam intensity, beam width evolution, length-scale correction, stability across pixel detector regions, afterglow, and scan-to-scan variation, where the uncertainties are summed in quadrature to obtain a total systematic uncertainty. The luminosity measurement uncertainty has the same normalization effect across all channels for background and signal yields. The luminosity systematic uncertainty does not have an effect on data-based background estimates.

Table 8.1: Summary of the systematic errors. When applicable, the percentage correction (on σ_{vis} for normalization effects and on the total luminosity for the afterglow effect) is also reported [47].

	Systematic	Correction (%)	Uncertainty (%)
Integration	Stability	—	1
	Dynamic inefficiencies	—	0.5
	Afterglow	~ 2	0.5
Normalization	Fit model	—	2
	Beam current calibration	—	0.3
	Ghosts and satellites	-0.4	0.2
	Length scale	-0.9	0.5
	Emittance growth	-0.1	0.2
	Orbit drift	0.2	0.1
	Beam-beam	1.5	0.5
	Dynamic- β	—	0.5
	Total	—	2.5

Cross section uncertainty

There are uncertainties associated with the calculation of theoretical cross sections that affect the predictions from signal and background simulations. Sources of uncertainties that impact the cross section calculations include choices made in the parton distribution functions, renormalization, and factorization scales. The exceptions to this are the WZ and ZZ cross sections, whose normalization is determined using data, as shown in Figures 7.1 (left) and 7.2. Systematic uncertainties of 6% and 12% are assigned to the production cross sections of WZ and ZZ processes, respectively, to account for the discrepancy between data and simulation, as previously described in Section 7.2.1.

Figure 7.3 shows that the number of $t\bar{t}$ events are well modeled in the dilepton control regions. However, background contributions from $t\bar{t}$ events that involve fake leptons originating from heavy-flavor jets are assigned a conservative 50% systematic uncertainty to account for the improper modeling of the process. The uncertainty value is chosen in order to cover the discrepancy between data and simulation in the relative isolation distribution of non-prompt muons for the $I_{\text{rel}} < 0.15$ GeV range, shown in Figure 7.4. We also include the theoretical uncertainty related to the $t\bar{t}$ cross section in the systematic uncertainty corresponding to the single lepton control region.

Lepton trigger efficiency uncertainty

The systematic uncertainty on the lepton trigger efficiency depends on the number of isolated electrons and muons and their p_T spectrum. The trigger efficiencies are measured using an unbiased H_T triggered data sample. Events are required to fire at least one of the double-muon, double-electron, or electron-muon triggers, as described in Section 6.2. This efficiency is measured by calculating the probability for at least one of the double-lepton triggers to be fired. An overall systematic uncertainty of 5% is determined for the lepton trigger efficiency. Both signal and background simulations are corrected for trigger efficiencies and are subject to this uncertainty.

Lepton isolation and identification efficiency uncertainties

Lepton efficiencies scale factors relating to the identification and isolation efficiency, as described in Section 6.3, are studied using the tag-and-probe method and are subject to uncertainties. These identification and isolation efficiencies do not completely match between data and simulation, with larger disagreement at low p_T values. Therefore, the lepton efficiency scale factor, given by the ratio between data and simulation efficiencies, is parametrized using a fit with respect to the probe lepton p_T . We determine the systematic uncertainty on the scale factor based on the deviation of the fit parameters calculated in various samples of a $Z + \text{jets}$ control region, for leptonically decaying Z bosons, in both data and simulation. Each sample is selected to have different number of jets or number of pileup vertices. The values of the fit parameters are compared and their differences are assigned as their uncertainty. The uncertainties on the fit parameters for the muon isolation efficiency scale factor are,

- $\Delta\sigma_\mu = 0.3416 \text{ (stat)} \pm 2.3697 \text{ (syst}_{\text{BE}}) \pm 1.8662 \text{ (syst}_{\text{jet}}) \pm 1.7979 \text{ (syst}_{\text{vert}})$,
- $\Delta SF_\mu^{\text{plateau}} = 0.0001 \text{ (stat)} \pm 0.002 \text{ (syst}_{\text{BE}}) \pm 0.0009 \text{ (syst}_{\text{jet}}) \pm 0.0002 \text{ (syst}_{\text{vert}})$,
- $\Delta SF_\mu^0 = 0.0039 \text{ (stat)} \pm 0.0371 \text{ (syst}_{\text{BE}}) \pm 0.1041 \text{ (syst}_{\text{jet}}) \pm 0.0166 \text{ (syst}_{\text{vert}})$,

where (stat) is the statistical uncertainty, (syst_{BE}) is the uncertainty due to differences between the barrel and endcap regions of the CMS detector, (syst_{jet}) is the uncertainty associated with number of jets in the event, and similarly (syst_{vert}) is the uncertainty related to the number of vertices in the event. The scale factor for the muon identification efficiency is constant, and, as a consequence, there are no uncertainties associated with any fit parameters but only to its overall value, which is given by,

- $\Delta SF_\mu = 0.0001 \text{ (stat)} \pm 0.0021 \text{ (syst}_{\text{BE}}) \pm 0.0018 \text{ (syst}_{\text{jet}}) \pm 0.0017 \text{ (syst}_{\text{vert}})$.

The uncertainties on fit parameters for the electron isolation efficiency scale factor are,

- $\Delta\sigma_e = 0.5597 \text{ (stat)} \pm 0.5075 \text{ (syst}_{\text{BE}}) \pm 1.9723 \text{ (syst}_{\text{jet}}) \pm 2.839 \text{ (syst}_{\text{vert}})$,

- $\Delta SF_e^{\text{plateau}} = 0.0001 \text{ (stat)} \pm 0.001 \text{ (syst}_{\text{BE}}) \pm 0.0004 \text{ (syst}_{\text{jet}}) \pm 0.0001 \text{ (syst}_{\text{vert}})$,
- $\Delta SF_e^0 = 0.0052 \text{ (stat)} \pm 0.0054 \text{ (syst}_{\text{BE}}) \pm 0.015 \text{ (syst}_{\text{jet}}) \pm 0.005 \text{ (syst}_{\text{vert}})$,

while uncertainties on fit parameters for the electron identification efficiency scale factor are,

- $\Delta\sigma_e = 0.3236 \text{ (stat)} \pm 9.582 \text{ (syst}_{\text{BE}}) \pm 0.3659 \text{ (syst}_{\text{jet}}) \pm 0.429 \text{ (syst}_{\text{vert}})$,
- $\Delta SF_e^{\text{plateau}} = 0.0002 \text{ (stat)} \pm 0.0116 \text{ (syst}_{\text{BE}}) \pm 0.0008 \text{ (syst}_{\text{jet}}) \pm 0.0013 \text{ (syst}_{\text{vert}})$,
- $\Delta SF_e^0 = 0.0151 \text{ (stat)} \pm 0.1283 \text{ (syst}_{\text{BE}}) \pm 0.054 \text{ (syst}_{\text{jet}}) \pm 0.0113 \text{ (syst}_{\text{vert}})$.

The uncertainties on the fit parameters for the τ -lepton isolation efficiency scale factor are,

- $\Delta\sigma_\tau = 0.79475 \text{ (stat)} \pm 2.742 \text{ (syst}_{\text{BE}}) \pm 1.15442 \text{ (syst}_{\text{jet}}) \pm 1.1089 \text{ (syst}_{\text{vert}})$,
- $\Delta SF_\tau^{\text{plateau}} = 0.000229813 \text{ (stat)} \pm 0.010013 \text{ (syst}_{\text{BE}}) \pm 0.00126 \text{ (syst}_{\text{jet}}) \pm 0.000301 \text{ (syst}_{\text{vert}})$,
- $\Delta SF_\tau^0 = 0.00517409 \text{ (stat)} \pm 0.023456 \text{ (syst}_{\text{BE}}) \pm 0.132649 \text{ (syst}_{\text{jet}}) \pm 0.007182 \text{ (syst}_{\text{vert}})$.

The total systematic uncertainty for the lepton efficiency scale factor is taken as the sum in quadratures of all the fit parameters uncertainties. Simulations are corrected by appropriate scale factors, whereby muon and electron isolation and identification efficiencies between data and simulation agree within less than a percent for lepton $p_T > 20 \text{ GeV}$. The muon identification/isolation scale factor uncertainties are 11% for p_T of 10 GeV and 0.2% at 100 GeV. The uncertainties for the electron identification/isolation scale factor are 14% for p_T of 10 GeV and 0.6% at 100 GeV. Lastly, the uncertainties for the tau identification/isolation scale factor are 2% for p_T of 10 GeV and 1.1% at 100 GeV.

E_T^{miss} resolution uncertainty

The E_T^{miss} resolution procedure previously discussed in Section 7.2.1 involves events without any genuine E_T^{miss} from neutrinos. We proceed to give a more detailed description of the method along with the evaluation of its systematic uncertainty [116]. The E_T^{miss} resolution

is parameterized based on the number of reconstructed vertices N_{vertex} and on the H_T of the event to account for pileup and jet activity, respectively. Any possible mis-measurement in the jet energy, for events with large values of H_T , can worsen the precision of the E_T^{miss} resolution. As previously mentioned, the x- and y-components of the E_T^{miss} can be approximated by Gaussian distributions. Consequently, we model the E_T^{miss} distribution as a sum of Rayleigh distributions, given by,

$$p(E_T^{\text{miss}}) = \sum_{ij} w_{ij} \frac{E_T^{\text{miss}}}{\sigma_{ij}^2} e^{-(E_T^{\text{miss}})^2 / 2\sigma_{ij}^2}, \quad (8.1)$$

where “ i ” represents the number of reconstructed vertices, and “ j ” corresponds to an H_T bin of width 40 GeV. The weight w_{ij} is the fraction of events in a particular E_T^{miss} bin, which have i vertices and H_T between $j \times 40$ GeV and $(j + 1) \times 40$ GeV. The E_T^{miss} resolution is characterized by σ_{ij} , which depends on N_{vertex} and H_T . The widths are fitted for in a Z-enriched dilepton sample by adjusting them to match that of data.

We apply smearing factors to simulated events, on an event-by-event basis depending on N_{vertex} and H_T , to correct the sample. The systematic uncertainty associated with the smearing factors applied to the E_T^{miss} distribution is obtained by investigating the migration of events due to the smearing. We evaluate the systematic uncertainty on the E_T^{miss} resolution by varying the fit parameter of the Rayleigh distribution and study the event migration between the different E_T^{miss} channels. The largest event migration across all search channels is taken as the systematic uncertainty on the E_T^{miss} resolution. The systematic uncertainty can be either correlated or anti-correlated across the various E_T^{miss} channels of the analysis, since the number of events is conserved between the search channels. The magnitude of the adjustment to the E_T^{miss} , as a result of the smearing, of the simulated events can be as low as a few percent to as high as 25%. The WZ background has about a 4% systematic uncertainty due to the E_T^{miss} resolution. A more comprehensive explanation on the modeling of the E_T^{miss} resolution can be found in Reference [122].

Jet energy scale uncertainty

The impact of the jet energy scale uncertainty on the background and signal yields are studied by shifting the jet p_T in the events up and down by a correction factor that depends on p_T and η . This leads to event migrations among the search channels affecting both the estimate on H_T and b-tagged jet events. We assign as the systematic uncertainty the largest variation, due to event migration, in the expected number of events across all channels. The jet energy scale uncertainty for WZ is approximately 0.5%.

b-tagging scale factor uncertainty

In Section 6.4 we described that b-tagging scale factors are applied to MC simulations to more accurately model collision data. The systematic uncertainty produced by the b-tagging scale factors is evaluated using additional fit functions provided by CMS collaboration. These auxiliary fit functions are obtained by varying the nominal functions with a $\pm 1\sigma$ uncertainty. This allows us to investigate event migrations between search channels, such that the difference is assigned as a systematic uncertainty for the channel due to the b-tagging scale factor. We evaluate the systematic uncertainties from b-tagging scale factors to be approximately 0.1% and 6% for WZ and $t\bar{t}$ samples, respectively. Analogous scale factors and systematic uncertainties exist for c-quark and light-quark jets, which may be b-tagged.

Initial state radiation uncertainty

Jets from initial state radiation can increase the H_T of an event. This effect is taken into account for background prediction from data-derived methods and simulated samples, where jet matching has been applied [123]. However, some signal models have not been subjected to the jet matching procedure, which means that jets from ISR may not have been correctly modeled. For such signal samples, we calculate a systematic uncertainty due to the improper modeling of ISR jets.

The main impact of these ISR jets is to migrate events out of the low H_T channels ($H_T < 200 \text{ GeV}$) and into the high H_T channels ($H_T > 200 \text{ GeV}$). We evaluate the systematic uncertainty by adding an H_T distribution from ISR jets to the signal and study how many events migrate from a low H_T range into a high H_T range. This uncertainty is naturally anti-correlated between the low- and high- H_T channels because of the event conservation. If the high H_T range gains events then the low H_T range must lose them.

We assume that the majority of jets in a $Z + \text{jets}$ sample are from ISR to obtain an H_T distribution for ISR jets. Based on this assumption, we use the H_T distribution from $Z + \text{jets}$ data as a template for the H_T distribution of ISR jets. From this H_T distribution we determine a probability for a signal event with a given H_T to be pushed above the 200 GeV threshold by ISR jets. For example, considering a signal with no hard jets, such that the H_T from the hard interaction is zero. In this case there is a 0.8% probability that ISR jets will push the event into the high H_T category, and as a result we assign a 0.8% systematic uncertainty. If on the other hand the signal always produces enough jets to have an H_T above 200 GeV, then the systematic uncertainty would be 0% since there would not be any events in the low H_T range to migrate up. As another example, a signal where 50% of the events have an H_T between 170 and 200 GeV, and the other 50% have an $H_T > 200 \text{ GeV}$. In this scenario there is a 20.8% chance that ISR jets migrate events from the low H_T range into the high H_T range. Therefore, there would be a -20.8% systematic uncertainty in the low H_T channel and a +20.8% systematic uncertainty in the high H_T channel. The ISR systematic uncertainty ranges from 0%, for models with lots of jets, to 10%, for models with few number of jets. Systematic uncertainties can be either correlated or anti-correlated depending on whether event migration showed an upward fluctuation in one channel with a corresponding deficit in another.

8.4 Summary of systematic uncertainties

Table 8.2 gives a summary of the most important systematics uncertainties used in the analysis and covered in this chapter.

Table 8.2: Brief summary of systematic uncertainty values.

Source of uncertainty	Magnitude (%)
Luminosity	2.6
$t\bar{t}$ cross section/isolation variable	50
Trigger efficiency	5
Muon ID/isolation at 10 (100) GeV	11 (0.2)
Electron ID/isolation at 10 (100) GeV	14 (0.6)
τ_h -lepton ID/isolation at 10 (100) GeV	2 (1.1)
E_T^{miss} resolution for WZ events	~ 4
Jet energy scale (WZ)	0.5
b-tagging scale factor	0.1 (WZ), 6 ($t\bar{t}$)
ISR modeling	0 – 5

Chapter 9

Statistical Analysis Method

We begin the chapter by introducing a statistical model for a multichannel counting experiment. Afterwards, we discuss a frequentist method for statistical inference and summarize the procedure that tests the agreement between observation and expected background contribution from SM processes. This is followed by a description of the approach involved in the limit setting procedure, where the limits are determined using a modified frequentist technique, to constrain signal cross sections and, thereby, establish lower mass limits.

9.1 Statistical technique

Different statistical methods are employed in order to describe the non-observation of a signal or to establish the significance of any excess in the number of observed events. The two commonly used statistical approaches for characterizing the absence of a signal are the *Bayesian* and *Frequentist* methods [17]. In the Bayesian method for statistical inference, the probability is interpreted as the *degree of belief* for an underlying parameter of interest, which is being constrained. Moreover, Bayesian inference introduces a prior probability distribution function, which reflects the a priori knowledge of where the parameter of interest should lie, such that Baye's theorem updates this distribution in view of any new relevant information. In the frequentist method for statistical inference, the concept of probability is understood as the frequency of occurrence of an outcome measured when the experiment is repeated.

Exclusion limits can be derived by considering two different hypotheses. The first corresponding to the signal plus background hypothesis (i.e. alternative hypothesis), H_{s+b} , and the second corresponding to background-only hypothesis (i.e. null hypothesis), H_b . By comparing these hypotheses against data we can place limits on the new physics signal using statistical inference to measure incompatibility with data.

9.2 Limit setting procedure

The statistical inference for limit settings depends on the statistical model, which describes the probability of observing a given dataset, and on the underlying parameters, such as the signal cross section, or more specifically the signal rate.

In this analysis, we rely on a multichannel counting experiment as the statistical modeling. The dataset is specified by the observed number of events, n , the number of expected background events, b , and the number of predicted signal events, s , all of which correspond to the parameters of interest. The probability of observing n events follows a Poisson distribution,

$$\mathcal{P}_{\text{poisson}}(n|\lambda) = \frac{\lambda^n \cdot e^{-\lambda}}{n!}, \quad (9.1)$$

where the mean of the distribution λ is given by b in the null hypothesis (i.e. background-only) and $s + b$ in the alternative hypothesis (i.e. background plus signal).

The probability distribution function for a multichannel analysis, assuming a background-only hypothesis, corresponds to the product of Poisson probabilities over all search channels, given by,

$$\mathcal{P}_{\text{poisson}}(\mathbf{n}|\mathbf{b}) = \prod_i^{N_{\text{channels}}} \frac{b_i^{n_i} e^{-b_i}}{n_i!}, \quad (9.2)$$

where N_{channels} are the number of search channels, $\mathbf{n} = \{n_i\}$ can be the observed number of events for the analysis or generated pseudo-data, such that n_i corresponds to the observed

number of events for a given channel “ i ”, and similarly $\mathbf{b} = \{b_i\}$ are the background estimates for the analysis, such that b_i is the expected number of events for a given channel “ i ”. The background estimate b_i depends on the integrated luminosity, SM background cross sections, and search channel selection efficiency. Additionally, the background estimate b_i can have contribution from several different sources of SM background. In that case,

$$b_i = \sum_j^{N_{\text{background}}} b_{ij}, \quad (9.3)$$

where b_{ij} is the j^{th} background contribution to the i^{th} search channel.

For the case of the signal plus background hypothesis, the probability distribution function for a multichannel search is defined by,

$$\mathcal{P}_{\text{poisson}}(\mathbf{n}|\mathbf{s} + \mathbf{b}) = \prod_i^{N_{\text{channels}}} \frac{(s_i + b_i)^{n_i} e^{-(s_i + b_i)}}{n_i!}. \quad (9.4)$$

where $\mathbf{s} = \{s_i\}$ are the signal estimates for the search, such that s_i corresponds to the number of signal events for a given channel “ i ”. The signal estimate s_i depends on the integrated luminosity, signal theory cross section, search channel selection efficiency, as well as on branching fractions.

9.3 LHC-type CL_s method

Exclusion limits are derived based on a modified frequentist technique, known as the CL_s method, introduced by the CMS and ATLAS collaborations [124]. A particular feature of the CL_s technique is that it protects against false exclusion in case that the experiment has little to no sensitivity for a signal and is accounted for by considering the compatibility of the background-only hypothesis with observation [125, 126, 127, 128]. This type of scenario can occur in searches which have a small signal on top of a large background, whereby overestimating the background can result in small upper limits on a signal cross section. The CL_s method sets limits on the cross section and masses derived from models

predicted by theory while constraining the possible ranges of the parameters of the model. We adopt the LHC-type CL_s method to set limits.

9.3.1 Likelihood function

In general, the procedure for setting limits depends on several parameters, which are characterized as either nuisance parameters or parameter of interest [129]. The nuisance parameters, denoted by $\boldsymbol{\theta} = \{\theta_i\}$, are any parameters that are not of interest in the experimental measurement, although they still having an effect on the predictions, and are associated with both statistical and systematic uncertainties. The background and signal predictions are subject to many systematic uncertainties and their effects are introduced by modifying the rates $s(\boldsymbol{\theta})$ and $b(\boldsymbol{\theta})$ with the “nuisance parameters” $\boldsymbol{\theta}$, where the uncertainties are modeled with log-normal probability distribution functions. The simulated signal sample is scaled by the *signal strength modifier* μ and can be related to the theoretical cross section of the signal by $\mu = \sigma^{(100-\alpha)\%}/\sigma_{\text{theory}}$, where $\alpha\%$ is the CL_s confidence level, which will be explained shortly. However, the parameter of interest, denoted by μ , is being constrained in the analysis in the absence of a signal.

The probability distribution in Equation 9.4 can be extended to include the parameter of interest μ and the nuisance parameters $\boldsymbol{\theta}$, and is referred to as the “likelihood function”. Furthermore, through Bayes’ theorem the likelihood function incorporates the *pdf* $p(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta})$ to model the nuisance parameter. The full likelihood function for a multichannel counting experiment corresponds to the product of the individual likelihood function for each search channel. In principle, the likelihood function $\mathcal{L}(\mathbf{n}|\mu, \boldsymbol{\theta})$ gives the probability to measure the dataset $\mathbf{n}$ given μ and $\boldsymbol{\theta}$. We define the likelihood function as,

$$\mathcal{L}(\mathbf{n}|\mu, \boldsymbol{\theta}) = \mathcal{P}_{\text{poisson}}(\mathbf{n}|\mu \cdot s(\boldsymbol{\theta}) + b(\boldsymbol{\theta})) \cdot p(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta}) \cdot \pi_{\boldsymbol{\theta}}(\boldsymbol{\theta}), \quad (9.5)$$

where $\mathbf{n}$ corresponds to the data either the observed or pseudo-data involved in constructing

the sampling distribution.

The frequentist auxiliary “measurement” *pdf* $p(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta})$, which are posteriors arising from some real or imaginary measurements $\tilde{\boldsymbol{\theta}}$, is used to constrain the likelihood function of the main measurement in a frequentist calculation, and is given by,

$$p(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta}) = \prod_i^{N_{\text{channels}}} \prod_j^{N_{\text{uncertainty}}} \frac{1}{\ln(\kappa_{ij})\sqrt{2\pi}} \frac{1}{\tilde{\theta}_{ij}} \exp\left(-\frac{(\ln(\tilde{\theta}_{ij}/\theta_{ij}))^2}{2(\ln \kappa_{ij})^2}\right), \quad (9.6)$$

where $\tilde{\boldsymbol{\theta}} = \{\tilde{\theta}_{ij}\}$ are the best estimates of the nuisance parameters $\boldsymbol{\theta}$, such that $\tilde{\theta}_{ij}$ corresponds to the best estimate of the nuisance parameter θ_{ij} for the i^{th} search channel and j^{th} source of uncertainty, and κ_{ij} is the statistical or systematic uncertainties of the i^{th} search channel from the j^{th} source. Examples of sources of uncertainties include the integrated luminosity measurement uncertainty, cross section uncertainty, or lepton selection efficiency uncertainties. The $\pi_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ functions are hyper-priors for those “measurements” and are assumed to be flat.

The signal rates in the likelihood function are modified as follows,

$$s_i = s(\theta_i) = \sum_j^{N_{\text{uncertainties}}} \tilde{s}_i \cdot (\kappa_{ij})^{\theta_{ij}}, \quad (9.7)$$

where $\tilde{s}_i$ is the number of signal events in the i^{th} search channel. While background rates are given by,

$$b_i = b(\theta_i) = \sum_j^{N_{\text{uncertainties}}} \tilde{b}_i \cdot (\kappa_{ij})^{\theta_{ij}}, \quad (9.8)$$

where $\tilde{b}_i$ is the number of background events in the i^{th} search channel.

9.3.2 Test-statistics

The general idea behind calculating an upper limit on the signal cross section is based on hypothesis testing, where the alternative hypothesis is the scenario with signal while the null

hypothesis corresponds to the scenario without signal. The alternative and null hypotheses are evaluated according to a test-statistic, which is a function of the parameter of interest, namely μ . By comparing the test-statistics value with data using the test-statics distribution for the alternative and null hypotheses, we can either discover or exclude a signal. The LHC-type CL_s approach uses a test-statistics based on the profiled log-likelihood ratio, which modifies the definition to include systematic uncertainties, to determine the compatibility of data against the alternative and null hypotheses, and is defined by,

$$Q_\mu = -2 \ln \frac{\mathcal{L}(\mathbf{n}|\mu, \hat{\boldsymbol{\theta}}_\mu)}{\mathcal{L}(\mathbf{n}|\hat{\mu}, \hat{\boldsymbol{\theta}})}, \quad (9.9)$$

where the numerator is the maximized conditional likelihood function and the denominator is the maximized unconditional likelihood function. In essence the test statistics summarizes the relevant properties of the data and is used to determine how signal-like or background-like data is.

The conditional maximum-likelihood estimator (MLE) $\hat{\boldsymbol{\theta}}_\mu$ denotes the value of $\boldsymbol{\theta}_\mu$ that maximizes the likelihood function $\mathcal{L}(\mathbf{n}|\mu, \boldsymbol{\theta}_\mu)$ for a specified signal strength modifier μ given a dataset $\mathbf{n}$. We determine the conditional MLEs, which are functions of μ , by solving the following equations,

$$\left. \frac{\partial}{\partial \theta_i} \mathcal{L}(\mathbf{n}|\mu, \boldsymbol{\theta}) \right|_{\theta_i = (\hat{\theta}_\mu)_i} = 0, \quad (9.10)$$

where $\mathbf{n}$ are the number of observed events from the collision data that is used in the search.

