Facultad de Ciencias

REINTERPRETACIÓN DE LAS BÚSQUEDAS DE CMS EN COLISIONES DE PROTONES CON ENERGÍA DE CENTRO DE MASA DE 13 TEV PARA EL MSSM FENOMENOLÓGICO

(Reinterpretation of CMS searches in proton-proton collisions at a center of mass energy of 13 TeV for the phenomenological MSSM)

Trabajo de Fin de Máster
para acceder al

MÁSTER EN FÍSICA DE PARTÍCULAS Y DEL COSMOS

Autor: Balogun Olufunke Abigail

Director: Luca Scodellaro

Septiembre - 2025

Chapter 1

Acknowledgements

I would like to express my deepest appreciation to Dr. Luca Scodellaro, my thesis advisor, for his tireless guidance, support, and motivation throughout this study. Dr. Scodellaro's leadership and adherence to excellence have been central to both the development of this thesis and my scholarly development.
I am very grateful to the particle physics group of IFCA (Instituto de Física de Cantabria), especially Dr. Alicia Calderón Tazón and Dr. Rocío Vilar Cortabitarte, for their scientific advice, positive feedback, and openness to collaboration. The expertise and generosity of the IFCA team enriched and made the research process enjoyable.
I would also like to thank Dr. Francisco Jesús Carrera, my graduate coordinator, and Dr. Nuria Castello Mor for the academic advising and the mentorship they provided me during my master's studies at the University of Cantabria.
My sincere thanks go to the Women for Africa Foundation for the scholarship that enabled me to pursue my master's studies in Spain. The commitment of the Foundation to empower women scientists in Africa is highly valued and deeply inspiring.
To my loved ones, thank you for your unwavering love, prayers, and support. Your confidence in me has been my greatest asset.
Finally, I express my heartfelt appreciation to all colleagues and friends who assisted with this work directly or indirectly. Your motivation has made this journey valuable and worthwhile.

"Nothing in life is to be feared; it is only to be understood. Now is the time to understand more so that we may fear less."
-Marie Curie

Abstract

Despite its remarkable success, the Standard Model (SM) leaves several fundamental questions unanswered, such as the nature of dark matter and the hierarchy problem. Supersymmetry (SUSY) is an extremely well-motivated extension of the SM that addresses these shortcomings by postulating a symmetry between bosons and fermions, predicting superpartners for all SM particles.

This thesis presents the reinterpretation of results from a previous search for SUSY, conducted at IFCA, within the framework of the 19-parameter phenomenological Minimal Supersymmetric Standard Model (phenomenological MSSM or pMSSM). The search was originally designed in the context of simplified SUSY models, targeting chargino and slepton pair production via electroweak interactions, as well as top-squark pair production in compressed spectra. The analysis used proton-proton collision data collected with the CMS detector at the CERN LHC between 2016 and 2018, corresponding to an integrated luminosity of 138 fb^{-1}. The search of interest consists of two isolated, oppositely charged leptons (electrons or muons) and large missing transverse momentum, a distinct signature of the presence of weakly interacting stable particles, such as the lightest neutralino.

The stransverse mass (m_{T2}) parameter was employed as the primary discriminant, and a maximum-likelihood combined fit across multiple signal regions, in terms of missing transverse momentum, jet multiplicity, and b-tagged jets, to constrain the Standard Model backgrounds and test for SUSY signals. No significant deviation was observed. Exclusion limits at 95% confidence level were set, ruling out chargino masses up to **1100 GeV** and slepton masses up to **700 GeV** for a massless neutralino.

This work describes the software framework required to reinterpret these results in the pMSSM, providing crucial input for constraining its parameter space.