The global maximum-likelihood estimators $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ correspond to the global maximum of the likelihood function $\mathcal{L}(\mathbf{n}|\mu, \boldsymbol{\theta})$ for the values μ and $\boldsymbol{\theta}$, respectively. We determine the global MLEs by solving the following simultaneous equations,

$$\left. \frac{\partial}{\partial \mu} \mathcal{L}(\mathbf{n}|\mu, \boldsymbol{\theta}) \right|_{\mu = \hat{\mu}, \boldsymbol{\theta} = \hat{\boldsymbol{\theta}}} = 0. \quad (9.11)$$

and

$$\left. \frac{\partial}{\partial \theta_i} \mathcal{L}(\mathbf{n}|\mu, \boldsymbol{\theta}) \right|_{\mu=\hat{\mu}, \theta_i=\hat{\theta}_i} = 0, \quad (9.12)$$

where $\mathbf{n}$ is defined in the same manner as for Equation 9.10.

The global MLE $\hat{\mu}$ is constrained to be non-negative to ensure that signal rate is physical and, therefore, only positively scaled cross sections are considered. The upper bound $\hat{\mu} \leq \mu$ is imposed such that the limits on the signal strength are a one-sided confidence interval, which implies that the test-statistic is defined to be zero only for $\hat{\mu} \geq \mu$. These requirements are such that only upper limits of the signal cross sections are considered. It also ensure that upward fluctuations of the data are not considered as evidence against the signal hypothesis with strength μ [124, 125]. The larger the likelihood fit value the more the dataset $\mathbf{n}$ disagrees with parameter of interest μ compared to the predicted value $\hat{\mu}$.

The procedure of calculating the nuisance parameters, by performing a constrained maximum likelihood fit for a given dataset and parameter of interest using Equations 9.10, 9.11, and 9.12, is referred to as profiling. For this reason the test-statistics is called a profiled log-likelihood ratio.

9.3.3 Observed limit

The approach used in the limit setting procedure is based on the LHC-type CL_s method, as prescribed by the CMS collaboration, using the profiled log-likelihood ratio as the test-statistics. We closely follow the outline in Section 2.1 of Reference [124] and briefly describe the procedure for evaluating the observed limit. The statistical significance of what is observed in data, whether it is due to chance, needs to be determined. We can construct *pdfs* of the test statistics under the alternative hypotheses through statistical sampling using simulated pseudo-data that follows a Poisson probability distribution. By using these *pdfs* we can evaluate the probability $\mathcal{P}(q_\mu \geq Q_\mu^{\text{obs}} | H_{s+b})$ for the observed test-statistic value Q_μ^{obs}

to be as or less compatible with the H_{s+b} hypothesis.

As a result, the CL_{s+b} is define as the p-value of the integral of the test-statistics sampling distribution, *pdfs*, given by,

$$CL_{s+b} = \mathcal{P}(q_\mu \geq Q_\mu^{\text{obs}} | H_{s+b}) = p_\mu = \int_{Q_\mu^{\text{obs}}}^{\infty} f(q_\mu | \lambda(\mu, \hat{\theta}_\mu^{\text{obs}})) dq_\mu. \quad (9.13)$$

The *pdf* $f(q_\mu | \lambda(\mu, \hat{\theta}_\mu^{\text{obs}}))$ for these test-statistic values are determined by conducting pseudo-experiments using a Poisson *pdf* with mean $\lambda(\mu, \hat{\theta}_\mu^{\text{obs}})$ when generating a dataset $\mathbf{n}$ for Equation 9.9. The conditional MLE value $\hat{\theta}_\mu^{\text{obs}}$ is calculated with Equations 9.10 using the observed number of events $\mathbf{n}^{\text{obs}}$ from collision data, such that it maximizes the likelihood function under the alternative hypothesis. Similarly, we can use Equations 9.11 and 9.12 to calculate the global MLEs $\hat{\mu}^{\text{obs}}$ and $\hat{\theta}^{\text{obs}}$ with the dataset $\mathbf{n}^{\text{obs}}$, as well. The test-statistics for the observed measurement Q_μ^{obs} is determined with Equation 9.9 using $\mathbf{n}^{\text{obs}}$, $\hat{\theta}_\mu^{\text{obs}}$, $\hat{\mu}^{\text{obs}}$ and $\hat{\theta}^{\text{obs}}$.

We define CL_b as the probability to obtain a result less compatible with the background-only hypothesis than the observed one, which is given by,

$$CL_b = \mathcal{P}(q_\mu \geq Q_\mu^{\text{obs}} | H_b) = 1 - p_b = \int_{Q_0^{\text{obs}}}^{\infty} f(q_\mu | \lambda(0, \hat{\theta}_0^{\text{obs}})) dq_\mu. \quad (9.14)$$

The *pdf* $f(q_\mu | \lambda(0, \hat{\theta}_0^{\text{obs}}))$ for these test-statistic values are determined by using instead a mean of $\lambda(0, \hat{\theta}_0^{\text{obs}})$ for the Poisson *pdf* when generating a dataset $\mathbf{n}$ for Equation 9.9. The conditional MLE value $\hat{\theta}_0^{\text{obs}}$ is calculated with Equations 9.10 (setting $\mu = 0$) using $\mathbf{n}^{\text{obs}}$, such that it maximizes the likelihood function under the null hypothesis. We use the same global MLEs as in the case for the CL_{s+b} value. The test-statistics Q_0^{obs} is determined with Equation 9.9 using $\mathbf{n}^{\text{obs}}$, $\mu = 0$, $\hat{\theta}_0^{\text{obs}}$, $\hat{\mu}^{\text{obs}}$ and $\hat{\theta}^{\text{obs}}$.

Figure 9.1 shows the test-statistics distributions for the alternative (red) and null (blue) hypotheses needed for the calculation of the CL_{s+b} and CL_b values, respectively. Rather than rejecting the null hypothesis in favor of the alternative hypotheses, a CL_s “confidence

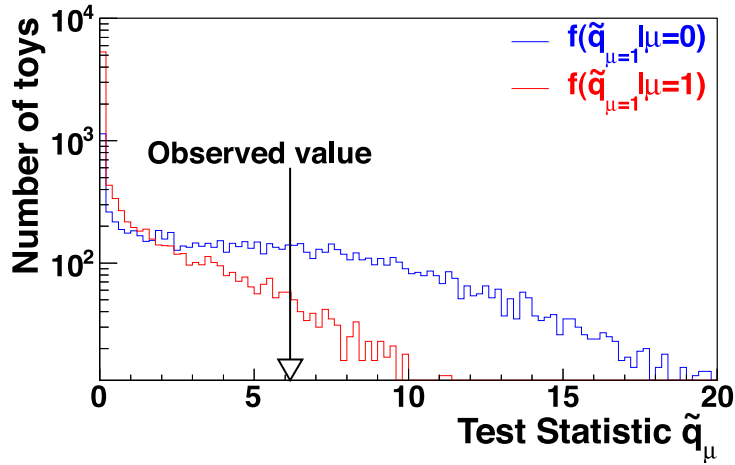


Figure 9.1: Test statistic distributions for ensembles of pseudo-data generated for the alternative (red) and null (blue) hypotheses [124]. See the text for definitions of the test statistic and methodology of generating pseudo-data.

level” (CL) is produced for alternative hypotheses corresponding a signal strength modifier value μ . The CL_s is defined as the ratio between CL_{s+b} and CL_b probabilities, and expressed as,

$$CL_s(\mu, Q_\mu^{\text{obs}}, Q_0^{\text{obs}}) = \frac{CL_{s+b}(\mu, Q_\mu^{\text{obs}})}{CL_b(\mu, Q_0^{\text{obs}})} \quad (9.15)$$

$$= \frac{p_\mu}{1 - p_b}. \quad (9.16)$$

Therefore, the alternative hypothesis can be rejected when,

$$CL_s(\mu, Q_\mu^{\text{obs}}, Q_0^{\text{obs}}) \leq \alpha, \quad (9.17)$$

meaning $\mu^{(100-\alpha)\%}$ is excluded with a $(100 - \alpha)\%$ CL. The CMS collaboration has chosen as a convention $\alpha = 0.05$, and so all $\mu^{95\%}$ are excluded at the 95% CL.

The CL_s technique prevents against unphysical limits in the event of large downward background fluctuations. For example, in the case of when the signal strength modifier is very close to zero or equals zero, and there is downward background fluctuations resulting

in a CL_{s+b} value of less than 5%, which implies that the signal is excluded. Therefore, this leads to an exclusion of the alternative hypothesis even when the predicted signal is very small, as is the case when the analysis has little to no signal sensitivity. The modified frequentist method accounts for this scenario, small signal on top of large background, by scaling it with the CL_b value, increasing the CL_s value above the 5% level and, therefore, preventing the exclusion.

9.3.4 Median expected limit and uncertainty bands

In addition to the observed limit obtained from the actual experiment (i.e. collision data), a median expected limit based on the background-only hypothesis and the integrated luminosity of the dataset is calculated, and is used to determine the expected level of sensitivity for new physics. We expect to exclude the signal as long as it can be accommodated by the background-only test hypothesis. The expected limits are calculated following the same prescription for the CL_s method, as described for the observed limit, whereby the value for the $\mu^{95\%}$ is determined from the test statistics distribution using background-only generated pseudo-data in both the numerator and denominator of Equation 9.9 and Q_μ^{bkg} instead of Q_μ^{obs} in the lower limits of the CL_s integrals 9.15.

A cumulative probability distribution (CDF) is extracted for these $\mu^{95\%}$ values. The median of all expected limits resulting from these pseudo experiments corresponds to the point where the cumulative probability distribution for $\mu^{95\%}$ crosses the 50% quantile. The $\pm 1\sigma$ (68%) uncertainties regions are given by the intersection of the CDF curve with the 84% and 16% quantiles, respectively. Similarly, the $\pm 2\sigma$ (95%) uncertainties regions are given by the intersection of the CDF curve with the 97.5% and 2.5% quantiles, respectively. Figure 9.2 shows CDF curve as a function of $\mu^{95\%}$, which is used to determine the median expected limit, 1σ (green region), and 2σ (yellow region) uncertainty bands.

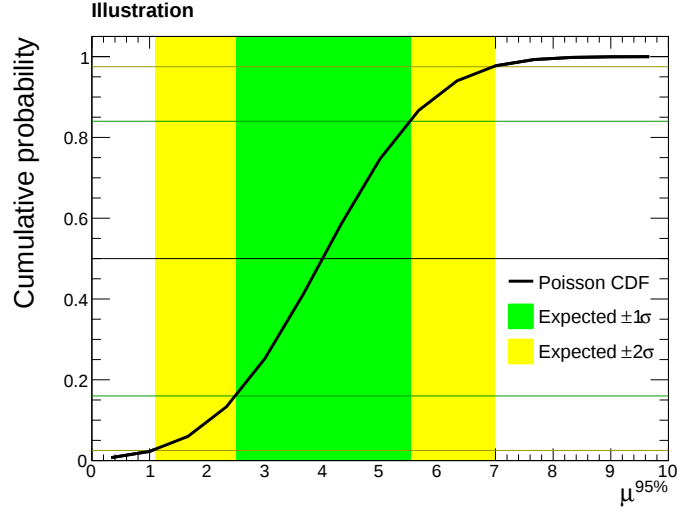


Figure 9.2: Cumulative probability distribution for possible experimental outcomes in terms of $\mu^{95\%}$ for the case of a high background rate. The median expected limit is given by the intersection of the 50% quantile with the CDF curve. The $\pm 1\sigma$ (68%) uncertainties bands (green region) are given by the intersection of the 84% and 16% quantiles with the CDF curve, respectively. Similarly, the $\pm 2\sigma$ (95%) uncertainties bands (yellow region) are given by the intersection of the 97.5% and 2.5% quantiles with the CDF curve, respectively. The quantiles are given by the horizontal lines.

9.4 Limit calculation with LandS

We observe no significant excess above the number of expected SM background events in this analysis, as discussed in Section 10.1. Therefore, we can proceed to extracting limits on the cross section for multilepton processes that could come from new physics signal. An upper limit is determined by calculating the maximum number of signal events that is statistically consistent with the observed number of events in data.

The exclusion limit calculation for this analysis is based on a modified frequentist CL_s method. We use the LandS software for setting the statistical limit at the 95% CL [130]. Furthermore, we follow the LHC-type CL_s prescription [124], which takes into account both statistical and systematic uncertainties. The LandS software combines the observed, background estimates, and predicted signal yields for each search channel, including their associated uncertainties, and computes the observed exclusion limit, as well as the expected median exclusion limit, along with its 1- and 2- σ uncertainty bands.

Nuisance parameters are provided for the search channels characterized by the sources of statistical and systematic uncertainties associated with the background and signal yields, using log-normal constraints. Examples of nuisance parameters include trigger and lepton efficiency uncertainties, luminosity uncertainty, and E_T^{miss} resolution systematic uncertainty, among others discussed in Section 8.3. Statistical uncertainties are uncorrelated across search channels, while all sources of systematic uncertainties are treated as either uncorrelated, fully correlated (i.e. positive), or fully anti-correlated (i.e. negative).

In a multichannel counting experiment not all search regions contribute equally for every particular signal scenario. Therefore, to reduce the necessary computational resources we calculate a combined limit based on the channels expected to provide the highest sensitivity, hence, channels expected to contain no signal are removed from the limit calculation. The search channels are added to the limit calculation in decreasing order of signal sensitivity until 90% of the expected signal yield is reached, whereby sensitivity of the channels are estimated based on their expected signal strength value $\sigma_{95\%}/\sigma_{\text{theory}}$. The search channels that are discarded have large SM background contributions, which if used can lead to a significant increase in the computational time needed to determine the limits for only a minimal gain in signal sensitivity. We also impose the additional constraint that no more than 42 search channels be considered in the limit calculations to further minimize the computational time.

Chapter 10

Experimental Results

In the first part of this chapter we present the experimental results of this analysis, by comparing the observed number of events with the background predictions as discussed in previous chapters. The second part focusses on the interpretation of the results in the context of various SUSY scenarios.

10.1 Multilepton results

This analysis is designed as a counting experiment based on a mutually exclusive multichannel approach, which allows for a model independent search for BSM physics in events with multiple leptons in the final states. Events are selected with at least 3 leptons including τ_h , corresponding to an integrated luminosity of 19.5 fb^{-1} of pp collisions at a center-of-mass energy $\sqrt{s} = 8 \text{ TeV}$, collected by the CMS detector during 2012 at the LHC. The data is divided into multiple exclusive search channels characterized by the number of leptons, the number of OSSF lepton pairs, the invariant mass of the OSSF lepton pair, the presence or absence of hadronically decaying τ leptons, the number of tagged b-jets, and the H_T and E_T^{miss} in the event.

A primary source of SM background for multilepton events arises from diboson production, such as WZ and ZZ, all estimated from MC simulations and corrected for pileup, jet multiplicity, trigger efficiencies, lepton efficiencies, E_T^{miss} resolution effects, and b-jet tagging efficiencies. Another important background contribution comes from $t\bar{t}$ production, which is also estimated using MC simulation, where the third prompt lepton, that is a

fake, originates from a jet (Section 7.2). Other SM backgrounds that provide fake leptons are determined using data-derived techniques (Section 7.3). We also consider contributions from rare SM processes such as $t\bar{t}W$, $t\bar{t}Z$, tbZ , and those involving Higgs production from gluon-gluon fusion, VBF, and in association with a W boson, Z boson, or $t\bar{t}$ pair, using MC simulations.

The main experimental results are presented in Tables 10.1 and 10.2 for events with exactly three leptons and events with four or more leptons, respectively. The tables show the number of observed events along with the expected number of SM background events. In the vast majority of the search channels examined, the observed number of events are in overall good agreement with the SM expectations. However, there are three channels where we observe an excess of data above the expectation from SM processes. In these channels events are required to have four-or-more leptons in the OSSF1, off-Z category with one τ_h -lepton candidate, no tagged b-jet, and $H_T < 200 \text{ GeV}$ ¹. Specifically, we observe 15 events with $0 < E_T^{\text{miss}} < 50 \text{ GeV}$, 4 events with $50 < E_T^{\text{miss}} < 100 \text{ GeV}$, and 3 events with $E_T^{\text{miss}} > 100 \text{ GeV}$, which have SM expectations of 7.5 ± 2.0 , 2.1 ± 0.5 , and 0.60 ± 0.24 , respectively. Figure 10.1 shows the E_T^{miss} distribution in data compared to the predictions from SM backgrounds for the search channels aforementioned. An expectation of 10.1 ± 2.4 events is estimated for the combined E_T^{miss} range. For the purposes of presentation, search channels in Tables 10.1 and 10.2 have been combined into coarser E_T^{miss} bins but not for the case of the limit setting procedure.

The joint probability to observe at least as large of an excess for all three channels considered individually is approximately 5%. Alternatively, we determine the single-measurement probability to observe 22 or more events when the expected number is 10.1 ± 2.4 events to be roughly 1%. However, once “trial factors” are incorporated to account for the 64 independent channels of the analysis, the probability to observe such a fluctuation increases to about 50% [131, 132, 133]. The systematic uncertainties and their correlations are taken into

¹Asymmetric internal photon conversion events were not rejected from the four or more lepton search channels in the off-Z category.

account when evaluating these probabilities. Although the data and background predictions agree when there is a Z boson candidate, the observations without a Z boson candidate are about two times the expectation. However, there is no clear discrepancy when examining the shape of kinematic distributions between data and SM estimations (Section C.4). Further studies of the excess are found in Appendix C.

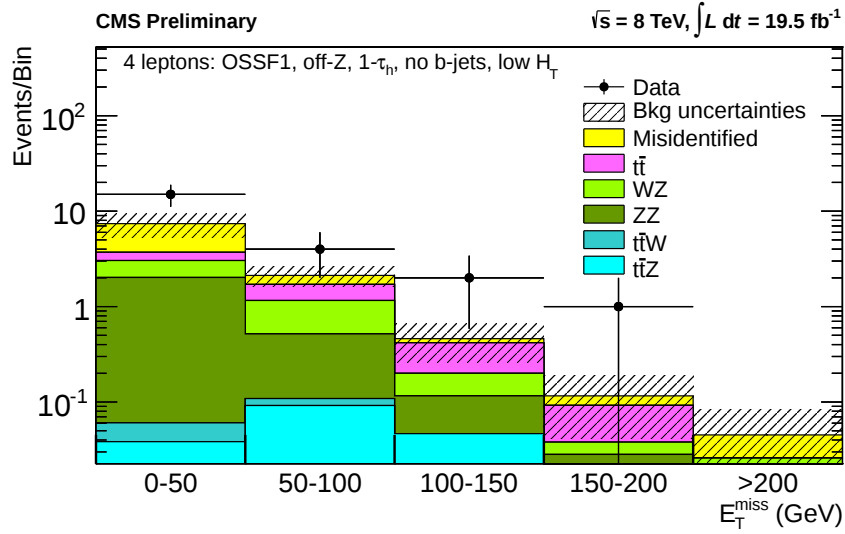


Figure 10.1: E_T^{miss} distribution of events with 4-leptons, one OSSF lepton pair off-Z, $1-\tau_h$, $H_T < 200$ GeV, and no b-jets.

Figure 10.2 shows the E_T^{miss} distributions of events with 3-leptons, no OSSF lepton pairs, and no τ_h , while Figures 10.3, 10.4, and 10.5 show the E_T^{miss} distributions of events with 3-leptons, one OSSF lepton pair below-Z, on-Z, or above-Z, respectively, and no τ_h . The plots have the SM background broken down by its various sources of contributions. If the signal from a model is expected to be in a channel with three electrons and muons, and large E_T^{miss} but without any OSSF lepton pairs, then the surrounding channels with low E_T^{miss} and OSSF lepton pairs on-Z become the control channels. Normally, the τ_h channels (moving from left to right) are control channels. However, for signals rich in τ_h -leptons the channels with no τ_h -leptons now become control channels. The E_T^{miss} distributions for the other remaining search channel are in Appendix A.

Table 10.1: Observed (Obs.) number of events with exactly three leptons from 19.5 fb^{-1} of 2012 data along with the expected (Exp.) number of SM background events. All channels are exclusive and all uncertainties include both systematic and statistical components. The OSSF_n designation indicates the number of e^+e^- and $\mu^+\mu^-$ pairs in the event, “On-Z” refers to events with at least one OSSF lepton pair with dilepton mass between 75 and 105 GeV, “Below-Z” and “Above-Z” refer to events with an OSSF lepton pair with mass below 75 GeV or above 105 GeV, respectively. Some search channels have been combined into coarser E_T^{miss} bins for the purposes of presentation. The channels marked with an asterisk are used as control regions and are therefore excluded from the limit calculations.

3 leptons $H_T > 200 \text{ GeV}$		$N_{\tau_h} = 0, N_{b\text{-jets}} = 0$		$N_{\tau_h} = 1, N_{b\text{-jets}} = 0$		$N_{\tau_h} = 0, N_{b\text{-jets}} \geq 1$		$N_{\tau_h} = 1, N_{b\text{-jets}} \geq 1$	
$m_{\ell\ell}$	E_T^{miss} (GeV)	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.
OSSF_0	—	5	3.7 ± 1.6	35	33 ± 14	1	5.5 ± 2.2	47	61 ± 30
OSSF_0	—	3	3.5 ± 1.4	34	36 ± 16	8	7.7 ± 2.7	82	91 ± 46
OSSF_0	—	4	2.1 ± 0.8	25	25 ± 10	1	3.6 ± 1.5	52	59 ± 29
OSSF_1	Above-Z	5	3.6 ± 1.2	2	10.0 ± 4.8	3	4.7 ± 1.6	19	22 ± 11
OSSF_1	Below-Z	7	9.7 ± 3.3	18	14.0 ± 6.4	8	9.1 ± 3.4	21	23 ± 11
OSSF_1	On-Z	39	61 ± 23	17	15.0 ± 4.9	9	14.0 ± 4.4	10	12.0 ± 5.8
OSSF_1	Above-Z	4	5.0 ± 1.6	14	11.0 ± 5.2	6	6.8 ± 2.4	32	30 ± 15
OSSF_1	Below-Z	10	11.0 ± 3.8	24	19.0 ± 6.4	10	9.9 ± 3.7	25	32 ± 16
OSSF_1	On-Z	78	80 ± 32	70	50 ± 11	22	22.0 ± 6.3	36	24.0 ± 9.8
OSSF_1	Above-Z	3	7.3 ± 2.0	41	33.0 ± 8.7	4	5.3 ± 1.5	15	23 ± 11
OSSF_1	Below-Z	26	25.0 ± 6.8	110	86 ± 23	5	10.0 ± 2.5	24	26 ± 11
OSSF_1	On-Z	*135	130 ± 41	542	540 ± 160	31	32.0 ± 6.5	86	75 ± 19
3 leptons $H_T < 200 \text{ GeV}$		$N_{\tau_h} = 0, N_{b\text{-jets}} = 0$		$N_{\tau_h} = 1, N_{b\text{-jets}} = 0$		$N_{\tau_h} = 0, N_{b\text{-jets}} \geq 1$		$N_{\tau_h} = 1, N_{b\text{-jets}} \geq 1$	
$m_{\ell\ell}$	E_T^{miss} (GeV)	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.
OSSF_0	—	7	11.0 ± 4.9	101	111 ± 54	13	10.0 ± 5.3	87	119 ± 61
OSSF_0	—	35	38 ± 15	406	402 ± 152	29	26 ± 13	269	298 ± 151
OSSF_0	—	53	51 ± 11	910	1035 ± 255	29	23 ± 10	237	240 ± 113
OSSF_1	Above-Z	18	13.0 ± 3.5	25	38 ± 18	10	6.5 ± 2.9	24	35 ± 18
OSSF_1	Below-Z	21	24 ± 9	41	50 ± 25	14	20 ± 10	42	54 ± 28
OSSF_1	On-Z	150	150 ± 26	39	48 ± 13	15	14.0 ± 4.8	19	23 ± 11
OSSF_1	Above-Z	50	46.0 ± 9.7	169	140 ± 48	20	18 ± 8	85	93 ± 47
OSSF_1	Below-Z	142	130 ± 27	353	360 ± 92	48	48 ± 23	140	133 ± 68
OSSF_1	On-Z	*773	780 ± 120	1276	1200 ± 310	56	47 ± 13	81	75 ± 32
OSSF_1	Above-Z	178	200 ± 35	1676	1900 ± 540	17	18.0 ± 6.7	115	94 ± 42
OSSF_1	Below-Z	510	560 ± 87	9939	9000 ± 2700	34	42 ± 11	226	228 ± 63
OSSF_1	On-Z	*3869	4100 ± 670	*50188	50000 ± 15000	*148	156 ± 24	906	925 ± 263

Table 10.2: Observed (Obs.) number of events with four or more leptons from 19.5 fb^{-1} of 2012 data along with the expected (Exp.) number of SM background events. All channels are exclusive and all uncertainties include both systematic and statistical components. The $\text{OSSF}n$ designation indicates the number of e^+e^- and $\mu^+\mu^-$ pairs in the event, “On-Z” refers to events with at least one OSSF lepton pair with dilepton mass between 75 and 105 GeV, and “Off-Z” refers to events where all possible OSSF lepton pairs fall outside this mass range. Some search channels have been combined into coarser E_T^{miss} bins for the purposes of presentation. The channel marked with an asterisk is used as a control region and is therefore excluded from the limit calculations

≥ 4 leptons $H_T > 200 \text{ GeV}$	$m_{\ell\ell}$	E_T^{miss} (GeV)	$N_{\tau_h} = 0, N_{b\text{-jets}} = 0$ Obs. Exp.	$N_{\tau_h} = 1, N_{b\text{-jets}} = 0$ Obs. Exp.	$N_{\tau_h} = 0, N_{b\text{-jets}} \geq 1$ Obs. Exp.	$N_{\tau_h} = 1, N_{b\text{-jets}} \geq 1$ Obs. Exp.
OSSF0	—	(100, ∞)	0	0	0	0
OSSF0	—	(50, 100)	0	0	0	0
OSSF0	—	(0, 50)	0	0	0	0
OSSF1	Off-Z	(100, ∞)	0	1	0	0
OSSF1	On-Z	(100, ∞)	1	0	0	0
OSSF1	Off-Z	(50, 100)	0	1	0	0
OSSF1	On-Z	(50, 100)	0	1	0	1
OSSF1	Off-Z	(0, 50)	0	0	0	0
OSSF1	On-Z	(0, 50)	0	0	0	1
OSSF2	Off-Z	(100, ∞)	0	—	0	—
OSSF2	On-Z	(100, ∞)	1	—	0	—
OSSF2	Off-Z	(50, 100)	0	—	0	—
OSSF2	On-Z	(50, 100)	0	—	0	—
OSSF2	Off-Z	(0, 50)	1	—	0	—
OSSF2	On-Z	(0, 50)	5	—	2	—
≥ 4 leptons $H_T < 200 \text{ GeV}$	$m_{\ell\ell}$	E_T^{miss} (GeV)	$N_{\tau_h} = 0, N_{b\text{-jets}} = 0$ Obs. Exp.	$N_{\tau_h} = 1, N_{b\text{-jets}} = 0$ Obs. Exp.	$N_{\tau_h} = 0, N_{b\text{-jets}} \geq 1$ Obs. Exp.	$N_{\tau_h} = 1, N_{b\text{-jets}} \geq 1$ Obs. Exp.
OSSF0	—	(100, ∞)	0	0	0	0
OSSF0	—	(50, 100)	0	2	0	0
OSSF0	—	(0, 50)	0	1	0	0
OSSF1	Off-Z	(100, ∞)	0	3	0	0
OSSF1	On-Z	(100, ∞)	1	2	1	0
OSSF1	Off-Z	(50, 100)	0	4	0	1
OSSF1	On-Z	(50, 100)	2	9	2	0
OSSF1	Off-Z	(0, 50)	2	15	0	0
OSSF1	On-Z	(0, 50)	4	41	1	2
OSSF2	Off-Z	(100, ∞)	0	—	0	—
OSSF2	On-Z	(100, ∞)	0	—	0	—
OSSF2	Off-Z	(50, 100)	2	—	0	—
OSSF2	On-Z	(50, 100)	4	—	0	—
OSSF2	Off-Z	(0, 50)	7	—	1	—
OSSF2	On-Z	(0, 50)	*156	—	4	—

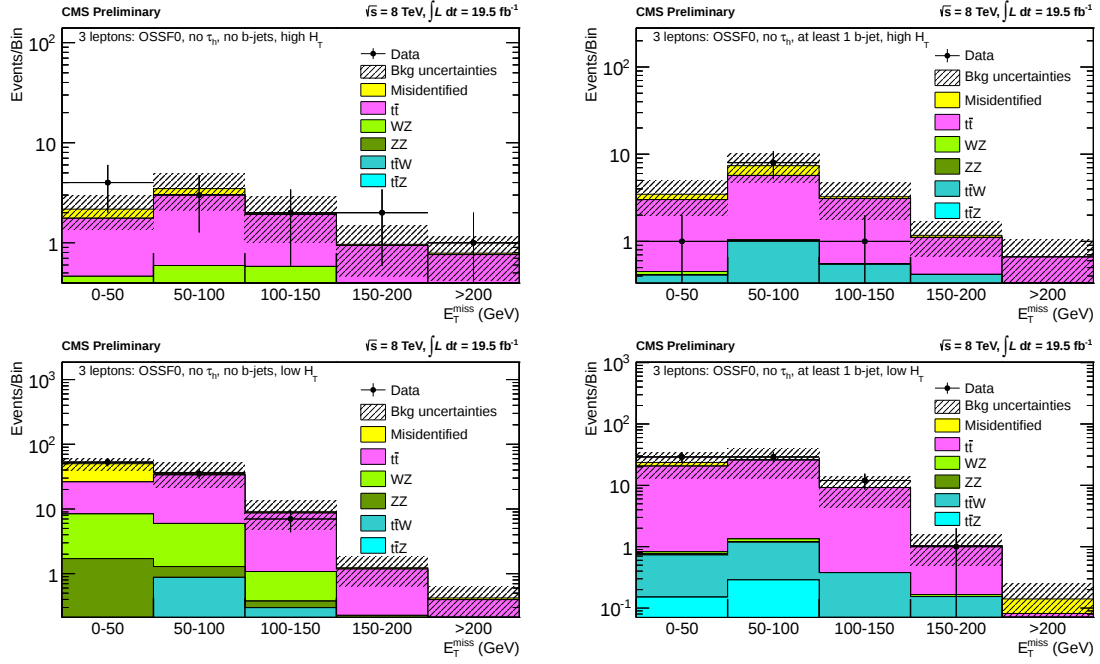


Figure 10.2: E_T^{miss} distribution of events with 3-leptons, no OSSF lepton pairs, no τ_h , $H_T > 200$ GeV (top) or $H_T < 200$ GeV (bottom), and 0 b-jets (left) or ≥ 1 b-jet (right).

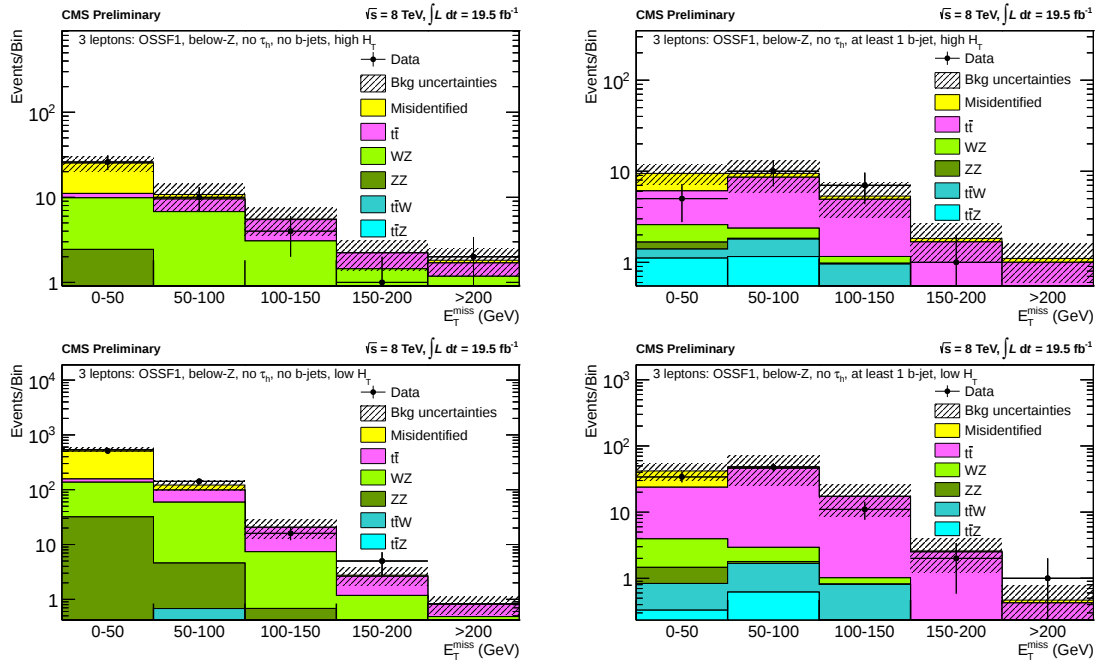


Figure 10.3: E_T^{miss} distribution of events with 3-leptons, one OSSF lepton pair below-Z, no τ_h , $H_T > 200$ GeV (top) or $H_T < 200$ GeV (bottom), and 0 b-jets (left) or ≥ 1 b-jet (right).

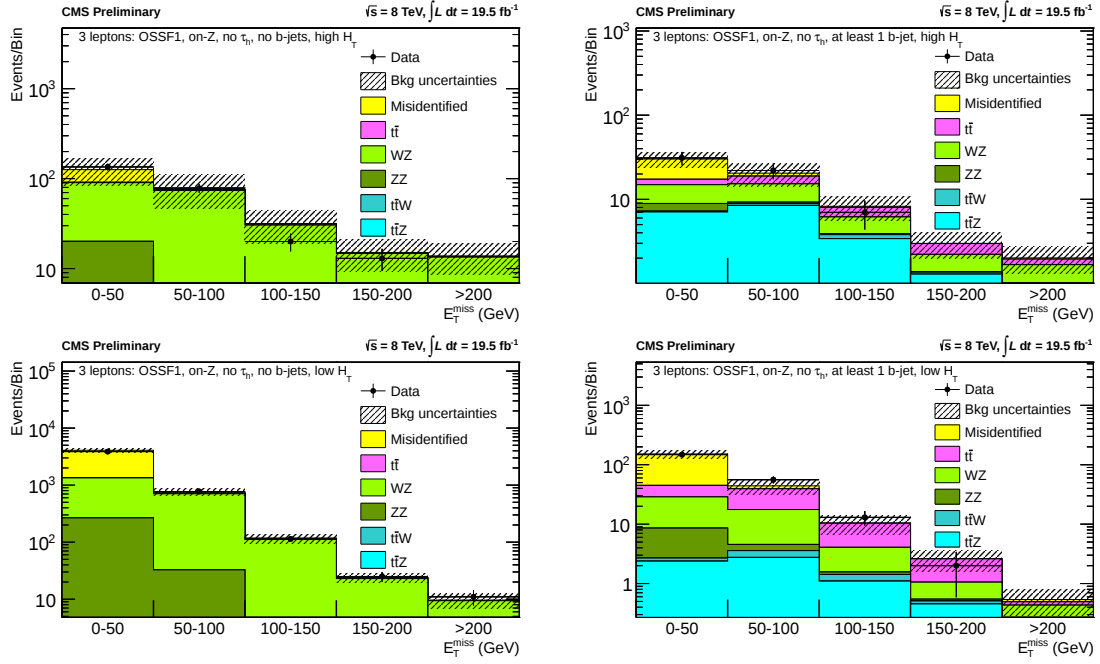


Figure 10.4: E_T^{miss} distribution of events with 3-leptons, one OSSF lepton pair on-Z, no τ_h , $H_T > 200$ GeV (top) or $H_T < 200$ GeV (bottom), and 0 b-jets (left) or ≥ 1 b-jet (right).

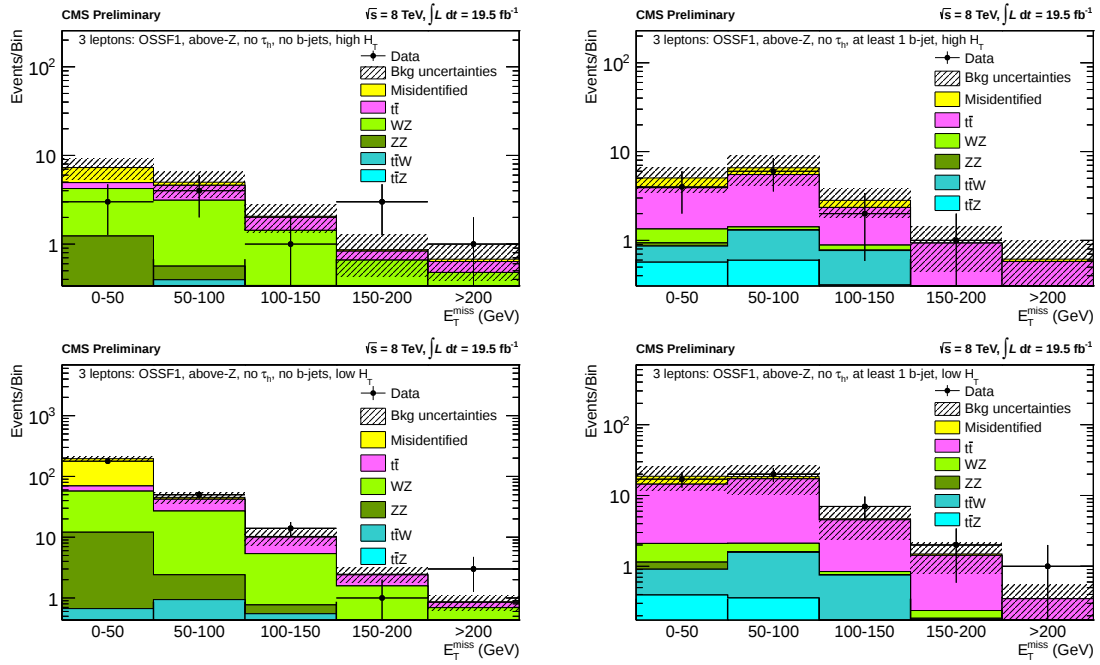


Figure 10.5: E_T^{miss} distribution of events with 3-leptons, one OSSF lepton pair above-Z, no τ_h , $H_T > 200$ GeV (top) or $H_T < 200$ GeV (bottom), and 0 b-jets (left) or ≥ 1 b-jet (right).

10.2 Interpretation of the results

We observe no significant excess, at the discovery level, above expected SM prediction. Therefore, we interpret the multilepton results in the context of six MSSM SUSY models and calculate their exclusion limits at the 95% confidence level. We perform a simultaneous fit across all exclusive search channels and compute the likelihood of observing a particular signal rate. With this approach we determine the sensitivity to different signal models for the purposes of excluding various cross sections, branching fractions, and mass ranges.

10.2.1 The natural higgsino NLSP scenario

We begin by describing a supersymmetric scenario that involves a “higgsino-like” neutralino, $\tilde{\chi}_1^0$, as the next-to-lightest supersymmetric particle (NLSP) known as the “natural higgsino NLSP” scenario [131, 132, 134, 133]. Gauge-mediated SUSY-breaking (GMSB) models give rise to this type of scenario. Both strong and electroweak SUSY production modes are considered. Figure 10.6 shows the Feynman diagram and the superpartner mass spectrum for the processes. The strong mechanism proceeds through right-handed top–anti-top squark pair production with the following cascade decays,

$$\begin{array}{c} \tilde{t}_R \tilde{t}_R^* \\ \text{b}\tilde{\chi}_1^+, \text{t}\tilde{\chi}_1^0, \text{t}\tilde{\chi}_2^0 \quad \leftarrow \quad \rightarrow \quad \bar{\text{b}}\tilde{\chi}_1^-, \bar{\text{t}}\tilde{\chi}_1^0, \bar{\text{t}}\tilde{\chi}_2^0, \end{array}$$

where $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$ are the lightest chargino and second-lightest neutralino, respectively, with both being “higgsino-like” as well, and where the $\tilde{t}_R^*$ state is the charge conjugate of the $\tilde{t}_R$ state.

The $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$ each decay to a $\tilde{\chi}_1^0$ and SM particles. Therefore, the final decay process in all cascades is $\tilde{\chi}_1^0 \rightarrow \text{H}\tilde{\text{G}}$ or $\text{Z}\tilde{\text{G}}$, assuming these are the only two decay modes for $\tilde{\chi}_1^0$, where $\tilde{\text{G}}$ is the Goldstino component of the massless and non-interacting gravitino (henceforth simply referred to as the gravitino) and the LSP. The particle H refers to the lightest

neutral CP-even SUSY Higgs boson, and if the other SUSY Higgs bosons are much heavier, then the SUSY Higgs boson is expected to have SM-like properties [24]. For this reason the Higgs boson is given a value of $m_H = 126 \text{ GeV}$. The cascade decays then yield HH, HZ, and ZZ final states [70]. The E_T^{miss} in the event comes from the undetected gravitinos.

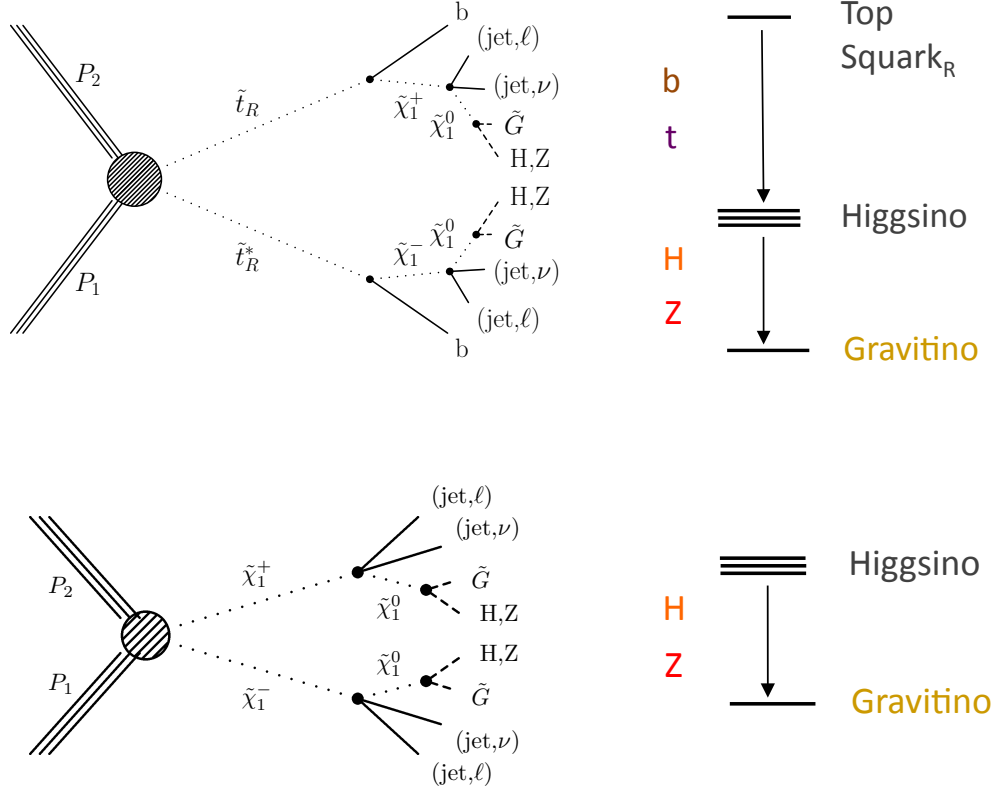


Figure 10.6: Feynman diagram and superpartner mass spectrum for the strong (top) and electroweak (bottom) SUSY production of the natural higgsino NLSP scenario, where the particles in parentheses (top left) have soft p_T spectrum.

The electroweak mechanism proceeds similarly to that of the strong interactions except that the $\tilde{\chi}_1^\pm \tilde{\chi}_2^0$ pair along with the other possible $\tilde{\chi}_1^\pm \tilde{\chi}_1^\mp$, $\tilde{\chi}_1^\pm \tilde{\chi}_1^0$, and $\tilde{\chi}_1^0 \tilde{\chi}_2^0$ pairs are produced directly from the proton-proton collisions and also leads to the same final states but with a lower number of jets in the event [70]. The $\tilde{\chi}_1^0 \tilde{\chi}_1^0$ and $\tilde{\chi}_2^0 \tilde{\chi}_2^0$ production processes vanish in the pure higgsino limit, as previously discussed in Section 4.2.3.

The natural higgsino NLSP scenario is particularly interesting and experimentally complex because of the many decay modes available to the Higgs boson that can yield events with three or more lepton, such as those from WW^* , ZZ^* , and $\tau\tau$, where W^* and Z^* indicate off-shell bosons². Hence, the final decay of the $\tilde{\chi}_1^0\tilde{\chi}_1^0$ pair to HH can lead to WW^*WW^* , ZZ^*ZZ^* , $\tau\tau\tau\tau$, WW^*ZZ^* , $WW^*\tau\tau$, $ZZ^*\tau\tau$, and ZZ^*bb final states, while the final decay to HZ can lead to WW^*Z , ZZ^*Z , and $\tau\tau Z$ final states. This scenario populates the largest range of the different search channels. The channels with the highest sensitivity are those with on-Z and off-Z selections for the decays through the HZ and ZZ states. Additionally for the decays through the HH state the most sensitive channels are the ones where at least one b-jet is required.

In Figure 10.7 we show one-dimensional exclusion limits on the cross section times branching fraction $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G}/Z\tilde{G})$ as a function of the top squark mass ($m_{\tilde{t}}$) for fixed $\tilde{\chi}_1^0$ branching fraction and chargino mass. The solid black curve corresponds to the observed 95% CL upper limit, the dashed curve is the median expected 95% CL upper limit, the green and yellow shaded regions correspond to the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands, respectively, on the expected limit. The solid blue curve is the theoretical prediction for the cross section as a function of $m_{\tilde{t}}$, with the dashed blue bands indicating the uncertainty on the cross section calculation. These curves are one dimensional projections of the plots in Figure 10.9. We have larger sensitivity at lower chargino masses since the cross section is larger but less sensitivity for the Higgs-boson-dominated mode in comparison with the Z-boson-dominated mode.

Figure 10.8 illustrates the 95% CL exclusion limits on branching fraction $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G})$ versus $m_{\tilde{t}}$ for different $\tilde{\chi}_1^\pm$ masses. We observe that we have greater sensitivity with the combination of strong and electroweak processes due to the fact that the gluon parton distribution functions drop as the top squark mass increases and strong production cross sections become comparable to electroweak production cross sections. As we go up on the

²The asterisks symbol * is used to denote both off-shell SM particles or the charge conjugate of SUSY particles depending on the context.

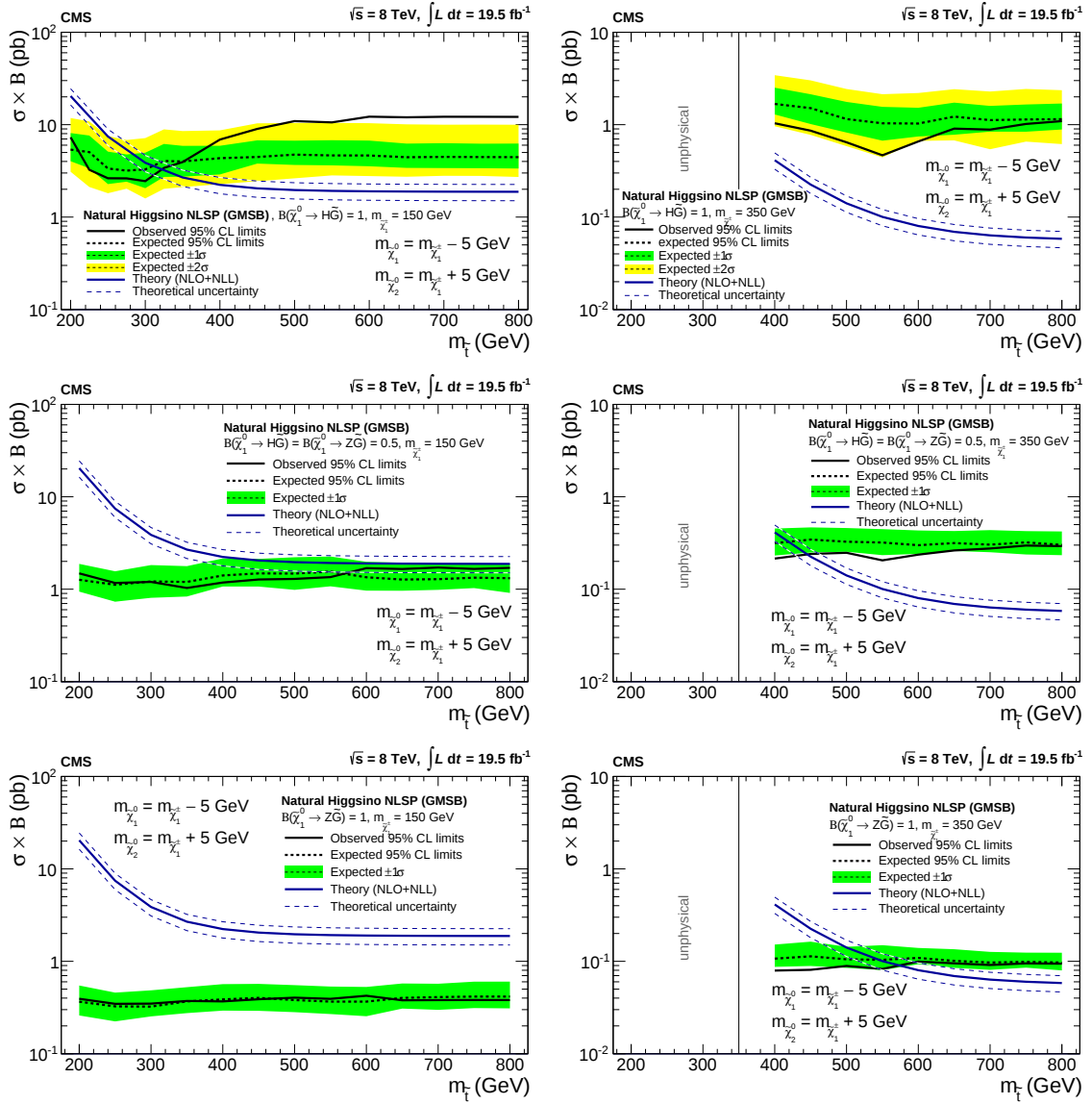


Figure 10.7: The 95% confidence level upper limits for the natural higgsino NLSP scenario on the cross section times branching fraction $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G}/Z\tilde{G})$. The three branching fractions of $\tilde{\chi}_1^0$ presented are $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G}) = 1.0$ (top), $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G}) = \mathcal{B}(\tilde{\chi}_1^0 \rightarrow Z\tilde{G}) = 0.5$ (middle), and $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow Z\tilde{G}) = 1.0$ (bottom). Both strong and electroweak SUSY production processes are taken into consideration. The $\tilde{\chi}_1^\pm$ mass is fixed at 150 GeV (left) and 350 GeV (right). The region of the bottom left of the contours is excluded. The region to the left of the vertical line is unphysical and limited by the $\tilde{\chi}_1^\pm$ mass. The green and yellow bands show the $\pm 1\sigma$ and $\pm 2\sigma$ variations on the expected limit due to experimental uncertainties, respectively.

vertical axis, the exclusion limit become worse, again, because the branching fraction to Z decreases, and we are left with pure Higgs boson production at the top of the figures.