Resumen

A pesar de su notable éxito, el Modelo Estándar (ME) deja varias preguntas fundamentales sin respuesta, como la naturaleza de la materia oscura y el problema de la jerarquía. La Supersimetría (SUSY) es una extensión del ME extremadamente bien motivada que aborda estas deficiencias al postular una simetría entre bosones y fermiones, prediciendo supercompañeros para todas las partículas del ME.
Esta tesis presenta la reinterpretación de los resultados de una búsqueda previa de SUSY, realizada en el IFCA, dentro del marco del Modelo Supersimétrico Mínimo Fenomenológico de 19 parámetros (MSSM fenomenológico o pMSSM). La búsqueda fue diseñada originalmente en el contexto de modelos SUSY simplificados, dirigida a la producción de pares de charginos y sleptones mediante interacciones electrodébiles, así como a la producción de pares de top-squarks en espectros comprimidos. El análisis utilizó datos de colisiones protón-protón recopilados con el detector CMS del LHC del CERN entre 2016 y 2018, correspondientes a una luminosidad integrada de 138 fb^{-1}.
La búsqueda de interés consiste en dos leptones (electrones o muones) aislados, de carga opuesta y gran momento transversal faltante, una firma distintiva de la presencia de partículas estables débilmente interactivas, como el neutralino más ligero.
El parámetro de masa transverso (mT2) se empleó como el discriminante principal, y un ajuste combinado de máxima verosimilitud a través de múltiples regiones de señal, en términos de momento transversal faltante, multiplicidad de jets y jets etiquetados como b, para restringir los fondos del Modelo Estándar y probar las señales de SUSY. No se observó ninguna desviación significativa. Se establecieron límites de exclusión con un nivel de confianza del 95%, que descartan masas de chargino de hasta 1100 GeV y masas de sleptón de hasta 700 GeV para un neutralino sin masa.
Este trabajo describe el marco de software necesario para reinterpretar estos resultados en el pMSSM, proporcionando una contribución crucial para restringir su espacio de parámetros.

Contents

List of Figures

List of Tables

Chapter 2

Introduction

2.1 Motivation

The Standard Model (SM) of particle physics has been remarkably successful in describing known fundamental particles and their interactions, providing precise predictions that have been thoroughly tested experimentally. Nevertheless, the SM has crucial shortcomings: It does not explain the nature of dark matter and does not address the hierarchy problem, i.e., why the electroweak scale is so much lower than the Planck scale. These limitations point to the need for new physics beyond the SM.

2.2 Supersymmetry and Its Relevance

Supersymmetry (SUSY) is one of the most extensively studied extensions of the SM. It postulates a symmetry between fermions and bosons, introducing a superpartner for every known particle that differs by half a unit of spin. SUSY offers a solution to the hierarchy problem by stabilizing the Higgs boson mass against large quantum corrections. Furthermore, if a discrete symmetry called R-parity is conserved, the lightest supersymmetric particle (LSP) is stable and can serve as a viable dark-matter candidate.

2.3 Objectives and Structure of the Thesis

This thesis focuses on a search for SUSY particle pair production, specifically charginos and sleptons, leading to final states with two oppositely charged leptons (electrons or muons) and large missing transverse momentum ($p_{\mathrm{T}}^{\mathrm{miss}}$). Such signatures are indicative of neutral, weakly interacting particles that

escape detection, such as the LSP.
Supersymmetric particles have been the subject of intensive experimental searches at the Large Hadron Collider (LHC). In particular, the CMS Collaboration has explored a broad range of final states and signatures. For each analysis, results are interpreted in terms of simplified models, where only a few parameters define idealized scenarios. However, simplified models do not capture the full phenomenology of SUSY, as they neglect the interplay of multiple production and decay channels and provide no insight into dark matter. To address this, the CMS Collaboration has initiated efforts to reinterpret results in the more general phenomenological MSSM (pMSSM), a 19-parameter realization of the minimal supersymmetric standard model.
One of the participating SUSY searches has been performed at the Instituto de Física de Cantabria, analyzing final states with two leptons of opposite electric charge (electrons or muons) and a large missing transverse momentum (p_T^{miss}) in events from a set of proton-proton collision data recorded by the CMS detector at a center of mass energy of $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 138 fb^{-1}. This thesis describes such an analysis and the adaptations needed for its reinterpretation.