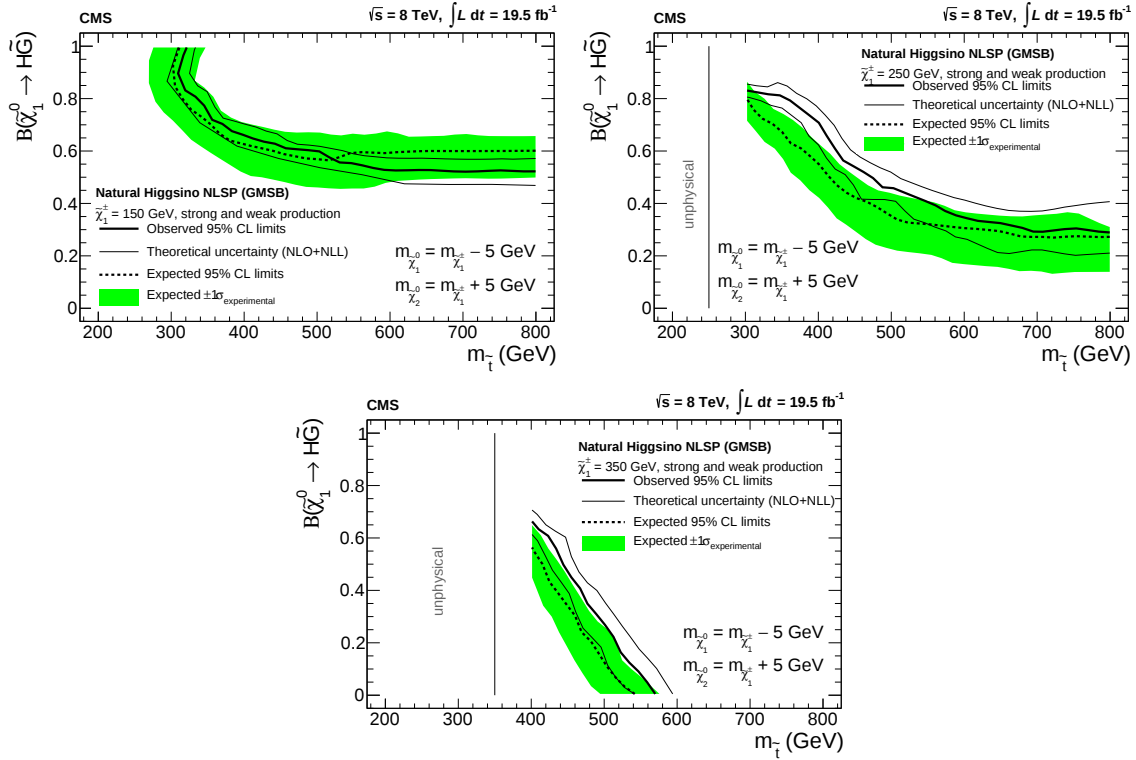


Figure 10.8: The 95% confidence level upper limits for the natural higgsino NLSP scenario on the branching fraction $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G})$ with fixed $\tilde{\chi}_1^\pm$ mass of 150 GeV (upper left), 250 GeV (upper right), and 350 GeV (bottom). It is assumed that $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G}) + \mathcal{B}(\tilde{\chi}_1^0 \rightarrow Z\tilde{G}) = 1.0$. Both strong and electroweak SUSY production processes are taken into consideration. The region to the left of the vertical line is unphysical and limited by the $\tilde{\chi}_1^\pm$ mass. The green band shows the $\pm 1\sigma$ variations on the expected limit due to experimental uncertainties.

Figure 10.9 shows the 95% CL exclusion limit in the $m_{\tilde{\chi}_1^\pm}$ versus $m_{\tilde{t}}$ mass plane for several $\tilde{\chi}_1^0$ branching fractions.

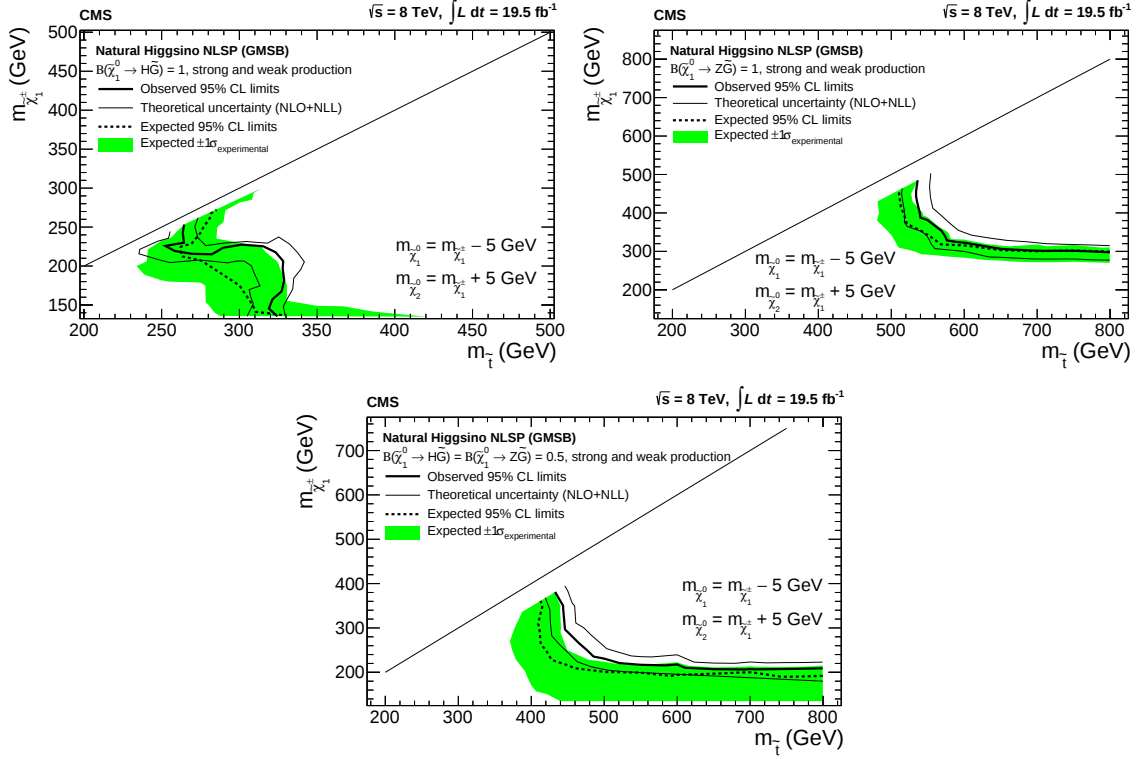


Figure 10.9: The 95% confidence level lower limits for the natural higgsino NLSP scenario in the top squark versus chargino mass plane. The three branching fractions of $\tilde{\chi}_1^0$ presented are $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G}) = 1.0$ (top left), $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow Z\tilde{G}) = 1.0$ (top right), and $\mathcal{B}(\tilde{\chi}_1^0 \rightarrow H\tilde{G}) = \mathcal{B}(\tilde{\chi}_1^0 \rightarrow Z\tilde{G}) = 0.5$ (bottom). Both strong and electroweak SUSY production processes are taken into consideration. The region to the left and below the contours is excluded. The region above the diagonal line is unphysical. The green band shows the $\pm 1\sigma$ variations on the expected limit due to experimental uncertainties.

10.2.2 The wino NLSP with a Higgs boson scenario

We next consider the wino NLSP with a Higgs boson scenario, in which both the “wino-like” chargino $\tilde{\chi}_1^\pm$ and neutralino $\tilde{\chi}_2^0$ serve as the NLSP. Figure 10.10 shows the Feynman diagram and the superpartner mass spectrum for the process. This process proceeds via electroweak SUSY chargino–neutralino pair production with the following cascade decays,

$$\tilde{\chi}_1^\pm \tilde{\chi}_2^0$$

$$W^\pm \tilde{\chi}_1^0 \quad \leftrightarrow \quad H \tilde{\chi}_1^0,$$

where the $\tilde{\chi}_1^0$ is “bino-like” and the LSP. The “wino-like” nature of the $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$ particles motivates the simplifying assumption that $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0}$. The particle H refers to the lightest neutral CP-even SUSY Higgs boson and is assumed to have a value of $m_H = 126$ GeV for the same reasons as given in the case of the natural Higgsino NLSP scenario. The branching fractions for the decays $\tilde{\chi}_1^\pm \rightarrow W^\pm \tilde{\chi}_1^0$ and $\tilde{\chi}_1^\pm \rightarrow H \tilde{\chi}_1^0$ are assumed to be unity. The $\tilde{\chi}_1^0$ is assumed to be stable and escapes detection, leading to large E_T^{miss} . The $\tilde{\chi}_1^\pm \tilde{\chi}_3^0$ pair production process vanishes in the pure wino limit, as previously discussed in Section 4.2.3.

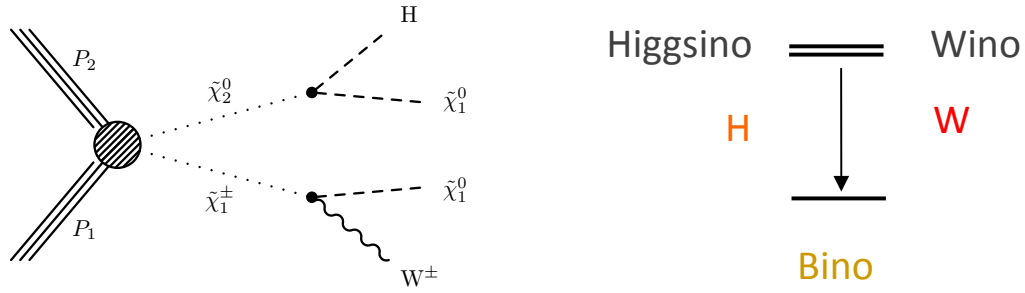


Figure 10.10: Feynman diagram and superpartner mass spectrum for the electroweak SUSY production of chargino-neutralino pairs for the wino NLSP with a Higgs boson scenario.

This process has been probed in previous CMS searches [135], which required that the neutralino decayed to a Z boson and $\tilde{\chi}_1^0$. Only recently has the decay to Higgs boson and $\tilde{\chi}_1^0$ been investigated [136, 137]. The wino NLSP with a Higgs boson scenario can produce multilepton final states, similarly to the natural higgsino NLSP scenario, if the Higgs boson decays to WW^* , ZZ^* , or $\tau\tau$, followed by leptonic decays of the W or Z boson.

We calculate the 95% CL exclusion limits on the production cross section and the results are presented in Figure 10.10. We expected to exclude the $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_1^0} = 130$ GeV masses for the chargino and neutralino, respectively, but our observed limit is slightly worse than anticipated. The most sensitive search channels for this scenario are those with exactly three leptons, no b-tagged jets, and low H_T .

The observed number of events, the expected SM backgrounds, and the predicted signal yields of the top five most sensitive search channels for the model point $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_1^0} =$

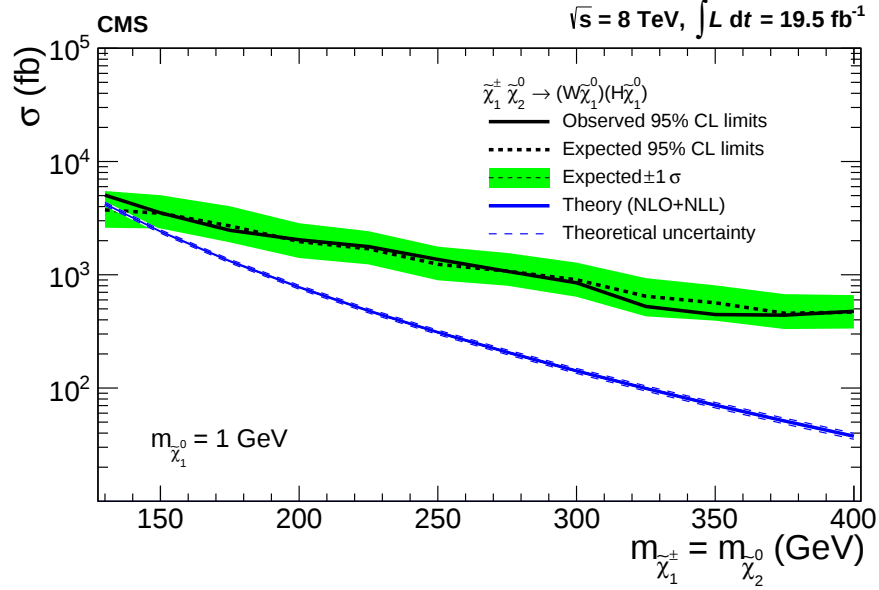


Figure 10.11: The 95% confidence level upper limits for the wino NLSP with a Higgs boson scenario on the $\tilde{\chi}_1^\pm \tilde{\chi}_1^0$ cross section times branching fraction $\mathcal{B}(\tilde{\chi}_1^\pm \tilde{\chi}_2^0 \rightarrow W^\pm H + \tilde{\chi}_1^0)$. The green band shows the $\pm 1\sigma$ variations on the expected limit due to experimental uncertainties.

130 GeV, $m_{\tilde{\chi}_1^0} = 1$ GeV, where the multilepton analysis has the best sensitivity, are presented in Table 10.3 with their corresponding E_T^{miss} distribution plots in Figure 10.12. The data yields in the channels are found to be broadly consistent with the expected SM backgrounds. Similar results tables for other model points are presented in Appendix D.

Table 10.3: Observed (Obs.) number of events and expected (Exp.) number of SM background events, along with the number of signal (Sig.) events, in the top five most sensitive search channels for the model point $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = 130$ GeV, and $m_{\tilde{\chi}_1^0} = 1$ GeV. All channels shown have exactly three selected leptons, a veto on tagged b-jets, and $H_T < 200$ GeV. The results are categorized by the presence of an OSSF lepton pair with invariant mass below 75 GeV (above 105 GeV) given by the “below-Z” (“above-Z”) designation, the E_T^{miss} in the event, and the number of τ_h candidates.

$m_{\ell+\ell-}$	E_T^{miss} (GeV)	N_{τ_h}	Obs.	Exp.	Sig.
Below-Z	(50, 100)	0	142	125 ± 28	24.4 ± 4.4
Below-Z	(100, 150)	0	16	21.3 ± 8.0	6.8 ± 1.2
—	(0, 50)	0	53	52 ± 12	8.7 ± 1.7
—	(50, 100)	0	35	38 ± 15	10.8 ± 2.0
—	(100, 150)	0	7	9.3 ± 4.3	3.37 ± 0.54

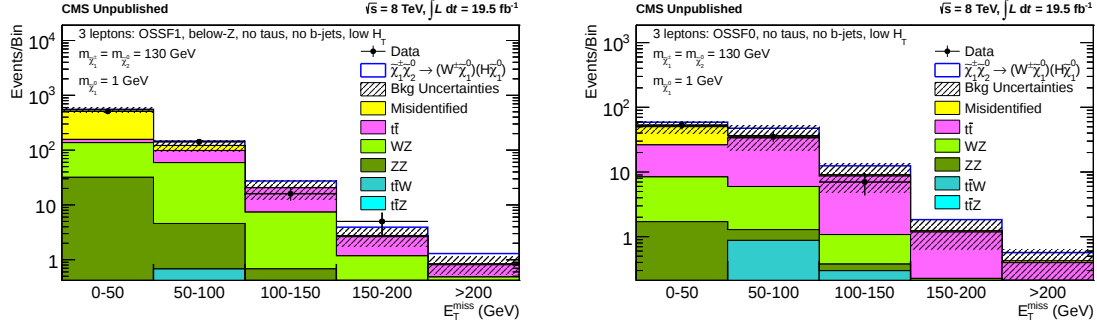


Figure 10.12: The E_T^{miss} distributions of the top five most sensitive search channels for the model point $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = 130$ GeV, and $m_{\tilde{\chi}_1^0} = 1$ GeV.

10.2.3 The slepton co-NLSP scenario

In supersymmetry models, multilepton final states arise naturally in GMSB with split messengers (GMSM) theories [138, 139, 140, 27]. An example of such a model is the slepton co-NLSP scenario, where we have mass-degenerate right-handed sleptons $\tilde{\ell}_R$ (selectron, smuon, and stau) that are at the bottom of the MSSM mass spectrum and server together as the NLSP [131, 132, 133]. Figure 10.13 provides the Feynman diagram and the superpartner mass spectrum for the process. Production proceeds mainly through pairs of squarks and/or gluinos but production through chargino-neutralino or right-handed slepton pairs is also possible, where the chargino is “wino-like”. The exact superpartner masses will determine the relative importance of the strong- and electroweak- production mechanisms. If production proceeds through a squark–squark pair state then the cascade decay is as follows,

$$\begin{aligned}
 & \tilde{q} \tilde{q} \\
 & q \tilde{\chi}_1^0 \leftarrow \rightarrow q \tilde{\chi}_1^0 \\
 & \ell^\mp \tilde{\ell}_R^\pm \leftarrow \rightarrow \ell^\mp \tilde{\ell}_R^\pm \\
 & \ell^\pm \tilde{G} \leftarrow \rightarrow \ell^\pm \tilde{G},
 \end{aligned}$$

where the neutralino $\tilde{\chi}_1^0$ is taken to be a bino, the superpartner of the B gauge boson, q and ℓ are quarks and leptons, respectively, and $\tilde{G}$ is the gravitino and the LSP, similar to the gravitino in the natural higgsino NLSP scenario except that it is almost massless, which leads to E_T^{miss} . The higgsinos are decoupled in this scenario, meaning they are much heavier than all other SUSY particles. Such scenarios have a large cross section with little SM background contribution [27].

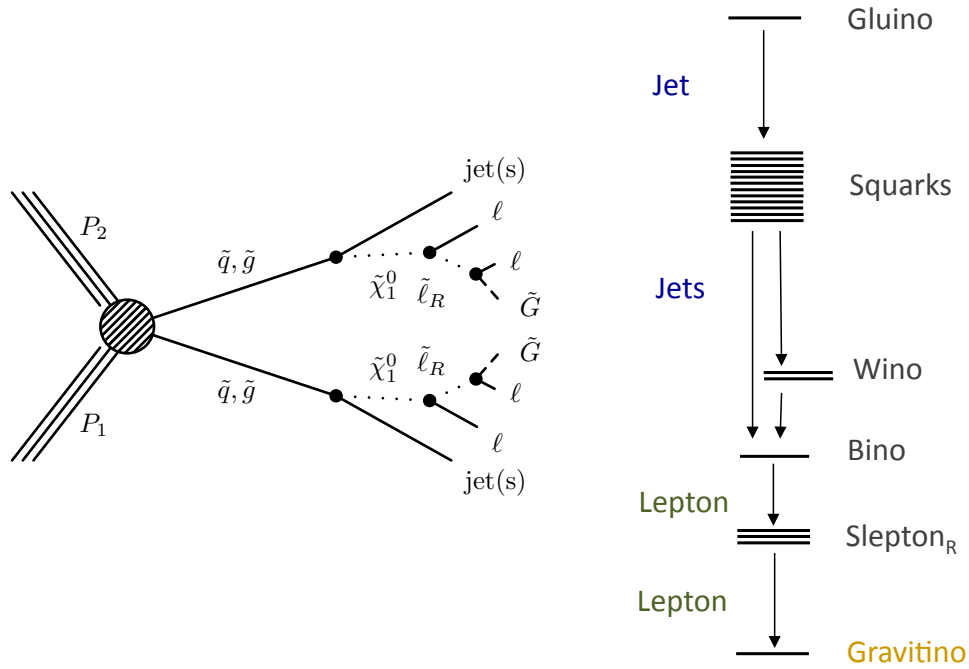


Figure 10.13: Feynman diagram and superpartner mass spectrum for the strong SUSY production of squark and/or gluino pairs for the slepton co-NLSP scenario.

Figure 10.14 shows the 95% CL exclusion limits for the slepton co-NLSP scenario in the gluino versus chargino mass plane. The region below and to the left of the solid black curve is excluded. In the parameter space dominated by strong production, the exclusion curve asymptotically approaches a horizontal line, while the exclusion curve tends towards a vertical line in the parameter space dominated by electroweak production.

We have the largest sensitivity for the slepton co-NLSP scenario in search channels with off-Z OSSF lepton pairs, no tagged b-jets, large E_T^{miss} , and depending on the mass spectrum it may or may not populate the large- H_T channels.

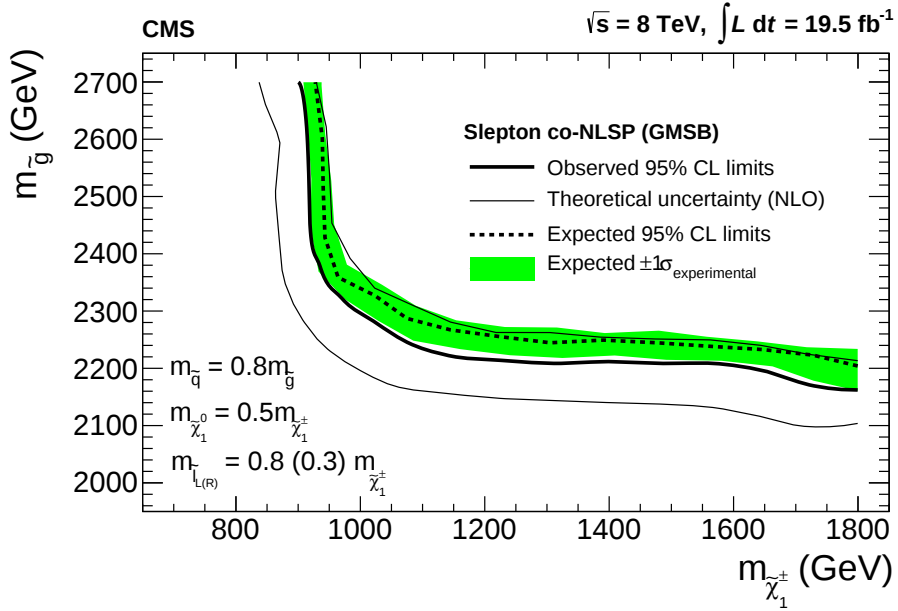


Figure 10.14: The 95% confidence level lower limits for the slepton co-NLSP scenario in the gluino versus chargino mass plane. Masses to the left of the solid black curve are excluded. The green band shows the $\pm 1\sigma$ variations on the expected limit due to experimental uncertainties.

10.2.4 The stau-(N)NLSP scenario

We now consider a model that involves two scenarios, namely the stau-NLSP and stau-next-to-next-to-lightest SUSY particle (NNLSP) scenarios. In the stau-NLSP scenario, the right-handed tau slepton $\tilde{\tau}_R$ is the NLSP and the mass-degenerate right-handed selectron $\tilde{e}_R$ and smuon $\tilde{\mu}_R$ are co-NLSPs and much heavier, while in the stau-NNLSP scenario the converse is true. Both scenarios together are referred to as the stau-(N)NLSP scenario [131, 132, 133]. The Feynman diagram and superpartner mass spectrum are presented in Figure 10.15 for both processes. This scenario arises for moderate to large values of the MSSM parameter $\tan \beta$ [21, 22], where $\tan \beta$ is the ratio of the vacuum expectation values of the two Higgs

doublets of the MSSM. The process proceeds through electroweak production of selectron–anti-selectron or smuon–anti-smuon pairs, which then undergo three-body decay processes. The following illustrates the cascade decay for selectron–anti-selectron pair production,

$$\begin{aligned} & \tilde{e}_R^+ \tilde{e}_R^- \\ & e^+ \tau^+ \tilde{\tau}_R^- \leftarrow \rightarrow e^- \tau^- \tilde{\tau}_R^+ \\ & \tau^- \tilde{G} \leftarrow \rightarrow \tau^+ \tilde{G}, \end{aligned}$$

where $\tilde{G}$ is the gravitino, which has the same properties as the one in the slepton co-NLSP scenario. The smuon–anti-smuon pair production undergoes a similar cascade decay. In the stau-NNLSP scenario we have electroweak pair production of staus. The staus then decay to the NLSP and a τ lepton. Lastly, the NLSPs decay to a light lepton and gravitino.

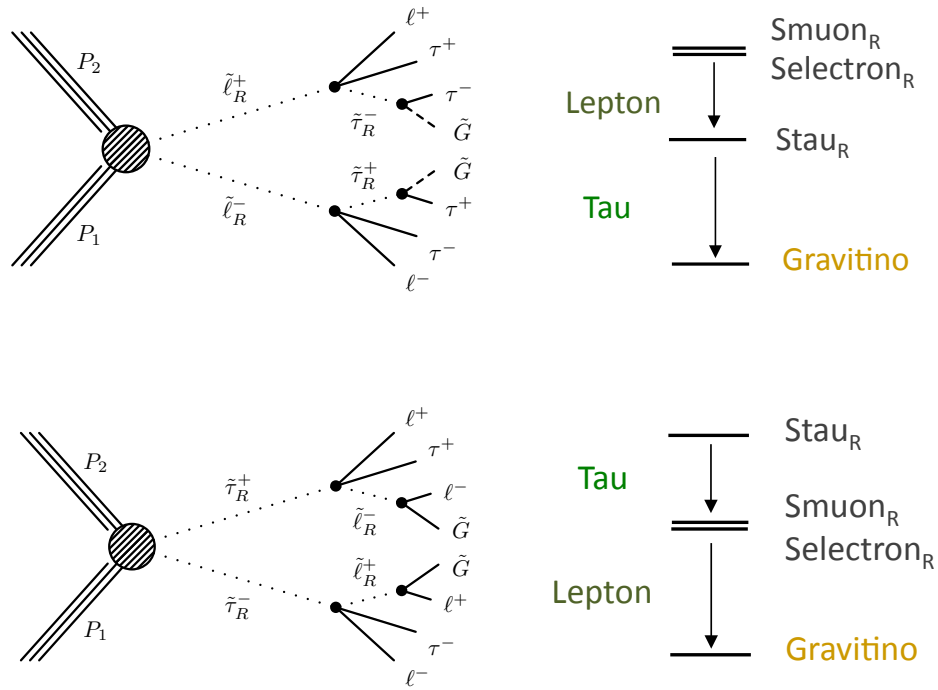


Figure 10.15: Feynman diagrams of right-handed smuon or right-handed selectron pair production (top left) and stau pair production (top right) in proton-proton collisions followed by decays leading to a final state with four leptons along with LSPs along with the a schematic superpartner mass spectrum for the GMSB stau-(N)NLSP scenario (bottom).

The 95% CL exclusion limits for the stau-(N)NLSP scenario in the $m_{\tilde{\tau}}$ versus $m_{\tilde{e}} = m_{\tilde{\mu}}$ mass plane are shown in Figure 10.16. The region above the diagonal line corresponds to the stau-NNLSP scenario, while the region below the diagonal line belongs to the stau-NLSP scenario. The search channels with the largest sensitivity for this scenarios contain off-Z OSSF lepton pairs, τ_h leptons, no tagged b-jets, and large E_T^{miss} .

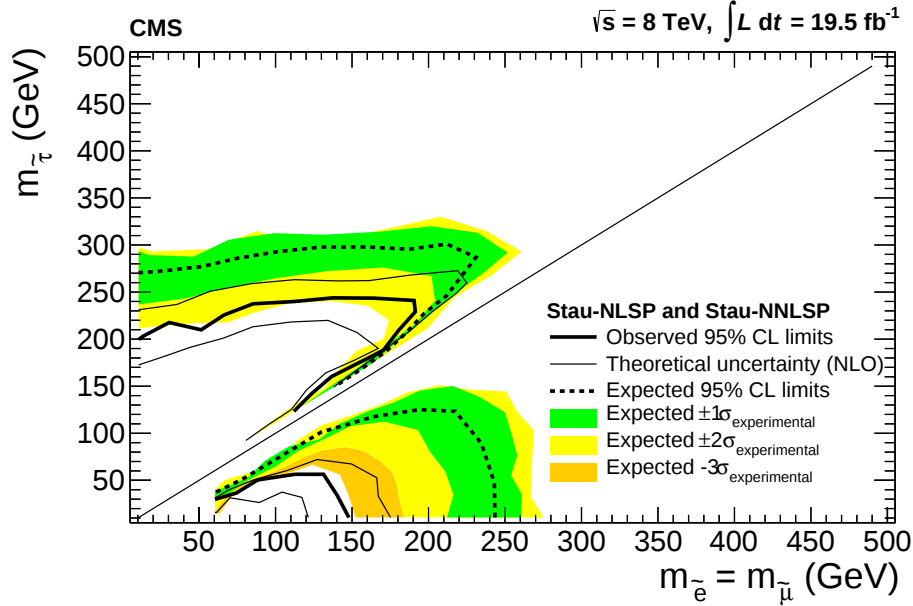


Figure 10.16: The 95% confidence level lower limits for the stau-NLSP (below diagonal line) and stau-NNLSP (above diagonal line) scenarios in the degenerate smuon and selectron versus stau mass plane. Masses to the left of the solid black curve are excluded. The green and yellow bands show the $\pm 1\sigma$ and $\pm 2\sigma$ variations on the expected limit due to experimental uncertainties, respectively.

In the stau-NLSP region there are four τ -leptons and two light leptons in the final state, whereas in the stau-NNLSP region there are four light leptons and two τ -leptons in the final state. As the curve approaches the diagonal line from either region, the mass difference between the stau and light sleptons decreases. Therefore within both regions, the p_T of two light leptons and two τ -leptons become very soft, resulting in a low signal efficiency and causing the exclusion curve to become nearly parallel to the diagonal from both sides. Within the stau-NNLSP mass region, the sensitivity above the diagonal line covers larger masses because light-leptons have higher reconstruction efficiencies and lower

background contributions compared to those of τ_h -leptons. The discrepancy between the observed and expected 95% CL limits in the stau-NLSP region is driven by the excesses in the four-or-more leptons, OSSF1, off-Z category with one τ_h -lepton candidate, no tagged b-jet, and $H_T < 200$ GeV mentioned in Section 10.1.

10.2.5 The third-generation SMS T1tttt scenario

A simplified model spectra (SMS) of distinct interest for its large number of b jets is the T1tttt scenario [27, 141, 142, 131, 132, 133]. This scenario involves pair production of gluinos $\tilde{g}$ with the squarks decoupled. The Feynman diagram and the mass spectrum for the process are shown in Figure 10.17. The cascaded decay is as follows,

$$\begin{aligned} & \tilde{g} \tilde{g} \\ & \tilde{t}\tilde{t} \leftarrow \quad \rightarrow \tilde{t}\tilde{t}^* \\ & t\tilde{\chi}_1^0 \leftarrow \quad \rightarrow \bar{t}\tilde{\chi}_1^0, \end{aligned}$$

where $\tilde{\chi}_1^0$ is the LSP, and the top squarks are off-shell and immediately decay to a top quark and the LSP. Therefore each gluino undergoes an effective three-body decay to two top quarks and the LSP.

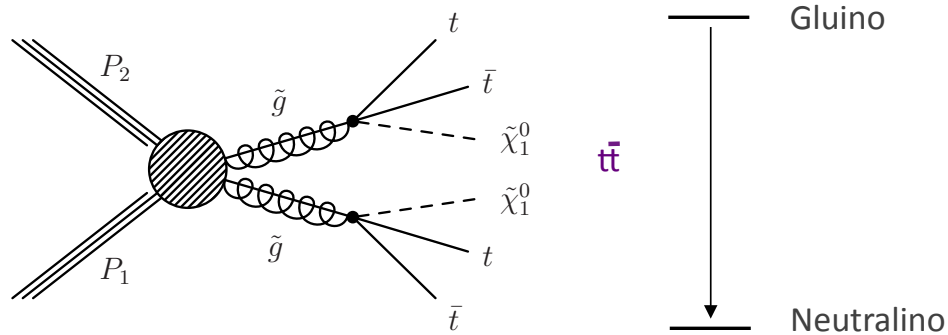


Figure 10.17: Feynman diagram and superpartner mass spectrum for the SMS T1tttt scenario.

The T1tttt scenario has four top quarks, which may result in up to four tagged b-jets once they decay and in a multilepton final state for leptonically decaying W bosons. In addition, due to the large number of jets the model produces large H_T , as well as large E_T^{miss} coming from the neutralinos. Events with multiple b-jets have little background contribution from SM processes.

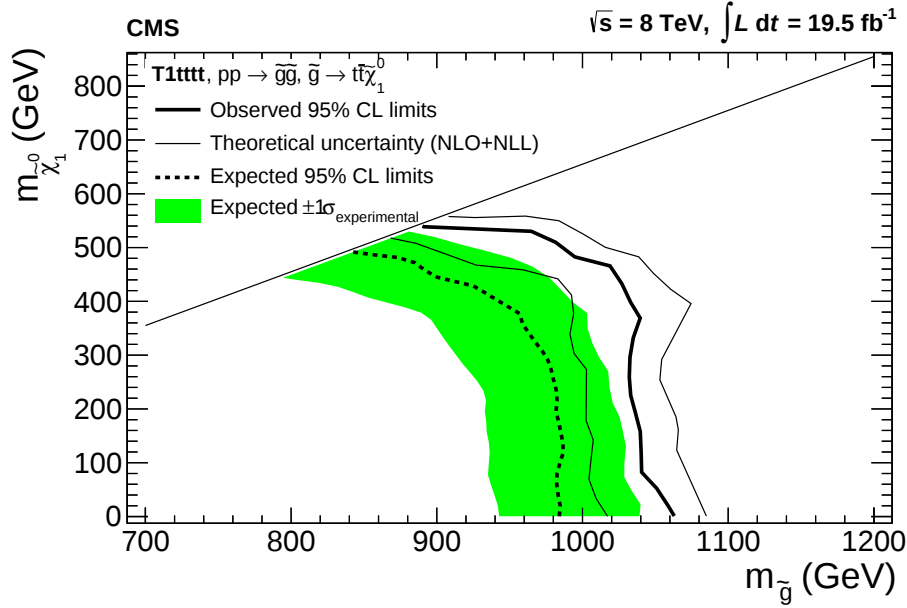


Figure 10.18: The 95% confidence level lower limits for the T1tttt scenario in the gluino versus neutralino mass plane. Masses to the left of the solid black curve and below the diagonal line are excluded. The green band shows the $\pm 1\sigma$ variations on the expected limit due to experimental uncertainties.

The 95% CL exclusion limits for the T1tttt scenario in the $m_{\tilde{g}}$ versus $m_{\tilde{\chi}_1^0}$ mass parameter space are shown in Figure 10.18. The T1tttt scenario populates channels with tagged b-jets with a preference for the high- H_T channels. We are able to exclude gluinos masses below 1 TeV.

10.2.6 The third-generation SMS T6ttWW scenario

We conclude the interpretation of the multilepton results in the context of MSSM models with the third-generation SMS T6ttWW scenario [143, 142, 131, 132, 133]. The process proceeds through direct bottom squark–anti-bottom squark pair production with the following

cascade decay,

$$\begin{aligned}
 & \tilde{b} \tilde{b}^* \\
 & \quad \downarrow \quad \quad \quad \downarrow \\
 & t \tilde{\chi}_1^- \quad \leftarrow \quad \rightarrow \quad \bar{t} \tilde{\chi}_1^+ \\
 & \quad \downarrow \quad \quad \quad \downarrow \\
 & W^- \tilde{\chi}_1^0 \quad \leftarrow \quad \rightarrow \quad W^+ \tilde{\chi}_1^0,
 \end{aligned}$$

where the neutralino $\tilde{\chi}_1^0$ is assumed to be 50 GeV, and the charginos $\tilde{\chi}_1^\pm$ are on-shell as a simplifying assumption, such that the W boson from its decay may either be on-shell or off-shell. Figure 10.19 shows the Feynman diagram and the superpartner mass spectrum.

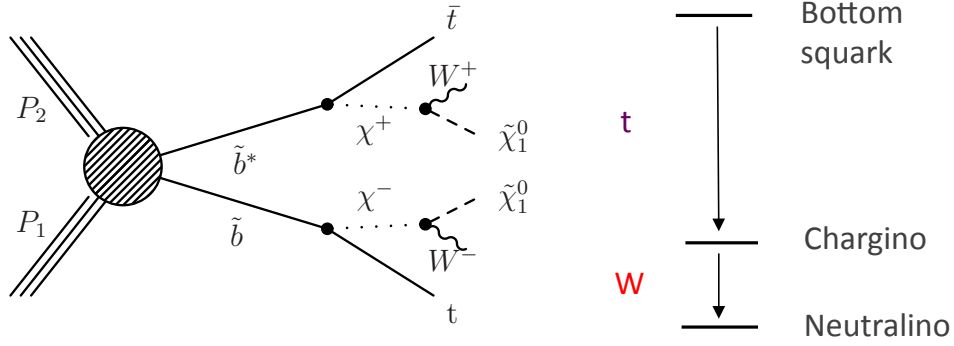


Figure 10.19: Feynman diagram and schematic superpartner mass spectrum for the SMS T6ttWW scenario.

Figure 10.20 shows the 95% CL exclusion limits for the T6ttWW scenario in the $m_{\tilde{\chi}_1^\pm}$ versus $m_{\tilde{b}}$ mass parameter space. The T6ttWW scenario populates channels with tagged b-jets with a preference for low- H_T channels. We are then able to exclude bottom squark masses with values less than 550 GeV.

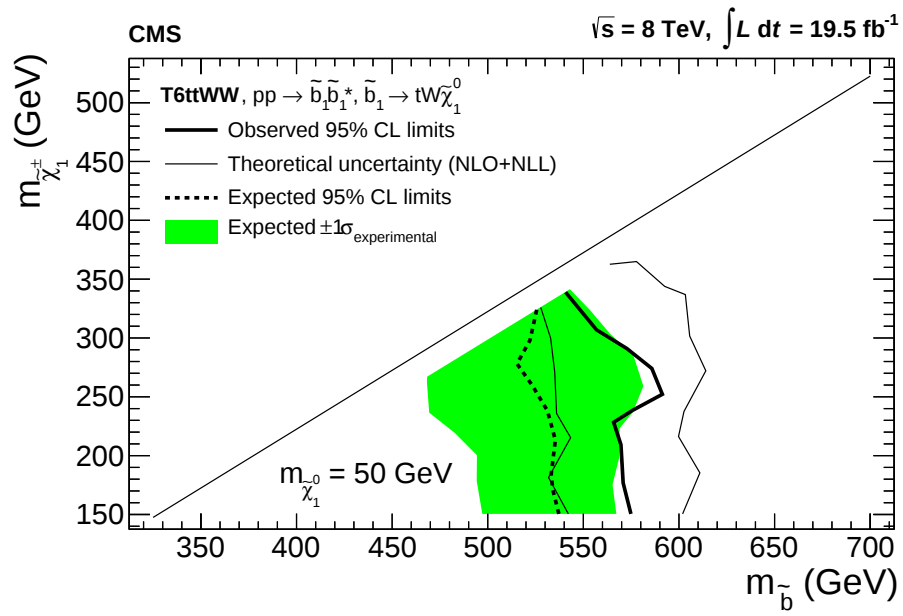


Figure 10.20: The 95% confidence level lower limits for the T6ttWW scenario in the bottom squark versus chargino mass. Masses to the left of the solid black curve and below the diagonal line are excluded. The green band shows the $\pm 1\sigma$ variations on the expected limit due to experimental uncertainties.

10.3 Conclusion

In this dissertation, a search for physics beyond the SM which manifests itself in a variety of possible multilepton final states has been performed using a total integrated luminosity of 19.5 fb^{-1} of pp collision data at a center-of-mass energy of $\sqrt{s} = 8 \text{ TeV}$ collected by the CMS detector at the LHC during 2012. We have estimated the background from SM processes using both MC simulations and data-derived methods. Simulated backgrounds were corrected for trigger, lepton, and b-jet efficiencies, pileup, and $E_{\text{T}}^{\text{miss}}$ resolution. Backgrounds from electrons and muons coming from b jets, jets faking hadronically decaying τ -leptons, and asymmetric internal conversion of photons into hard and soft leptons, where the soft leptons are not reconstructed, were estimated using data-derived techniques. The search was carried out by selecting and classifying events into multiple exclusive channels. We saw an overall good agreement between observations and expectations from SM processes. We utilized several MSSM SUSY scenarios as benchmarks for new physics with multiple leptons in the final states, and we were able to probe new regions of GMSB space, such as the natural Higgsino NLSP, slepton co-NLSP, and stau-(N)NLSP models, and a few SMS scenarios like T1tttt , T6ttWW , and electroweak SUSY production of charginos and neutralinos in final states with a Higgs boson. In the absence of any discovery of new physics we placed 95% CL lower limits on supersymmetric mass spectrum and upper limits on cross sections and branching fractions. The models span a range of parameters, and the limits on intermediate and final state masses illustrate the reach of this analysis. Finally, the sensitivity of this broad open search can be applied to many other models not addressed in this thesis.

Appendix A

Observation and SM Background Plots

The Figures below show the observation and background E_T^{miss} distributions of the analysis for the three lepton and four or more lepton search channels.

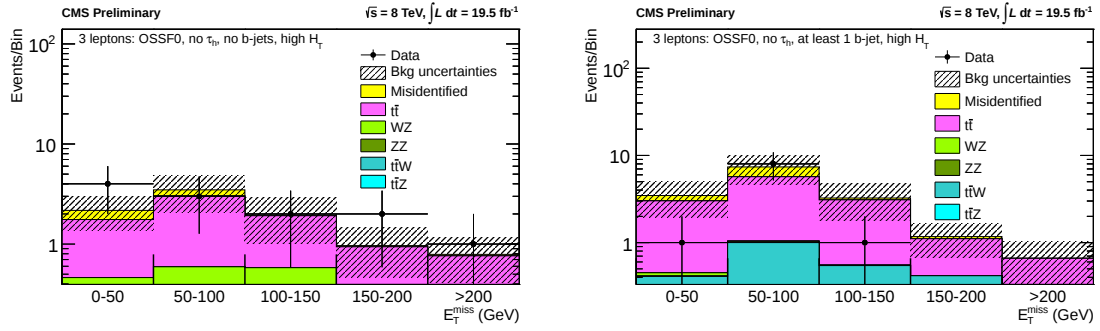


Figure A.1: 3-leptons + OSSFO + Tau0 + $H_T > 200 \text{ GeV}$ and 0 b-jets (left) or ≥ 1 b-jet (right)

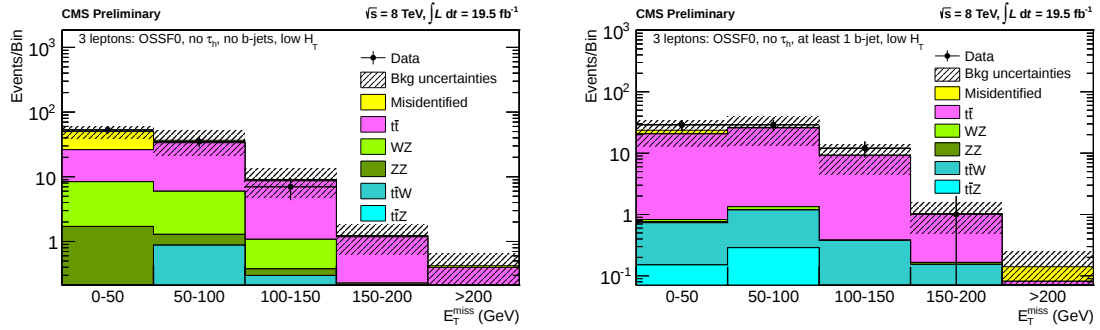


Figure A.2: 3-leptons + OSSFO + Tau0 + $H_T < 200 \text{ GeV}$ and 0 b-jets (left) or ≥ 1 b-jet (right)

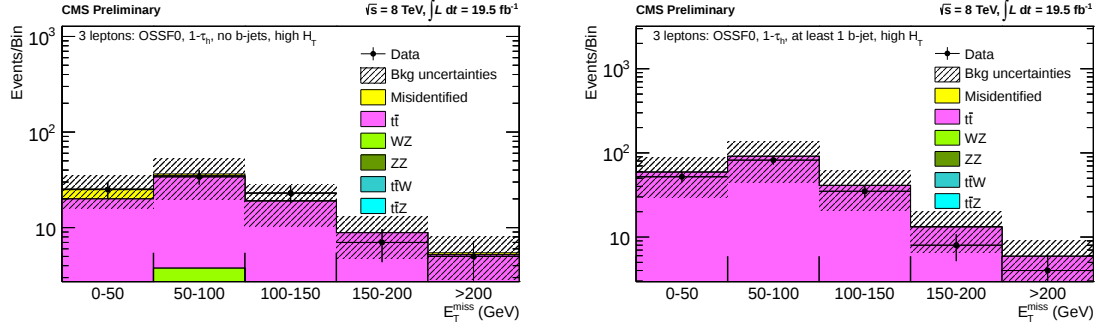


Figure A.3: 3-leptons + OSSF0 + Tau1 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

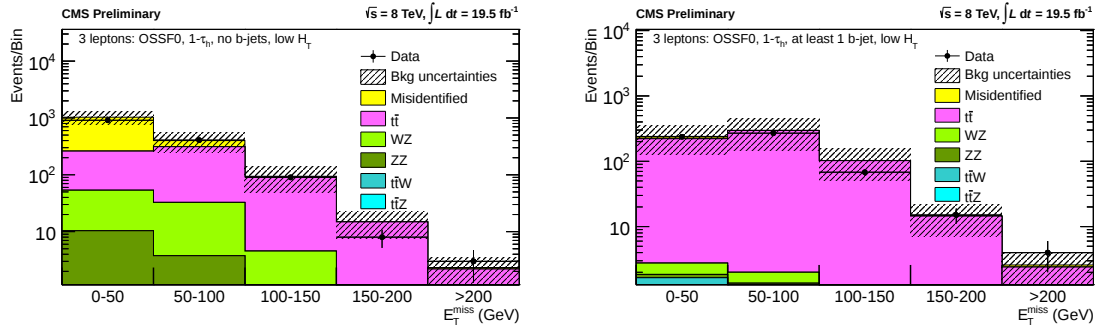


Figure A.4: 3-leptons + OSSF0 + Tau1 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

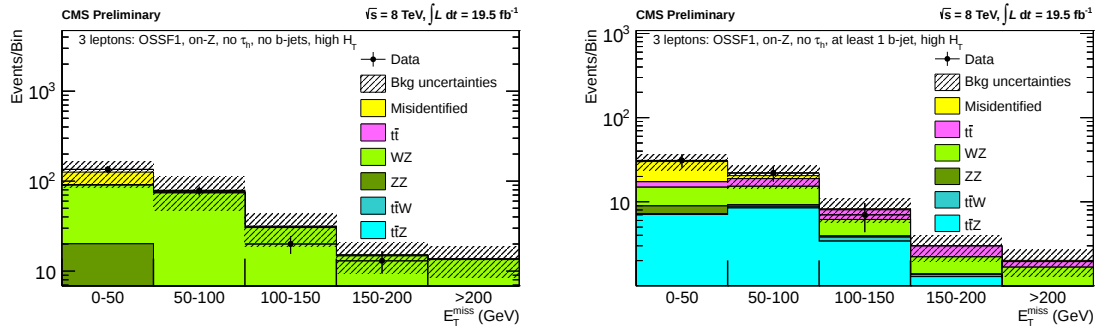


Figure A.5: 3-leptons + OSSF1 + on-Z + Tau0 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

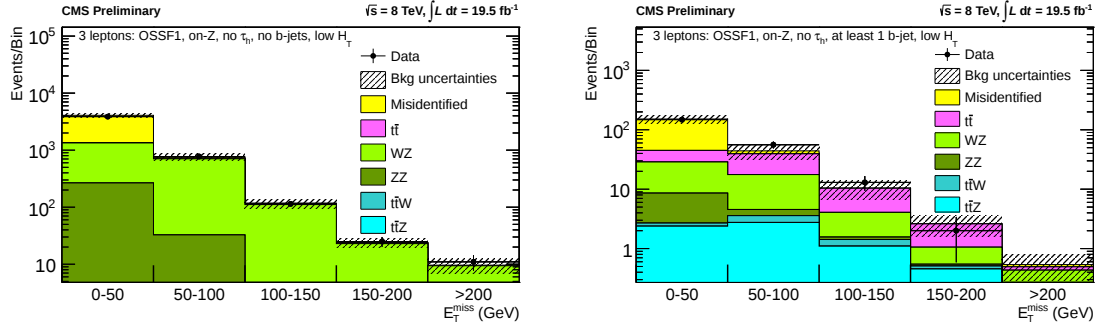


Figure A.6: 3-leptons + OSSF1 + on-Z + Tau0 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

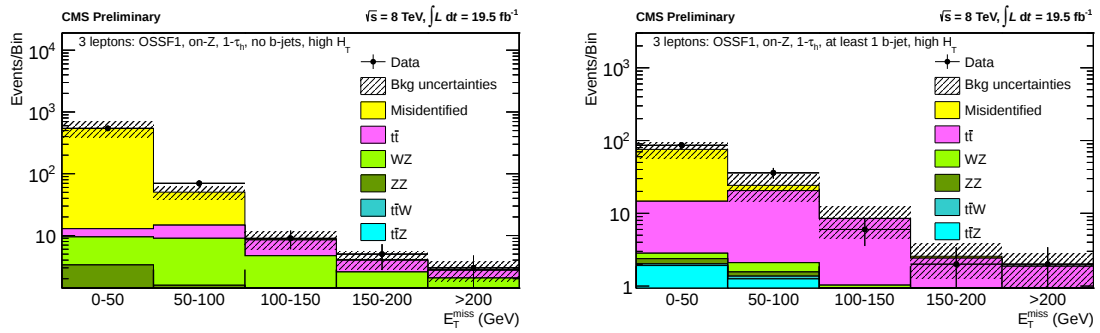


Figure A.7: 3-leptons + OSSF1 + on-Z + Tau1 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

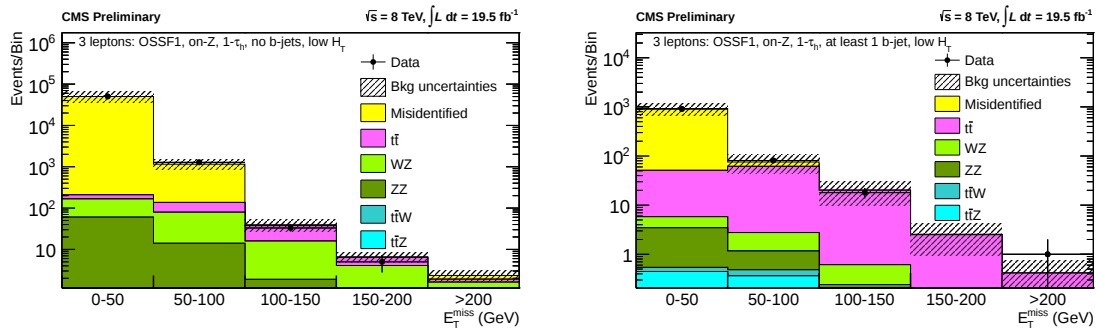


Figure A.8: 3-leptons + OSSF1 + on-Z + Tau1 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