Chapter 3

Theoretical Framework

3.1 The Standard Model and Its Limitations

The standard model (SM) of particle physics is a quantum field theory describing three of the four fundamental interactions, which are electromagnetic, weak, and strong interactions, through the gauge group $U(1)\times SU(2)\times SU(3)$. It also classifies all known elementary particles, including quarks, leptons, gauge bosons, and the Higgs boson. Although the SM is extremely successful in explaining a very wide range of experimental measurements with very high accuracy, it has several theoretical and observational shortcomings:

- It lacks a description of gravity.
- It does not explain the reason behind the small but non-zero neutrino masses.
- It cannot explain the matter–antimatter asymmetry in the universe.
- It does not offer a particle candidate for dark matter.
- It contains the hierarchy problem: why the Higgs boson mass is so much less than the Planck scale despite large quantum corrections.

These failings prompt the exploration of theories outside the SM, one of which is supersymmetry, a highly motivated and highly explored candidate. These limitations motivate physics beyond the SM, with supersymmetry being one of the most studied candidates.

3.2 Supersymmetric Models

Supersymmetry (SUSY) is an extension of the SM that introduces a symmetry between fermions and bosons. For each particle of the SM, SUSY predicts the existence of a superpartner with spin one-half unit greater or lesser:

- The SM fermions (quarks and leptons) are paired with a bosonic superpartners: sleptons and squarks.
- The SM bosons (gauge bosons and Higgs boson) are associated with fermionic superpartners, higgsinos, and gauginos.

The simplest realization of SUSY is the minimal supersymmetric standard model (MSSM), which contains the minimal superpartner content plus two extra Higgs doublets to cancel anomalies and generate up- and down-type quark masses. Upon electroweak symmetry breaking, higgsinos and gauginos mix to form the charginos and neutralinos.
SUSY can also solve the hierarchy problem by canceling out the large loop corrections to the mass of the Higgs boson. Also, in scenarios where a discrete quantum number called R-parity is conserved, the lightest supersymmetric particle (LSP) is stable and can be an interesting dark matter candidate. Thus, SUSY not only addresses key shortcomings of the SM but also provides natural dark matter candidates, as discussed next.

3.3 R-parity and Dark Matter Candidates

The R-parity of a particle is defined as:

$$R_p = (-1)^{3B+L+2S}$$

Where B is the particle baryon number
L is its lepton number, and
S is its spin.
All standard model particles have $R_p = +1$
While their SUSY partners have $R_p = -1$.
If R-parity is conserved:

- SUSY particles are produced in pairs.
- The lightest supersymmetric particle (LSP) is stable.

This makes the LSP an excellent dark matter candidate because it is neutral, weakly interacting, and cosmologically stable. The LSP in most models is the lightest neutralino, $\tilde{\chi}_1^0$, a mixture of the superpartners of the neutral Higgs and gauge bosons.

3.4 Simplified SUSY Models

To interpret experimental data in a tractable way, simplified models are often used. These models isolate a small number of SUSY particles and decay chains while fixing the rest of the spectrum to high masses to reduce model dependence.
In an analysis recently performed by a team at IFCA, three simplified scenarios are considered:

- Pair production of charginos $\tilde{\chi}_1^{\pm}$, each decaying either through intermediate sleptons or directly into a W boson and a neutralino.
- Pair production of sleptons $\tilde{\ell}$, each decaying into a lepton and the lightest neutralino $\tilde{\chi}_1^0$.
- Pair production of top squarks $\tilde{t}$, each decaying either via a top quark or via a chargino.

The analysis assumes R-parity conservation and focuses on final states with two oppositely charged (OC) leptons, plus two neutrinos and two neutralinos (LSPs) that lead to large missing transverse momentum (p_T^{miss}). In the case of the model with top squark pair production, two additional bottom quarks can be present in the final states.

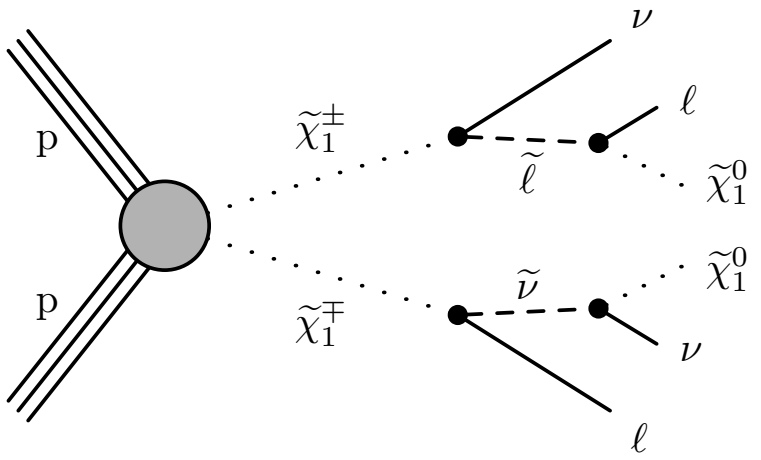


Figure 3.1: Decays via intermediate sleptons or sneutrinos to leptons and neutralinos

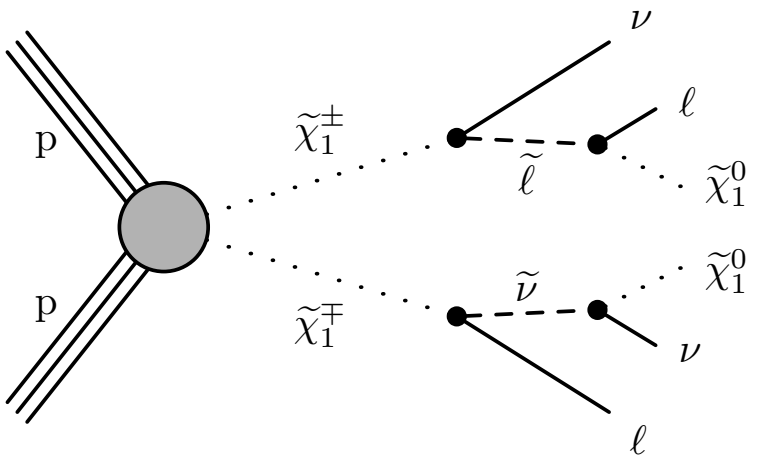


Figure 3.2: Direct decays into a W boson and the lightest neutralino $\tilde{\chi}_1^0$.

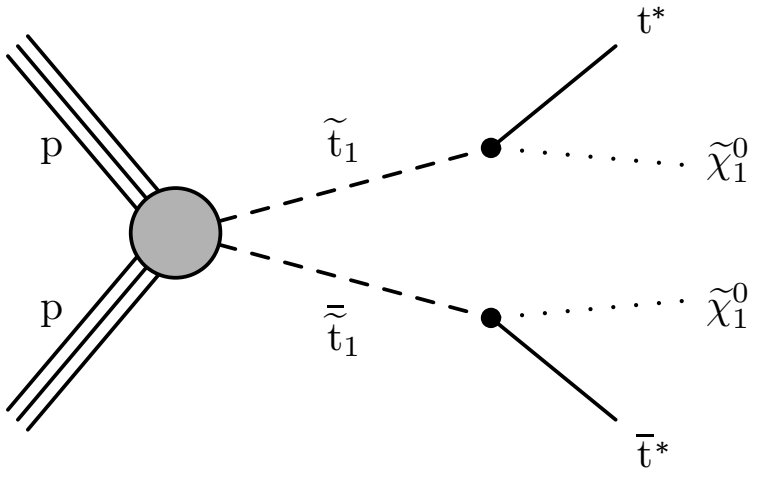


Figure 3.3: Decays into a top quark and the lightest neutralino

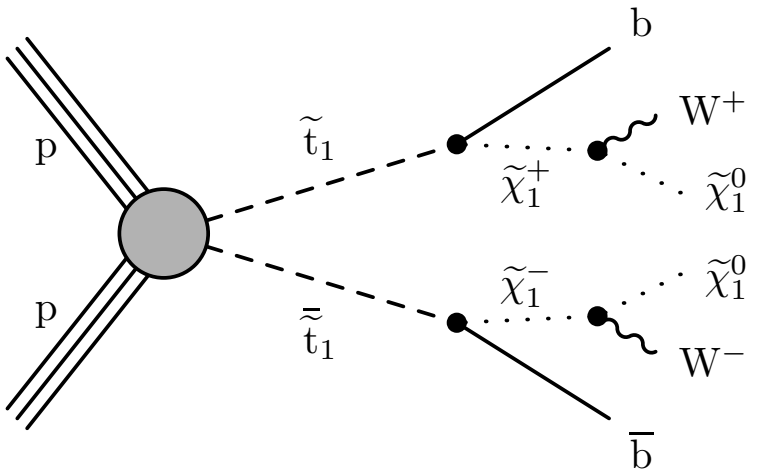


Figure 3.4: Direct decays into a bottom quark and a chargino, further decaying into a neutralino and a W boson