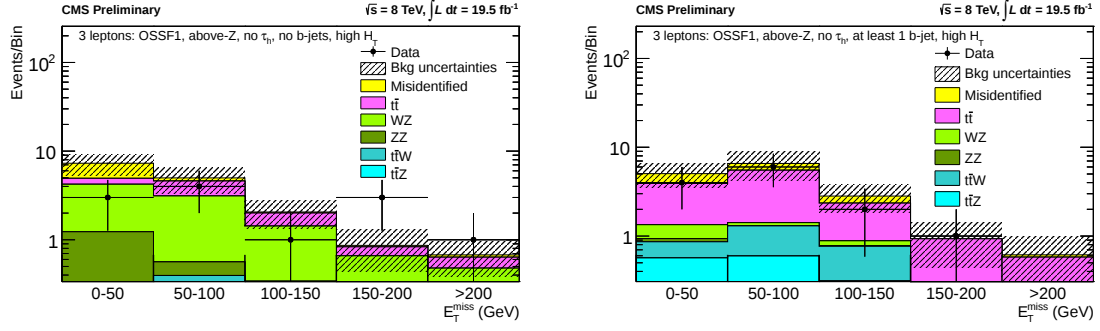


Figure A.9: 3-leptons + OSSF1 + above-Z + Tau0 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

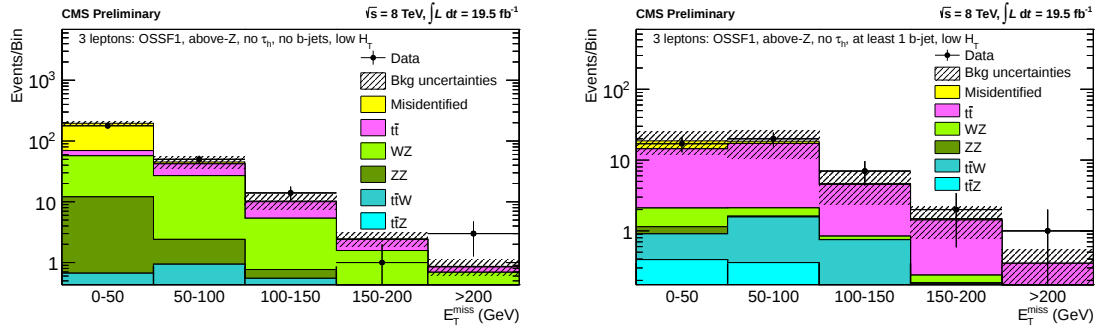


Figure A.10: 3-leptons + OSSF1 + above-Z + Tau0 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

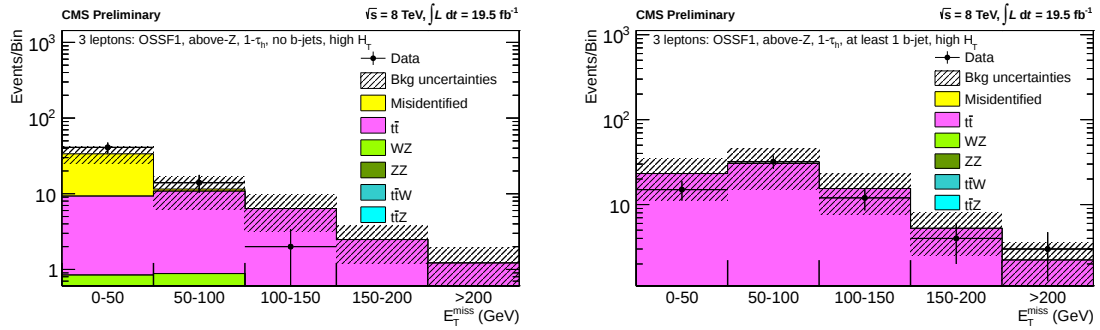


Figure A.11: 3-leptons + OSSF1 + above-Z + Tau1 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

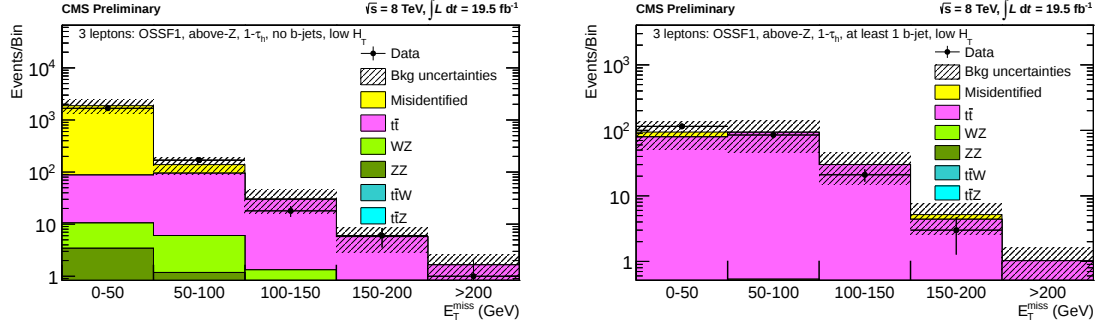


Figure A.12: 3-leptons + OSSF1 + above-Z + Tau1 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

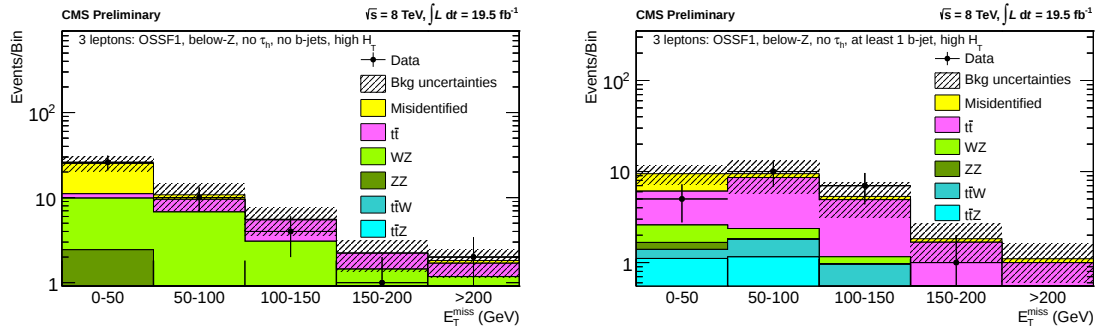


Figure A.13: 3-leptons + OSSF1 + below-Z + Tau0 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

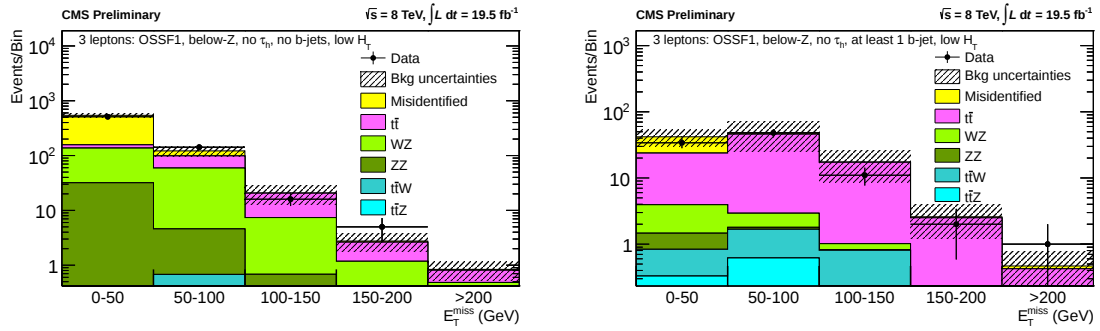


Figure A.14: 3-leptons + OSSF1 + below-Z + Tau0 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

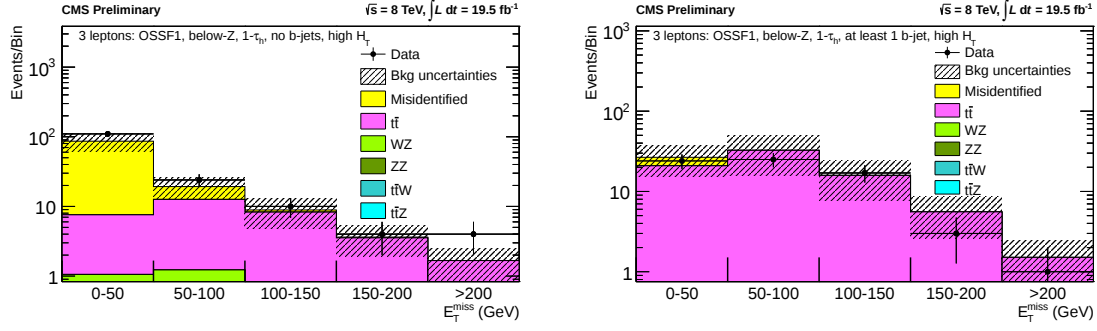


Figure A.15: 3-leptons + OSSF1 + below-Z + Tau1 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

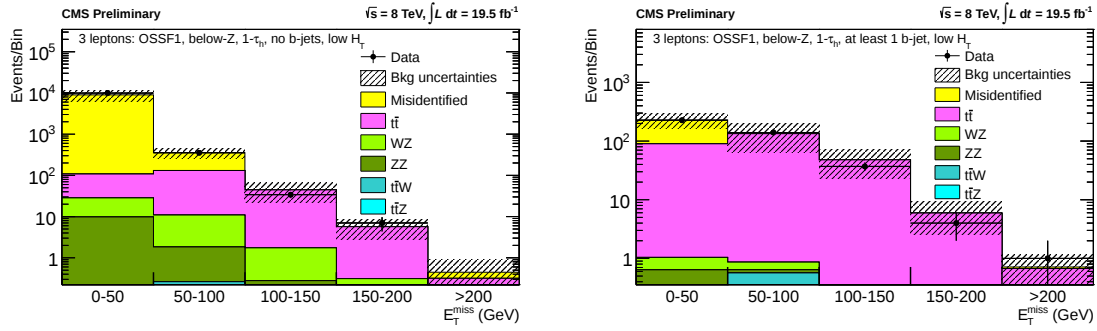


Figure A.16: 3-leptons + OSSF1 + below-Z + Tau1 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

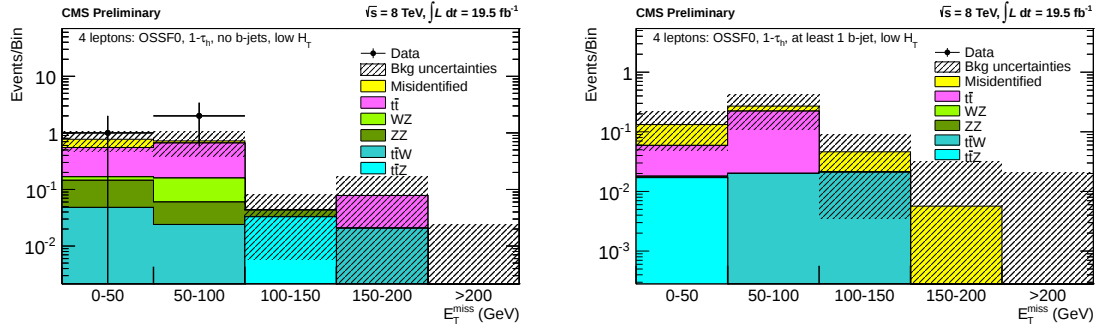


Figure A.17: 4-leptons + OSSF0 + Tau1 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

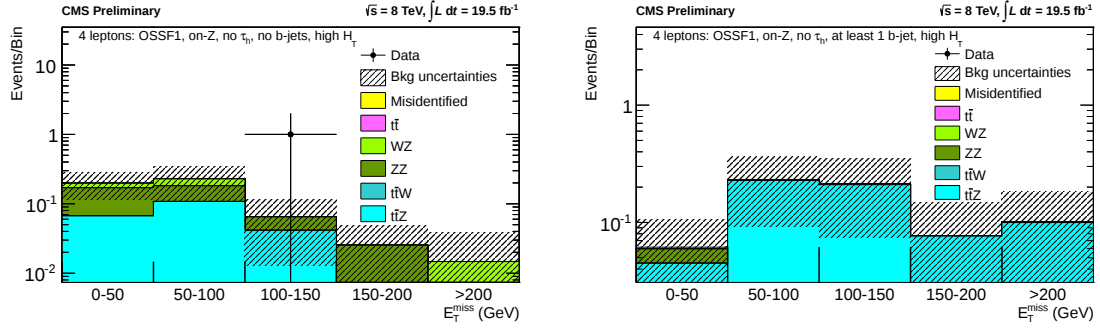


Figure A.18: 4-leptons + OSSF1 + on-Z + Tau0 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

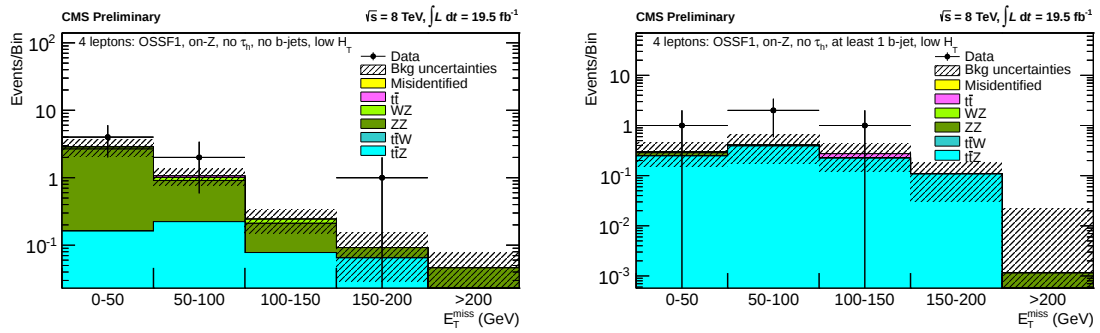


Figure A.19: 4-leptons + OSSF1 + on-Z + Tau0 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

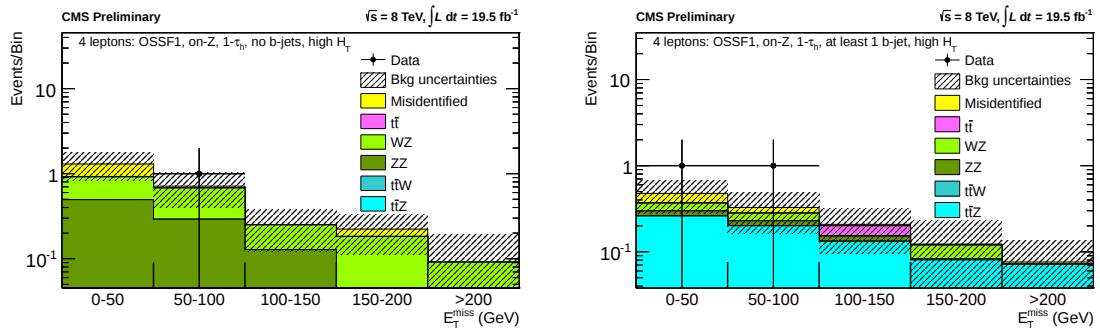


Figure A.20: 4-leptons + OSSF1 + on-Z + Tau1 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

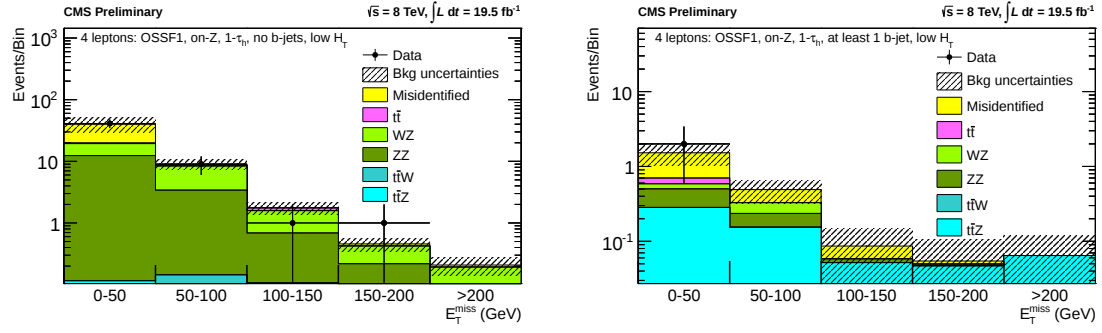


Figure A.21: 4-leptons + OSSF1 + on-Z + Tau1 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

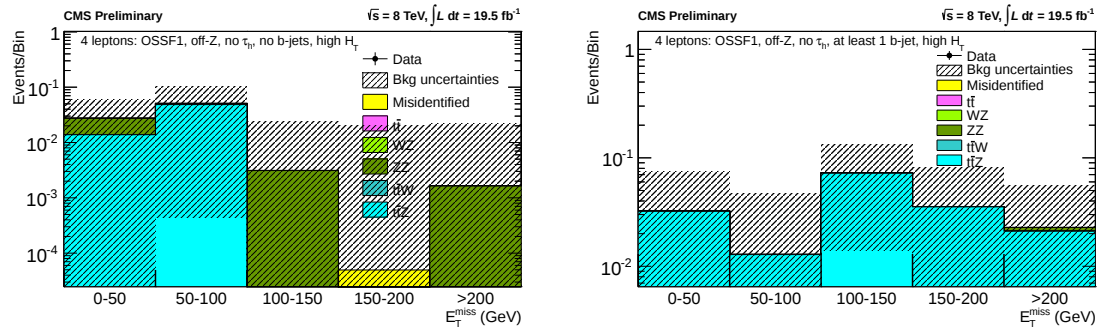


Figure A.22: 4-leptons + OSSF1 + off-Z + Tau0 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

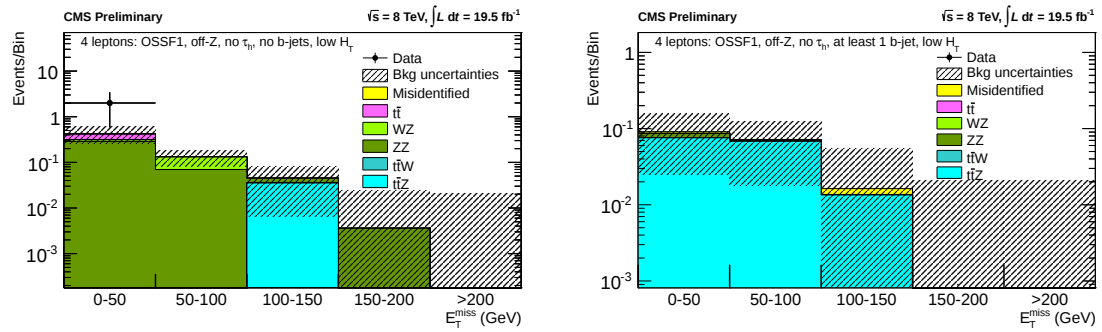


Figure A.23: 4-leptons + OSSF1 + off-Z + Tau0 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

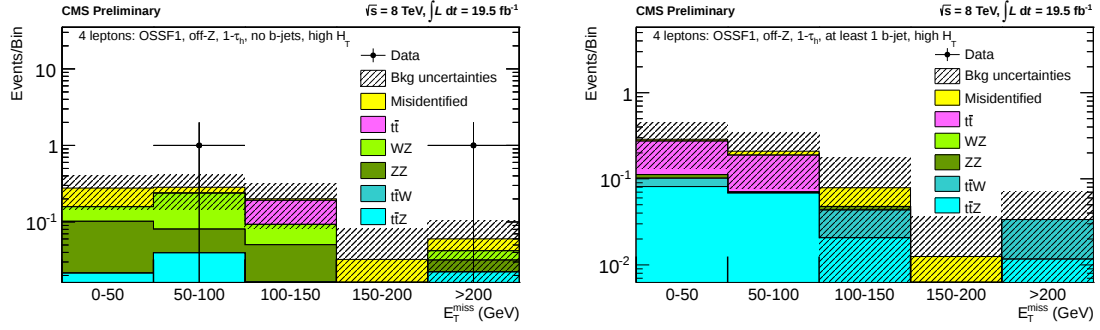


Figure A.24: 4-leptons + OSSF1 + off-Z + Tau1 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

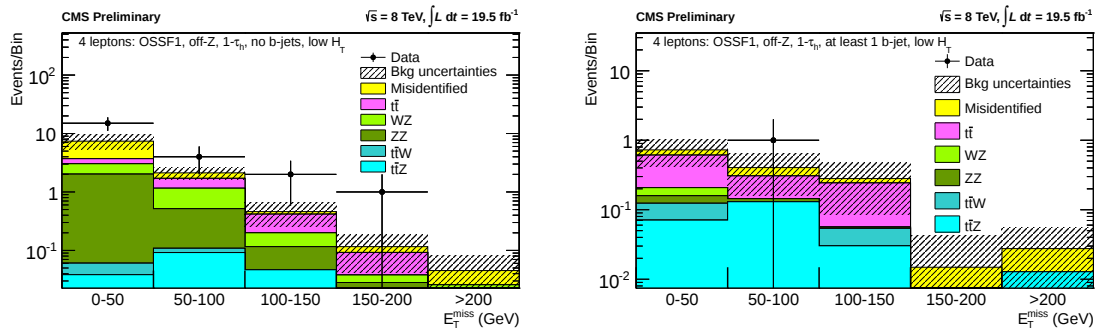


Figure A.25: 4-leptons + OSSF1 + off-Z + Tau1 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

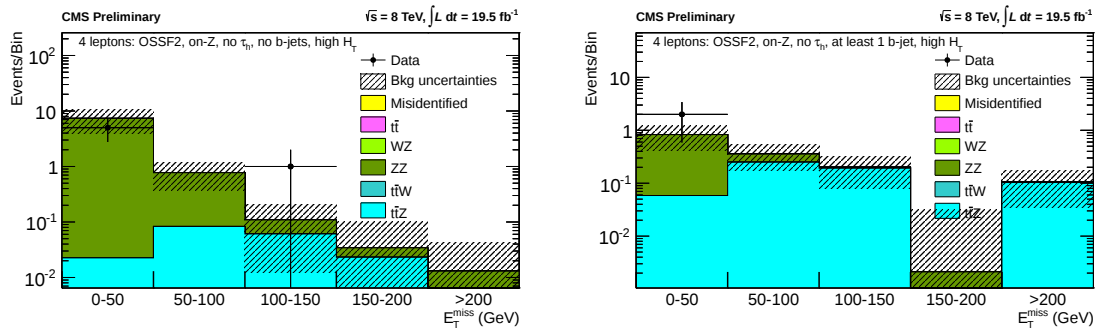


Figure A.26: 4-leptons + OSSF2 + on-Z + Tau0 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

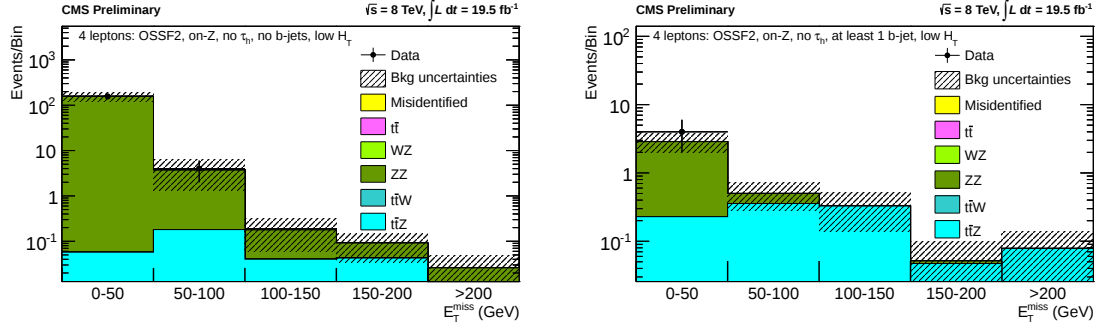


Figure A.27: 4-leptons + OSSF2 + on-Z + Tau0 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

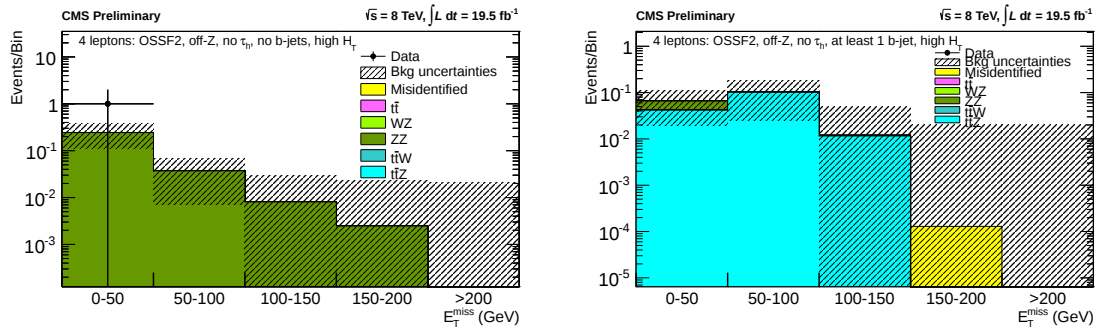


Figure A.28: 4-leptons + OSSF2 + off-Z + Tau0 + $H_T > 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

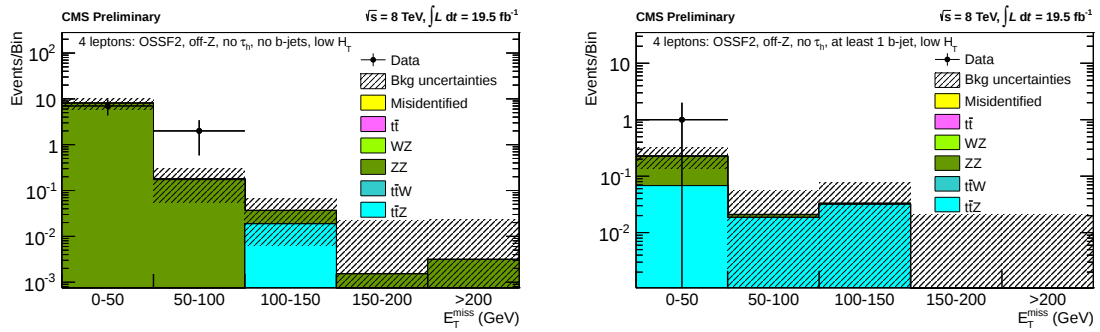


Figure A.29: 4-leptons + OSSF2 + off-Z + Tau0 + $H_T < 200$ GeV and 0 b-jets (left) or ≥ 1 b-jet (right)

Appendix B

Trigger List

We present all the un-prescaled triggers the analysis used, where “v*” indicates the various versions of the same trigger. The different versions of the same trigger are due to the minor updates the triggers undergo during the data taking period.