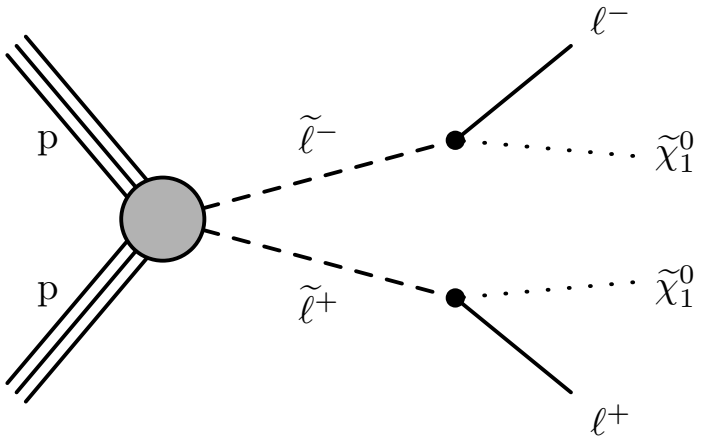


Figure 3.5: Direct decay of slepton into lepton and neutralino

3.5 The Phenomenological Minimal Supersymmetric Standard Model

Although the MSSM provides the minimal field content and interactions sufficient to introduce supersymmetry, it is accompanied by more than one hundred new parameters in the guise of soft-breaking masses, trilinear couplings, and phases. Such a large number of dimensions makes global analysis cumbersome. To avoid this, the phenomenological MSSM (pMSSM) is a reduced but general model.
The pMSSM is built on a set of simplifying assumptions that render the parameter space more tractable without losing the phenomenology of supersymmetry. The most widely used formulation assumes:

- There are no additional sources of CP violation, other than the one described by the SM.
- Minimal flavor violation (MFV), i.e., the only source of flavor-changing processes is the CKM matrix.
- Degeneracy of the first two generations of sfermions with extremely small Yukawa couplings, suppressing flavor-violating effects.
- R-parity conservation, rendering the lightest supersymmetric particle (LSP) stable.

With these assumptions, there are 19 free parameters (20 if the gravitino mass is also counted as one). They include gaugino masses, scalar masses for squarks and sleptons, third-generation trilinear couplings, the Higgsino mass parameter (μ), the Higgs vacuum expectation ratio ($\tan\beta$), and the mass of the pseudoscalar Higgs boson.
The pMSSM serves as a flexible framework for collider phenomenology, dark matter studies, and global SUSY fits. It strikes a balance between the predictiveness of highly constrained models and the generality of the full MSSM.

Chapter 4

The CMS Experiment

4.1 The Large Hadron Collider (LHC)

The CERN Large Hadron Collider (LHC) is the highest-energy particle accelerator in the world. It is a 27-kilometer superconducting magnet ring with several accelerating structures to boost the energy of the particles. Proton-proton collisions at the LHC are used to investigate the basic structure of matter at the highest laboratory energies. During Run 2 (2016–2018), the LHC operated at a center-of-mass energy of $\sqrt{s} = 13$ TeV and delivered on average an integrated luminosity of around 138 fb^{-1} to each of the four main experiments, one of which is CMS.

4.2 Overview of the CMS Detector

The Compact Muon Solenoid (CMS) is a multi-purpose detector designed to investigate a wide range of physics phenomena, such as the discovery of the Higgs boson, supersymmetry, and extra dimensions. The CMS detector is designed around a strong solenoid magnet, which allows for precise momentum measurement of charged particles.
The CMS detector is arranged in layers around the beam line to detect particles and measure their energies as they move away from the collision vertex. The detector consists of the following major subsystems: an inner tracker, electromagnetic and hadronic calorimeters, a superconducting solenoid, and a muon detection system.