DoubleMu

- HLT_Mu17_Mu8_v*

DoubleElectron

- HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*

MuEG

- HLT_Mu8_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*
- HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*

SingleMu

- HLT_Mu40_eta2p1_v*
- HLT_Mu50_eta2p1_v*
- HLT_IsoMu24_eta2p1_v*
- HLT_IsoMu30_eta2p1_v*
- HLT_IsoMu34_eta2p1_v*

- HLT_IsoMu40_eta2p1_v*

SingleElectron

- HLT_Ele80_CaloIdVT_TrkIdT_v*
- HLT_Ele100_CaloIdVT_TrkIdT_v*
- HLT_Ele90_CaloIdVT_GsfTrkIdT_v*

Appendix C

Study of excess

If there are enough events we can distinguish a statistical fluctuation from new physics by examining different distributions and properties of the events and compare them to the background expectations. If the distributions and properties are consistent with the background prediction, and the only discrepancy is the total number of events, then we can say the excess is consistent with a statistical fluctuation. On the other hand, if the distributions are very different from the background expectations, then that could point to a signal.

Table C.1: Most discrepant channels for stau-NLSP scenario for the grid point with smuon/selectron mass of 150 GeV and stau mass of 50 GeV (lower half) on the observed lower exclusion contour. All channels have no b jets and $H_T < 200$ GeV.

N_ℓ	OSSF n	N_{τ_h}	$m_{\ell+\ell-}$	E_T^{miss} (GeV)	Obs.	Exp.	Sig.
4	OSSF1	1	Off-Z	(50, 100)	4	2.2 ± 0.5	5.2 ± 1.1
4	OSSF1	0	Off-Z	(0, 50)	2	0.5 ± 0.2	3.2 ± 0.9
4	OSSF1	1	Off-Z	(0, 50)	15	7.5 ± 2.0	4.9 ± 1.2
4	OSSF2	0	Off-Z	(50, 100)	2	0.18 ± 0.14	0.94 ± 0.32
4	OSSF1	1	Off-Z	(100, 150)	2	0.5 ± 0.2	1.3 ± 0.4

Table C.2: Most discrepant channels for stau-NNLSP for the grid point with smuon/selectron mass of 170 GeV and stau mass of 230 GeV (upper half) on the observed upper exclusion contour. All channels have no b jets and $H_T < 200$ GeV.

N_ℓ	OSSF n	N_{τ_h}	$m_{\ell+\ell-}$	E_T^{miss} (GeV)	Obs.	Exp.	Sig.
4	OSSF2	0	Off-Z	(50, 100)	2	0.11 ± 0.09	0.78 ± 0.13
3	OSSF1	0	Below-Z	(150, 200)	3	1.79 ± 0.56	0.62 ± 0.10
4	OSSF1	1	Off-Z	(100, 150)	2	0.40 ± 0.14	0.26 ± 0.06
4	OSSF1	0	On-Z	(50, 100)	2	0.90 ± 0.26	0.30 ± 0.08
3	OSSF1	0	Above-Z	(100, 150)	11	7.1 ± 1.5	0.46 ± 0.08

Tables C.1 and C.2 show the most discrepant channels for a point on the observed

stau-NLSP and stau>NNLSP exclusion limits, respectively. Several studies are performed in order to investigate the excess observed in the stau-NLSP scenario, more specifically, the background estimation is broken down by its sources of SM contributions, the observed number of events is broken down by lepton flavor, we check the total lepton charge, we study the impact of a systematic miss-estimation on the other channels, we study the effect of allowing each background to float, and, lastly, we plot the H_T , electron p_T , muon p_T , and τ_h -lepton p_T distributions of the excess. All studies are consistent with an upward fluctuation. Further details can be found in Reference [116].

C.1 Breakdown of background estimation by SM contribution

There are a total of 10.3 expected events in the 4 lepton, OSSF1, off-Z, 1 τ_h -lepton, no b jets, $H_T < 200$ GeV, and E_T^{miss} between 0 – 50 GeV channel, see Fig. 10.1. Table C.3 shows the breakdown, by the SM contribution, of the background estimation for this channel.

Table C.3: SM background breakdown for the 4-lepton channel with OSSF1, off-Z, 1 τ_h -lepton, no b jets, $H_T < 200$ GeV, and E_T^{miss} between 0 – 50 GeV.

Background type	Exp.
Misidentified	4.0
$t\bar{t}$	1.5
WZ	1.8
ZZ	2.5
$t\bar{t}Z$	0.2
$t\bar{t}W$	0.04
SM Higgs boson	0.3
tbZ	0.03

C.2 Lepton flavor breakdown of observations

Table C.4 shows the observed number of events, for the 4 lepton, OSSF1, off-Z, 1 τ_h -lepton, no b jets, $H_T < 200$ GeV, and E_T^{miss} between 0 – 50 GeV channel, broken down by lepton flavor.

Table C.4: Flavor breakdown of the 4-lepton channel with OSSF1, off-Z, 1 τ_h -lepton, no b jets, $H_T < 200$ GeV, and E_T^{miss} between 0 – 50 GeV.

Flavor	Obs.	Exp.	$2.1 \times \text{Exp.}$
$3\mu, 1\tau_h$	5	2.5	5.3
$2\mu, 1e, 1\tau_h$	11	4.3	9.2
$1\mu, 2e, 1\tau_h$	2	1.9	4.0
$3e, 1\tau_h$	4	1.6	3.4

C.2.1 Check total charge of leptons

- Total charge expectations:

SM Backgrounds with real τ , $|Q| = 0$

SM Backgrounds with fake τ , $|Q| = 0$ or 2 (50%/50%)

- For large statistics charge breakdown could help indicate identity of excess.

Table C.5: Check total charge of leptons. Need much more collision data to distinguish between the different cases based on charges.

Total charge	Obs.	Statistical fluctuation	Real Tau excess	Fake tau excess
$Q = 0$	16	14.1	18.3	12.4
$Q = \pm 2$	6	7.9	3.7	9.6

C.2.2 Systematic misestimation

We investigated how other channels would be affected if this were a systematic misestimation.

- Misidentified background:

Use f_{SB} to check for fake rate changes. No indication of rate changes between $3\ell + \tau_h$ and $2\ell + \tau_h$ or between on-Z and off-Z. At most events migrate by 20-30%.

We would need 400% migration to fill in the excess with backgrounds from fakes.

- $t\bar{t}$ background:

MC matches data well in $e^+\mu^-\tau_h$ channels (Obs. = 148, Exp. =162). We would need to require 1000% (10 times) the fake tau rate.

- ZZ background:

The tau efficiencies would have to be underestimated by factor of 6.

C.2.3 Allow background sources to float

We float the scale factor for each background ($t\bar{t}$, WZ, ZZ, misidentified) individually while keeping the other backgrounds fixed. We then proceeded to minimize negative log-likelihood treating all channels independently to get best fit for the scale factor. A scale factor of 1 corresponds to scaling to cross-section. We summarize the results below.

- Floating misidentified background:

Scale factor that minimizes negative log-likelihood: 101%

Statistical uncertainty on scale factor: 0.4%

- Floating $t\bar{t}$:

Scale factor that minimizes negative log-likelihood : 94%

Statistical uncertainty on scale factor: 2%

- Floating WZ:

Scale factor that minimizes negative log-likelihood : 95.2%

Statistical uncertainty on scale factor: 2.7%

- Floating ZZ:

Scale factor that minimizes negative log-likelihood : 96.4%

Statistical uncertainty on scale factor: 6.4%

C.3 Expected number of channels with deviations

The channel with $3(e/\mu)+1\tau_h$ and no Z candidate has an 2.5σ upward fluctuation in all E_T^{miss} ranges. We ask the question how many E_T^{miss} distributions are expected to have all ranges high. First, we calculate all of the E_T^{miss} distributions in this analysis. The three lepton channels are divided into OSSF0 and OSSF1 categories. Each OSSF1 three lepton is divided into below-Z, on-Z, and above-Z categories. This gives four groups that are divided into two H_T ranges, two b-jet categories, and two τ_h -lepton categories. For the three leptons channels there are 32 E_T^{miss} distributions. The four lepton channels are divided into OSSF0, OSSF1, and OSSF2 categories. The OSSF1 and OSSF2 categories are divided into off-Z and on-Z categories. The OSSF0 and OSSF1 categories are divided into two τ_h -lepton categories. This gives eight groups that are divided into two H_T ranges and two b-jet categories. The four leptons channels give an additional 32 E_T^{miss} distributions. There is a total of 64 E_T^{miss} distributions between the 3 and 4 lepton channels.

For the $3(e/\mu)+1\tau_h$, off-Z, low H_T , E_T^{miss} distribution we have 22 events with 10 ± 2.4 expected. Using the method of Z_{bi} we get a p-value of 1.04% (2.3σ). Using equation C.1 we can calculate the probability of getting n E_T^{miss} distributions out of 64 with a p-value $\leq 1.04\%$. Table C.6 gives the probabilities of getting n discrepant E_T^{miss} distributions. We find that we had a 49% chance of getting at least one E_T^{miss} distribution like the $3(e/\mu)+1\tau_h$, off-Z, low H_T , E_T^{miss} distribution. Since this is the only E_T^{miss} distribution where the total number of events is this far off from the expected number of events, we are within the expectations for statistical fluctuations.

$$p(n) = \binom{64}{n} \times (1.04\%)^n \times (100\% - 1.04\%)^{64-n} \quad (\text{C.1})$$

Table C.6: The probability of having n out of 64 E_T^{miss} distributions with p-values $\leq 1.04\%$ for various values of n .

n	Probability
0	51.2%
1	34.4%
2	11.4%
3	2.5%

C.4 $3\ell + 1\tau_h$ channel comparison plots

There is an excess of events in the four lepton channel where one of the four leptons is a τ_h -lepton, the number of OSSF lepton pairs is one, and no Z candidates. We examine distributions for the events with a Z candidate, which are in good agreement with data, and the no Z candidate events, which are in poor agreement with data. If the events with no Z candidate are a statistical fluctuation then one would expect that if the background were scaled to equal area with the data the shape of the distributions would match.

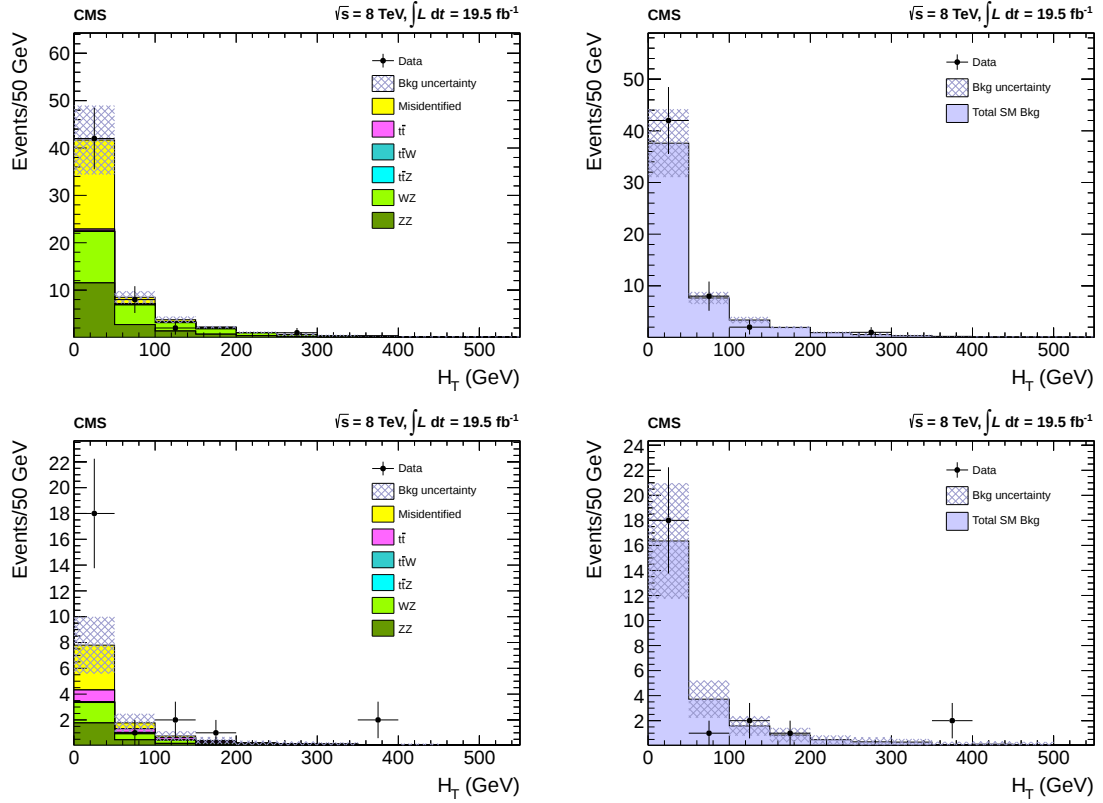


Figure C.1: Investigation of excess in four leptons, one of which is a τ_h -lepton. The top row has a Z candidate and the bottom row does not. On the left is the H_T distribution with expected backgrounds. The right plot shows the H_T distribution where the background has been normalized to match the observation.

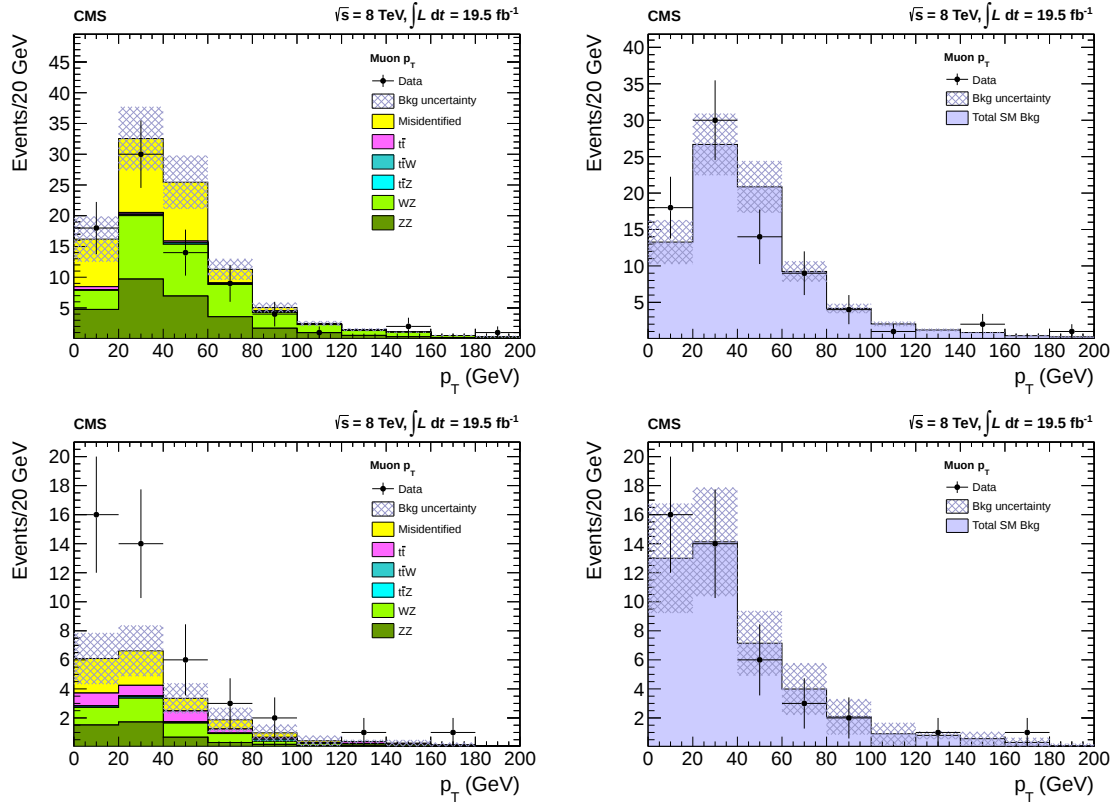


Figure C.2: Investigation of excess in four leptons, one of which is a τ_h -lepton. The top row has a Z candidate and the bottom row does not. The left is the muon p_T distribution with expected backgrounds. The right plot shows the muon p_T distribution where the background has been normalized to match the observation.

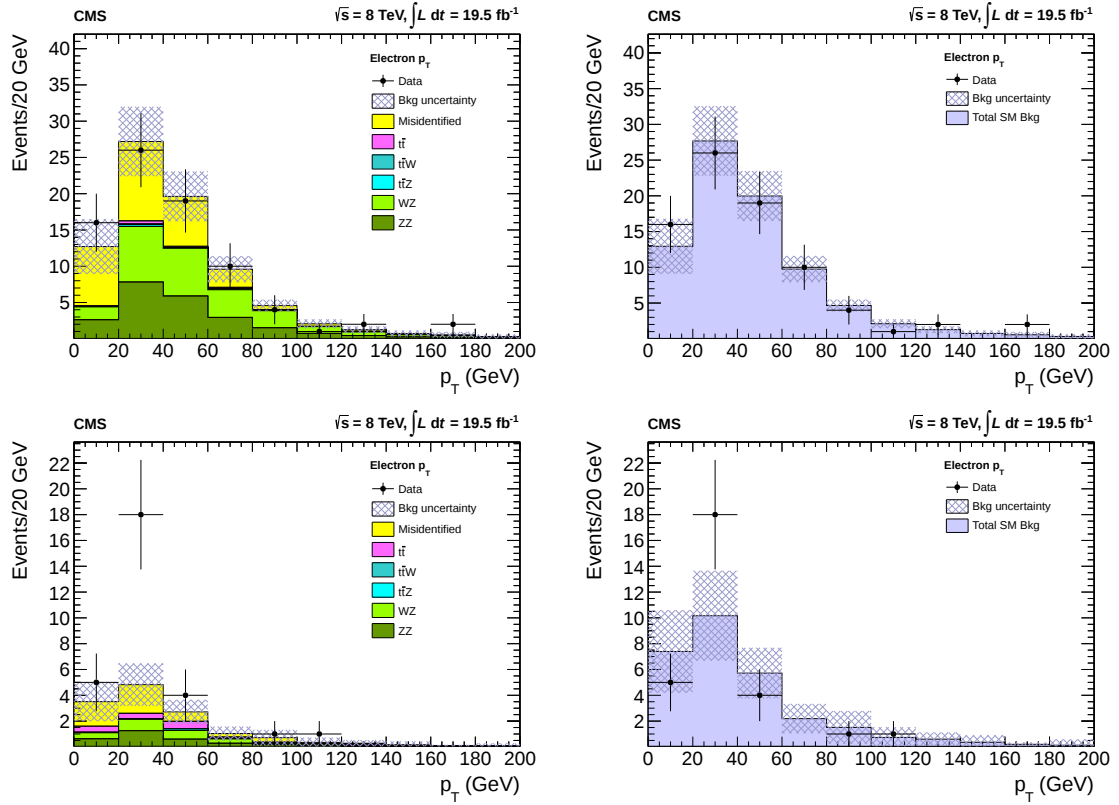


Figure C.3: Investigation of excess in four leptons, one of which is a τ_h -lepton. The top row has a Z candidate and the bottom row does not. On the left is the electron p_T distribution with expected backgrounds. The right plot shows the electron p_T distribution where the background has been normalized to match the observation.

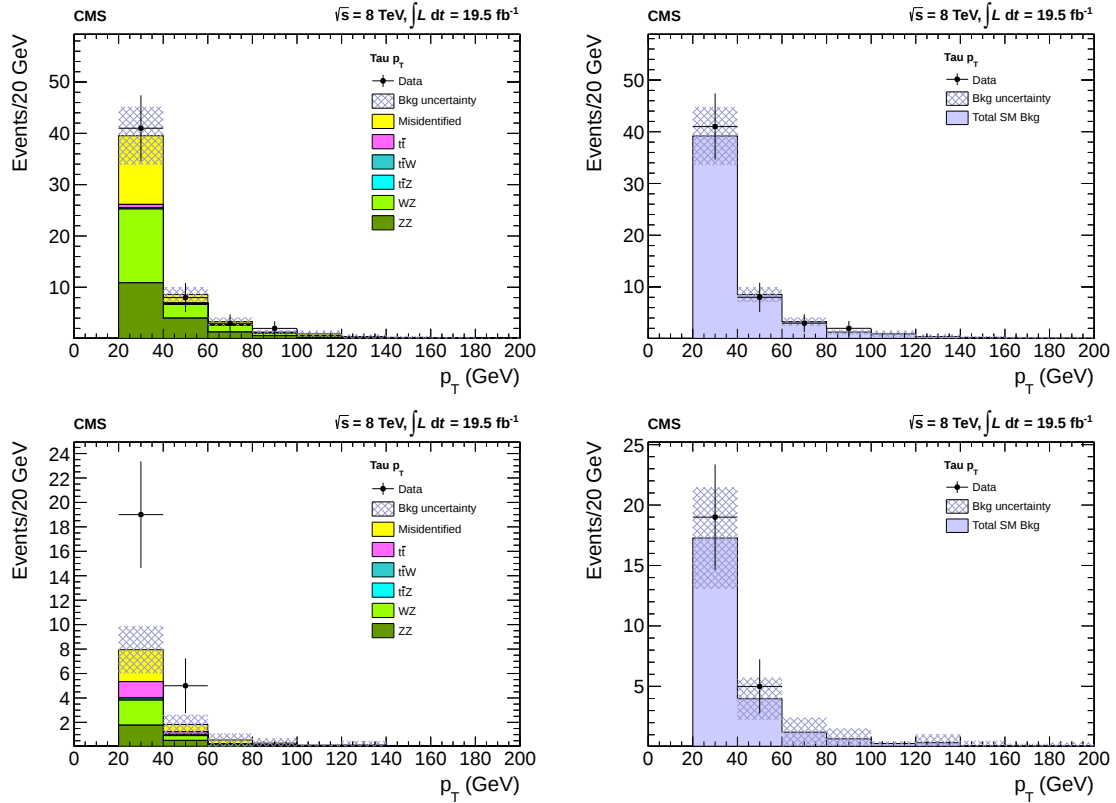


Figure C.4: Investigation of excess in four leptons, one of which is a τ_h -lepton. The top row has a Z candidate and the bottom row does not. On the left is the τ_h -lepton p_T distribution with expected backgrounds. The right plot shows the τ_h -lepton p_T distribution where the background has been normalized to match the observation.

Appendix D

Additional wino NLSP with a Higgs boson results

This section presents additional result tables for the wino NLSP with a Higgs boson scenario not found in Section 10.2.2.

Table D.1: Observed (Obs.) number of events and expected (Exp.) number of SM background events, along with the number of signal (Sig.) events, in the 5 best search channels for the model point $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = 150$ GeV, and $m_{\tilde{\chi}_1^0} = 1$ GeV. All channels shown have exactly three selected leptons, a veto on tagged b-jets, and $H_T < 200$ GeV. The results are categorized by the presence of an OSSF pair with invariant mass below 75 GeV (above 105 GeV) given by the “below-Z” (“above-Z”) designation, the E_T^{miss} in the event, and the number of τ_h candidates.

$m_{\ell+\ell-}$	E_T^{miss} (GeV)	N_{τ_h}	Obs.	Exp.	Sig.
Below-Z	(50, 100)	0	142	125 ± 28	14.9 ± 2.8
Below-Z	(100, 150)	0	16	21.3 ± 8.0	5.06 ± 0.86
—	(0, 50)	0	53	52 ± 12	4.61 ± 0.99
—	(50, 100)	0	35	38 ± 15	6.5 ± 1.1
—	(100, 150)	0	7	9.3 ± 4.3	2.32 ± 0.43

Table D.2: Results for the $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = 200$ GeV, $m_{\tilde{\chi}_1^0} = 1$ GeV model point. Details are the same as in Table D.1.

$m_{\ell+\ell-}$	E_T^{miss} (GeV)	N_{τ_h}	Obs.	Exp.	Sig.
Below-Z	(50, 100)	0	142	125 ± 28	4.90 ± 0.91
Below-Z	(100, 150)	0	16	21.3 ± 8.0	2.63 ± 0.43
Below-Z	(150, 200)	0	5	2.9 ± 1.0	0.61 ± 0.16
—	(50, 100)	0	35	38 ± 15	2.31 ± 0.43
—	(100, 150)	0	7	9.3 ± 4.3	1.31 ± 0.26

Table D.3: Results for the $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = 300 \text{ GeV}$, $m_{\tilde{\chi}_1^0} = 1 \text{ GeV}$ model point. Details are the same as in Table D.1.

$m_{\ell^+\ell^-}$	$E_{\text{T}}^{\text{miss}}$ (GeV)	$N_{\tau_{\text{h}}}$	Obs.	Exp.	Sig.
Below-Z	(100, 150)	0	16	21.3 ± 8.0	0.70 ± 0.13
Below-Z	(150, 200)	0	5	2.9 ± 1.0	0.348 ± 0.067
Below-Z	(200, ∞)	0	0	0.88 ± 0.31	0.218 ± 0.041
Above-Z	(150, 200)	0	1	2.48 ± 0.68	0.180 ± 0.045
—	(150, 200)	1	8	15.1 ± 7.4	0.44 ± 0.12

Table D.4: Results for the $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = 400 \text{ GeV}$, $m_{\tilde{\chi}_1^0} = 1 \text{ GeV}$ model point. Details are the same as in Table D.1.

$m_{\ell^+\ell^-}$	$E_{\text{T}}^{\text{miss}}$ (GeV)	$N_{\tau_{\text{h}}}$	Obs.	Exp.	Sig.
Below-Z	(100, 150)	0	16	21.3 ± 8.0	0.167 ± 0.028
Below-Z	(150, 200)	0	5	2.9 ± 1.0	0.138 ± 0.025
Below-Z	(200, ∞)	0	0	0.88 ± 0.31	0.137 ± 0.025
—	(200, ∞)	0	0	0.42 ± 0.22	0.057 ± 0.011
—	(200, ∞)	1	3	2.4 ± 1.1	0.152 ± 0.038

Bibliography

- [1] J. C. Street and E. C. Stevenson, “New Evidence for the Existence of a Particle of Mass Intermediate Between the Proton and Electron,” *Phys. Rev.*, vol. 52, p. 1003, Nov 1937.
- [2] C. Cowan, F. Reines, F. Harrison, H. Kruse, and A. McGuire, “Detection of the free neutrino: A Confirmation,” *Science*, vol. 124, p. 103, 1956.
- [3] G. Danby, J.-M. Gaillard, K. Goulianos, L. M. Lederman, N. Mistry, M. Schwartz, and J. Steinberger, “Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos,” *Phys. Rev. Lett.*, vol. 9, p. 36, 1962.
- [4] M. Gell-Mann, “A schematic model of baryons and mesons,” *Phys. Lett.*, vol. 8, no. 3, p. 214, 1964.
- [5] G. Zweig, “An SU(3) model for strong interaction symmetry and its breaking,” 1964.
- [6] C.-N. Yang and R. L. Mills, “Conservation of Isotopic Spin and Isotopic Gauge Invariance,” *Phys. Rev.*, vol. 96, p. 191, 1954.
- [7] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons,” *Phys. Rev. Lett.*, vol. 13, p. 508, 1964.
- [8] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons,” *Phys. Rev. Lett.*, vol. 13, p. 321, 1964.
- [9] S. Weinberg, “A Model of Leptons,” *Phys. Rev. Lett.*, vol. 19, p. 1264, 1967.
- [10] S. Glashow, “Partial Symmetries of Weak Interactions,” *Nucl. Phys.*, vol. 22, p. 579, 1961.
- [11] G. Aad *et al.*, “Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments,” 2015.
- [12] S. Chatrchyan *et al.*, “Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV,” *JHEP*, vol. 1306, p. 081, 2013.
- [13] S. Chatrchyan *et al.*, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Phys. Lett. B*, vol. 716, p. 30, 2012.
- [14] G. Aad *et al.*, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,” *Phys. Lett. B*, vol. 716, p. 1, 2012.
- [15] Wikipedia, “Standard model of particle physics.” http://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

- [16] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*. Advanced book classics, Addison-Wesley Publishing Company, 1995.
- [17] J. Beringer *et al.*, “Review of Particle Physics,” *Phys. Rev. D*, vol. 86, p. 010001, 2012.
- [18] Goldstone, Jeffrey and Salam, Abdus and Weinberg, Steven, “Broken symmetries,” *Phys. Rev.*, vol. 127, p. 965, 1962.
- [19] Wikipedia, “Dark matter.” http://en.wikipedia.org/wiki/Dark_matter.
- [20] P. Athron, *Aspects of Electroweak Symmetry Breaking in Physics Beyond the Standard Model*. PhD thesis, University of Glasgow, November 2008.
- [21] H. P. Nilles, “Supersymmetry, Supergravity and Particle Physics,” *Phys. Rept.*, vol. 110, p. 1, 1984.
- [22] H. E. Haber and G. L. Kane, “The Search for Supersymmetry: Probing Physics Beyond the Standard Model,” *Phys. Rept.*, vol. 117, p. 75, 1985.
- [23] W. de Boer, “Grand unified theories and supersymmetry in particle physics and cosmology,” *Prog. Part. Nucl. Phys.*, vol. 33, p. 201, 1994.
- [24] S. P. Martin, “A Supersymmetry primer,” *Adv. Ser. Direct. High Energy Phys.*, vol. 21, p. 1, 2010.
- [25] F. Quevedo, S. Krippendorff, and O. Schlotterer, “Cambridge Lectures on Supersymmetry and Extra Dimensions,” 2010.
- [26] Wikipedia, “Minimal Supersymmetric Standard Model.” http://en.wikipedia.org/wiki/Minimal_Supersymmetric_Standard_Model.
- [27] D. Alves *et al.*, “Simplified Models for LHC New Physics Searches,” *J. Phys. G*, vol. 39, p. 105005, 2012.
- [28] J. Alwall, P. Schuster, and N. Toro, “Simplified Models for a First Characterization of New Physics at the LHC,” *Phys. Rev. D*, vol. 79, p. 075020, 2009.
- [29] S. Kraml, S. Kulkarni, U. Laa, A. Lessa, W. Magerl, *et al.*, “SModelS: a tool for interpreting simplified-model results from the LHC and its application to supersymmetry,” *Eur. Phys. J. C*, vol. 74, p. 2868, 2014.
- [30] L. Evans and P. Bryant, “LHC Machine,” *JINST*, vol. 3, p. S08001, 2008.
- [31] F. Marcastel, “CERN’s Accelerator Complex. La chaîne des accélérateurs du CERN,” Oct 2013. General Photo.
- [32] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. Geneva: CERN, 2004.
- [33] S. Chatrchyan *et al.*, “Commissioning of the CMS Experiment and the Cosmic Run at Four Tesla,” *JINST*, vol. 5, p. T03001, 2010.
- [34] S. Chatrchyan *et al.*, “The CMS experiment at the CERN LHC,” *JINST*, vol. 3, p. S08004, 2008.

- [35] G. L. Bayatian, S. Chatrchyan, G. Hmayakyan, *et al.*, *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS, Geneva: CERN, 2006. There is an error on cover due to a technical problem for some items.
- [36] V. Khachatryan *et al.*, “CMS tracking performance results from early LHC operation,” *Eur. Phys. J. C*, vol. 70, p. 1165, 2010.
- [37] S. Chatrchyan *et al.*, “Description and performance of track and primary-vertex reconstruction with the CMS tracker,” *JINST*, vol. 9, no. 10, p. P10009, 2014.
- [38] V. Karimäki, M. Mannelli, P. Siegrist, H. Breuker, A. Caner, R. Castaldi, K. Freudenreich, G. Hall, R. Horisberger, M. Huhtinen, and A. Cattai, *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS, Geneva: CERN, 1997.
- [39] S. Chatrchyan *et al.*, “Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV,” *JINST*, vol. 8, p. P09009, 2013.
- [40] S. Chatrchyan *et al.*, “Performance of the CMS Hadron Calorimeter with Cosmic Ray Muons and LHC Beam Data,” *JINST*, vol. 5, p. 03012, 2010.
- [41] S. Abdullin *et al.*, “The CMS barrel calorimeter response to particle beams from 2 to 350 GeV/c,” *Eur. Phys. J. C*, vol. 60, p. 359, 2009.
- [42] G. Baiatian *et al.*, “Design, Performance, and Calibration of CMS Hadron Endcap Calorimeters,” vol. CMS-NOTE-2008-010, Mar 2008.
- [43] S. Abdullin *et al.*, “Design, performance, and calibration of the CMS Hadron-outer calorimeter,” *Eur. Phys. J. C*, vol. 57, p. 653, 2008.
- [44] G. Bayatian, A. Sirunian, I. Emelyanchik, V. Massolov, N. Shumeiko, *et al.*, “Design, performance, and calibration of CMS forward calorimeter wedges,” *Eur. Phys. J. C*, vol. 53, no. 1, p. 139, 2008.
- [45] J. Brooke, “Performance of the CMS Level-1 Trigger,” *PoS*, vol. ICHEP2012, p. 508, 2013.
- [46] T. Lenzi, “Development and Study of Different Muon Track Reconstruction Algorithms for the Level-1 Trigger for the CMS Muon Upgrade with GEM Detectors,” 2013.
- [47] CMS Collaboration, “CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update,” *CMS-PAS-LUM-13-001*, 2013.
- [48] CMS Collaboration, “Public CMS Luminosity Information.” <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [49] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, “MADGRAPH 5 : Going Beyond,” *JHEP*, vol. 06, p. 128, 2011.
- [50] F. Maltoni and T. Stelzer, “MADGRAPH: Automatic event generation with MADGRAPH,” *JHEP*, vol. 02, p. 027, 2003.

- [51] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method,” *JHEP*, vol. 11, p. 070, 2007.
- [52] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual,” *JHEP*, vol. 05, p. 026, 2006.
- [53] T. Sjöstrand, S. Mrenna, and P. Z. Skands, “A brief introduction to PYTHIA 8.1,” *Comput. Phys. Commun.*, vol. 178, p. 852, 2008.
- [54] S. Kretzer, H. L. Lai, F. I. Olness, and W. K. Tung, “CTEQ6 parton distributions with heavy quark mass effects,” *Phys. Rev. D*, vol. 69, p. 114005, 2004.
- [55] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, *et al.*, “Implications of CTEQ global analysis for collider observables,” *Phys. Rev. D*, vol. 78, p. 013004, 2008.
- [56] S. Agostinelli *et al.*, “GEANT4—a simulation toolkit,” *Nucl. Instrum. Meth. A*, vol. 506, p. 250, 2003.
- [57] S. Abdullin, P. Azzi, F. Beaudette, P. Janot, and A. Perrotta, “The fast simulation of the CMS detector at LHC,” *J. Phys. Conf. Ser.*, vol. 331, p. 032049, 2011.
- [58] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, “W Physics at the LHC with FEWZ 2.1,” *Comput. Phys. Commun.*, vol. 184, p. 208, 2013.
- [59] N. Kidonakis, “Differential and total cross sections for top pair and single top production,” pp. 831–834, 2012.
- [60] J. M. Campbell and R. K. Ellis, “ $t\bar{t}W^\pm$ production and decay at NLO,” *JHEP*, vol. 07, p. 052, 2012.
- [61] M. Garzelli, A. Kardos, C. 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*, vol. 11, p. 056, 2012.
- [62] J. M. Campbell and R. Ellis, “MCFM for the Tevatron and the LHC,” *Nucl. Phys. Proc. Suppl.*, vol. 205, p. 10, 2010.
- [63] S. Frixione and B. R. Webber, “Matching NLO QCD computations and parton shower simulations,” *JHEP*, vol. 06, p. 029, 2002.
- [64] S. Frixione, P. Nason, and B. R. Webber, “Matching NLO QCD and parton showers in heavy flavor production,” *JHEP*, vol. 08, p. 007, 2003.
- [65] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau, *et al.*, “Four-lepton production at hadron colliders: MC@NLO predictions with theoretical uncertainties,” *JHEP*, vol. 02, p. 099, 2012.
- [66] A. Djouadi, J.-L. Kneur, and G. Moultaka, “SUSPECT: A Fortran code for the supersymmetric and Higgs particle spectrum in the MSSM,” *Comput. Phys. Commun.*, vol. 176, p. 426, 2007.

- [67] P. Z. Skands, B. Allanach, H. Baer, C. Balazs, G. Belanger, *et al.*, “SUSY Les Houches accord: Interfacing SUSY spectrum calculators, decay packages, and event generators,” *JHEP*, vol. 07, p. 036, 2004.
- [68] J. Alwall, A. Ballestrero, P. Bartalini, S. Belov, E. Boos, *et al.*, “A Standard format for Les Houches event files,” *Comput. Phys. Commun.*, vol. 176, p. 300, 2007.
- [69] P. Meade and M. Reece, “BRIDGE: Branching ratio inquiry/decay generated events,” 2007.
- [70] K. T. Matchev and S. D. Thomas, “Higgs and Z boson signatures of supersymmetry,” *Phys. Rev. D*, vol. 62, p. 077702, 2000.
- [71] 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*, vol. 1210, p. 081, 2012.
- [72] 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*, vol. 73, p. 2480, 2013.
- [73] M. Kramer, A. Kulesza, R. van der Leeuw, M. Mangano, S. Padhi, *et al.*, “Supersymmetry production cross sections in pp collisions at $\sqrt{s} = 7$ TeV,” 2012.
- [74] W. Beenakker, M. Klasen, M. Krämer, T. Plehn, M. Spira, and P. M. Zerwas, “Production of Charginos, Neutralinos, and Stopped at Hadron Colliders,” *Phys. Rev. Lett.*, vol. 83, p. 3780, 1999.
- [75] W. Beenakker, M. Klasen, M. Kramer, T. Plehn, M. Spira, and P. M. Zerwas, “Erratum: Production of Charginos, Neutralinos, and Stopped at Hadron Colliders [Phys. Rev. Lett. 83, 3780 (1999)],” *Phys. Rev. Lett.*, vol. 100, p. 029901, 2008.
- [76] W. Beenakker, R. Höpker, and M. Spira, *PROSPINO: A program for the production of supersymmetric particles in next-to-leading order QCD*, 1996.
- [77] W. Beenakker, R. Höpker, M. Spira, and P. M. Zerwas, “Squark and gluino production at hadron colliders,” *Nucl. Phys. B*, vol. 492, p. 51, 1997.
- [78] A. Kulesza and L. Motyka, “Threshold resummation for squark-antisquark and gluino-pair production at the LHC,” *Phys. Rev. Lett.*, vol. 102, p. 111802, 2009.
- [79] A. Kulesza and L. Motyka, “Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC,” *Phys. Rev. D*, vol. 80, p. 095004, 2009.
- [80] W. Beenakker, S. Brensing, M. Krämer, A. Kulesza, E. Laenen, and I. Niessen, “Soft-gluon resummation for squark and gluino hadroproduction,” *JHEP*, vol. 12, p. 041, 2009.
- [81] W. Beenakker, S. Brensing, M. Krämer, A. Kulesza, E. Laenen, L. Motyka, and I. Niessen, “Squark and gluino hadroproduction,” *Int. J. Mod. Phys. A*, vol. 26, p. 2637, 2011.

- [82] F. Beaudette, “The CMS Particle Flow Algorithm,” p. 295, 2014.
- [83] CMS Collaboration, “Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector,” *CMS Physics Analysis Summary*, vol. CMS-PAS-PFT-10-001, 2010.
- [84] CMS Collaboration, “Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV,” *CMS Physics Analysis Summary*, vol. CMS-PAS-PFT-10-002, 2010.
- [85] R. Frühwirth, “Application of Kalman filtering to track and vertex fitting,” *Nucl. Instrum. Meth. A*, vol. 262, p. 444, 1987.
- [86] W. Erdmann, “Vertex reconstruction at the cms experiment,” *J. Phys. Conf. Ser.*, vol. 110, no. 9, p. 092009, 2008.
- [87] P. W. Erdmann, “Offline Primary Vertex Reconstruction with Deterministic Annealing Clustering,” *CMS Internal Note*, vol. CMS IN-2011/014, 2011.
- [88] W. W. R. Frühwirth, “Adaptive Vertex Fitting,” *CMS Note*, vol. CMS NOTE-2007/008, 2007.
- [89] CMS Collaboration, “Photon reconstruction and identification at $\sqrt{s} = 7$ TeV,” *CMS Physics Analysis Summary*, vol. CMS-PAS-EGM-10-005, 2010.
- [90] CMS Collaboration, “Isolated Photon Reconstruction and Identification at $\sqrt{s} = 7$ TeV,” *CMS Physics Analysis Summary*, 2011.
- [91] S. Chatrchyan *et al.*, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV,” *JINST*, vol. 7, p. P10002, 2012.
- [92] CMS Collaboration, “Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV,” *CMS Physics Analysis Summary*, vol. CMS-PAS-MUO-10-002, 2010.
- [93] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, *et al.*, “Electron reconstruction in CMS,” *Eur. Phys. J. C*, vol. 49, p. 1099, 2007.
- [94] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET,” *CMS Physics Analysis Summary*, vol. CMS-PAS-PFT-09-001, 2009.
- [95] CMS Collaboration, “Electron Reconstruction and Identification at $\sqrt{s} = 7$ TeV,” *CMS Physics Analysis Summary*, vol. CMS-PAS-EGM-10-004, 2010.
- [96] CMS Collaboration, “Performance of τ -lepton reconstruction and identification in CMS,” *JINST*, vol. 7, p. P01001, 2012.
- [97] A. K. Nayak, “Reconstruction of physics objects in the CMS detector,” *PoS*, vol. CHARGED2012, p. 010, 2012.
- [98] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet User Manual,” *Eur. Phys. J. C*, vol. 72, p. 1896, 2012.

- [99] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_T jet clustering algorithm,” *JHEP*, vol. 04, p. 063, 2008.
- [100] Y. L. Dokshitzer, G. Leder, S. Moretti, and B. Webber, “Better jet clustering algorithms,” *JHEP*, vol. 9708, p. 001, 1997.
- [101] S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions,” *Phys. Rev. D*, vol. 48, p. 3160, 1993.
- [102] CMS Collaboration, “Determination of jet energy calibration and transverse momentum resolution in CMS,” *JINST*, vol. 6, p. P11002, 2011.
- [103] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas,” *Phys. Lett. B*, vol. 659, p. 119, 2008.
- [104] S. Chatrchyan *et al.*, “Identification of b-quark jets with the CMS experiment,” *JINST*, vol. 8, p. P04013, 2013.
- [105] W. Waltenberger, “Adaptive vertex reconstruction,” *CERN-CMS-NOTE-2008-033*, 2008.
- [106] K. A. Olive *et al.*, “Review of Particle Physics,” *Chin. Phys. C*, vol. 38, no. 9, p. 090001, 2014.
- [107] V. Khachatryan *et al.*, “Performance of the missing transverse energy reconstruction by the CMS experiment in $\sqrt{s} = 8$ TeVpp data,” 2014.
- [108] CMS Collaboration, “Missing transverse energy performance of the cms detector,” *JINST*, vol. 6, no. 09, p. P09001, 2011.
- [109] CMS Collaboration, “Missing Transverse Energy Performance in Minimum-Bias and Jet Events from Proton-Proton Collisions at $\sqrt{s} = 7$ TeV,” 2010.
- [110] CMS Collaboration, “CMS E_T^{miss} Performance in Events Containing Electroweak Bosons from pp Collisions at $\sqrt{s} = 7$ TeV,” 2010.
- [111] S. Arora, J. Chou, C. Contreras-Campana, E. Contreras-Campana, R. Gray, A. Lath, S. Panwalkar, S. Schnetzer, S. Somalwar, S. Thomas, P. Thomassen, M. Walker, and P. Zywicki, “Background and Efficiency Determination Methods for Multilepton Analyses,” *CMS Analysis Note*, vol. CMS AN-2012/257, 2012.
- [112] S. Arora, J. Chou, C. Contreras-Campana, E. Contreras-Campana, R. Gray, A. Lath, S. Panwalkar, S. Schnetzer, S. Somalwar, S. Thomas, P. Thomassen, M. Walker, and P. Zywicki, “A Search for Direct Chargino Neutralino production with three or more leptons using $\sqrt{s} = 8$ TeV CMS data,” *CMS Analysis Note*, vol. CMS AN-2012/256, 2012.
- [113] M. Chen, D. Dobur, A. Korytov, K. Matchev, G. Mitselmakher, L. Muniz, R. Remington, L. Shchutska, N. Skhirtladze, J. Yelton, P. Everaerts, R. Cousins, and J. Hauser, “Search for Direct Electroweak Production of Charginos and Neutralinos with the Triplepton Plus Missing Energy Final State,” *CMS Analysis Note*, vol. CMS AN-2012/248, 2012.

- [114] D. Barge, C. Campagnari, P. Kalavase, D. Kovalskyi, V. Krutelyov, J. Ribnik, W. Andrews, G. Cerati, D. Evans, F. Golf, I. MacNeill, S. Padhi, Y. Tu, F. Wurthwein, A. Yagil, J. Yoo, L. Bauerdick, I. Bloch, K. Burkett, I. Fisk, Y. Gao, O. Gutsche, B. Hooberman, S. Jindariani, and J. Linacre, “Search for New Physics with Same-Sign Dileptons using the 2011 dataset of CMS,” *CMS Analysis Note*, vol. CMS AN-2011/258, 2011.
- [115] CMS Internal twiki, “Apply SF for b/c/light tags to directly tagged jets in MC.” <https://twiki.cern.ch/twiki/bin/view/CMS/BTagSFUtil>.
- [116] S. Arora, J. Chou, C. Contreras-Campana, E. Contreras-Campana, R. Gray, A. Lath, S. Panwalkar, S. Schnetzer, S. Somalwar, S. Thomas, P. Thomassen, M. Walker, and P. Zywicki, “A search for anomalous production of events with three or more leptons using 19.5 fb^{-1} of $\sqrt{s} = 8\text{ TeV}$ LHC data,” *CMS Analysis Note*, vol. CMS AN-2012/343, 2012.
- [117] S. Chatrchyan *et al.*, “Search for anomalous production of multilepton events in pp collisions at $\sqrt{s} = 7\text{ TeV}$,” *JHEP*, vol. 06, p. 169, 2012.
- [118] S. Chatrchyan *et al.*, “Search for electroweak production of charginos and neutralinos using leptonic final states in pp collisions at $\sqrt{s} = 7\text{ TeV}$,” *JHEP*, vol. 1211, p. 147, 2012.
- [119] S. Chatrchyan *et al.*, “Search for Physics Beyond the Standard Model Using Multilepton Signatures in pp Collisions at $\sqrt{s} = 7\text{ TeV}$,” *Phys. Lett. B*, vol. 704, p. 411, 2011.
- [120] R. C. Gray, C. Kilic, M. Park, S. Somalwar, and S. Thomas, “Backgrounds to Higgs Boson searches from $W\gamma^* \rightarrow \ell\nu\ell(\ell)$ asymmetric internal conversion,” 2011.
- [121] S. Chatrchyan *et al.*, “Search for top-squark pair production in the single-lepton final state in pp collisions at $\sqrt{s} = 8\text{ TeV}$,” *Eur. Phys. J. C*, vol. 73, p. 2677, 2013.
- [122] S. Arora, J. Chou, C. Contreras-Campana, E. Contreras-Campana, R. Gray, A. Lath, S. Panwalkar, S. Schnetzer, S. Somalwar, S. Thomas, P. Thomassen, M. Walker, and P. Zywicki, “ E_T^{miss} resolution dependence on pileup,” *CMS Analysis Note*, vol. CMS AN-2012/342, 2012.
- [123] J. Alwall, S. de Visscher, and F. Maltoni, “QCD radiation in the production of heavy colored particles at the LHC,” *JHEP*, vol. 0902, p. 017, 2009.
- [124] ATLAS and CMS Collaboration, “Procedure for the LHC Higgs boson search combination in summer 2011,” tech. rep., CERN, 2011.
- [125] B. Mistlberger and F. Dulat, “Limit setting procedures and theoretical uncertainties in Higgs boson searches,” 2012.
- [126] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” *Eur. Phys. J. C*, vol. 71, p. 1554, 2011.
- [127] A. L. Read, “Modified frequentist analysis of search results (the CL_s method) in “Workshop on Confidence Limits”, Eds. F. James, L. Lyons, and Y. Perrin,” 2000. p. 81.

- [128] A. L. Read, “Presentation of search results: The CL(s) technique,” *J. Phys. G*, vol. 28, p. 2693, 2002.
- [129] Bob Cousins, et al., “Probability density functions for positive nuisance parameters,” 2010.
- [130] Mingshui Chen, et al., “Lands software package,” 2013.
- [131] S. Chatrchyan *et al.*, “Search for anomalous production of events with three or more leptons in pp collisions at $\sqrt{s} = 8$ TeV,” *Phys. Rev. D*, vol. 90, p. 032006, 2014.
- [132] S. Chatrchyan *et al.*, “Search for anomalous production of events with three or more leptons in pp collisions at $\sqrt{s} = 8$ TeV,” *Phys. Rev. D*, vol. 90, p. 032006, 2014.
- [133] CMS Collaboration, “A search for anomalous production of events with three or more leptons using 19.5 fb^{-1} of $\sqrt{s} = 8$ TeV LHC data,” *CMS Physics Analysis Summary*, vol. CMS-PAS-SUS-13-002, 2013.
- [134] S. Chatrchyan *et al.*, “Search for top squark and higgsino production using diphoton Higgs boson decays,” *Phys. Rev. Lett.*, vol. 112, p. 161802, 2014.
- [135] CMS Collaboration, “Search for electroweak production of charginos, neutralinos, and sleptons using leptonic final states in pp collisions at 8 TeV,” *CMS Physics Analysis Summary*, vol. CMS-PAS-SUS-13-006, 2013.
- [136] V. Khachatryan *et al.*, “Searches for electroweak production of charginos, neutralinos, and sleptons decaying to leptons and W, Z, and Higgs bosons in pp collisions at 8 TeV,” 2014.
- [137] CMS Collaboration, “Search for electroweak production of charginos and neutralinos in final states with a Higgs boson in pp collisions at 8 TeV,” *CMS Physics Analysis Summary*, vol. CMS-PAS-SUS-13-017, 2013.
- [138] S. Dimopoulos, S. D. Thomas, and J. D. Wells, “Implications of low energy supersymmetry breaking at the Fermilab Tevatron,” *Phys. Rev. D*, vol. 54, p. 3283, 1996.
- [139] R. L. Culbertson *et al.*, “Low scale and gauge mediated supersymmetry breaking at the Fermilab Tevatron Run II,” 2000.
- [140] J. T. Ruderman and D. Shih, “Slepton co-NLSPs at the Tevatron,” *JHEP*, vol. 11, p. 046, 2010.
- [141] R. Essig, E. Izaguirre, J. Kaplan, and J. G. Wacker, “Heavy Flavor Simplified Models at the LHC,” *JHEP*, vol. 01, p. 074, 2012.
- [142] S. Chatrchyan *et al.*, “Interpretation of searches for supersymmetry with simplified models,” *Phys. Rev. D*, vol. 88, p. 052017, 2013.
- [143] H. M. Lee, V. Sanz, and M. Trott, “Hitting sbottom in natural SUSY,” *JHEP*, vol. 05, p. 139, 2012.