4.3 Subsystems: Tracker, Calorimeters, Superconducting Magnet, and Muon System

4.3.1 Tracker System

The silicon tracker is the nearest subsystem to the interaction point and is used to measure the trajectories of charged particles. It consists of pixel and strip detectors and operates in a 3.8 T magnetic field.

4.3.2 Calorimeters

The electromagnetic calorimeter (ECAL) is designed to measure the energy of electrons and photons. It consists of a barrel and two endcap sections made of lead tungstate ($PbWO_4$) crystals. With its high granularity and rapid response, the ECAL provides excellent energy resolution for electromagnetic showers. Outside of the ECAL is the hadronic calorimeter (HCAL), which measures the energy of hadrons (e.g., neutrons, protons, pions). The HCAL is composed of scintillator and brass absorber tiles, optimized for high efficiency in detecting hadronic showers.

4.3.3 Muon System

The muon detector system is CMS's outermost component and consists of gas-ionization detectors implanted in the steel return yoke of the magnet. The system provides efficient identification and precise momentum measurement for muons, which are important signatures in many new physics searches, including the one for supersymmetric particles.

4.3.4 Magnet and Solenoid

A central element of the CMS detector is its superconducting solenoid magnet, which generates a powerful magnetic field of 3.8 T. This strong magnetic field causes the trajectories of charged particles to curve, enabling precise measurement of their momentum.

4.4 Trigger and Data Acquisition

Due to the extremely high rate of collisions at the LHC, only a fraction of events can be stored. The event rate is reduced from approximately 40 MHz to about 1 kHz by the CMS trigger system in two stages:

- The **Level-1 (L1) trigger**: a hardware-based system that makes decisions in microseconds.
- The **High-Level Trigger (HLT)**: a software-based system that uses advanced reconstruction algorithms to further select events of interest.

4.5 Data-Taking Period and Luminosity

The search is based on proton-proton collision data collected by the CMS experiment between 2016 and 2018, during Run 2 of the LHC. The integrated luminosity of the data used in this search is 138 fb^{-1}, recorded with stable beams and well-performing detectors throughout the period.

Chapter 5

Data and Simulation

5.1 Collision Data

This analysis uses proton-proton collision data collected by the CMS detector during the Run 2 era of the LHC from 2016 to 2018. The data correspond to an integrated luminosity of 138 fb^{-1} at a center-of-mass energy of $\sqrt{s} =$ 13 TeV. The analysis considers only periods of good beam conditions and detector performance.

5.2 Monte Carlo Simulation Samples

To estimate Standard Model (SM) backgrounds and verify the efficacy of the analysis technique, simulated Monte Carlo (MC) samples are used. These samples are generated for both SM processes and theoretical supersymmetry (SUSY) signal models. The major SM backgrounds, such as Drell–Yan, top quark pair production ($t\bar{t}$), single top, diboson (WW, WZ, ZZ), and triboson production, are simulated using packages such as MadGraph5_aMC@NLO, Pythia8, and Powheg, depending on the specific process.

5.3 Simulated Signal Models

Signal samples for simplified SUSY scenarios are simulated to model the expected event topologies and kinematic features. The scenarios considered include:

- Pair production of charginos ($\tilde{\chi}_1^{\pm}$), decaying either via intermediate sleptons or sneutrinos to leptons and neutralinos, or via W bosons.

- Slepton pair production ($\tilde{\ell}$) decaying to leptons and neutralinos ($\tilde{\chi}_1^0$).
- Pair production of top squarks ($\tilde{t}$), decaying either via top quarks or via intermediate charginos that subsequently decay into W bosons and neutralinos.

These simulated events are processed through a detailed CMS detector simulation based on GEANT4 and are reconstructed using the same software chain as the real collision data.

5.4 Pileup and Event Weighting

Simulated events are reweighted to match the pileup conditions observed in the data. Pileup refers to the presence of additional proton-proton interactions within the same or nearby bunch crossings. The distribution of the number of pileup interactions in the simulation is reweighted to match the measured distribution in data. Additional corrections are applied to account for differences between simulation and data in trigger efficiencies, lepton and jet reconstruction and identification efficiencies, and jet energy calibrations. These corrections are applied as event weights during the analysis to ensure proper normalization and accurate modeling of kinematic distributions.

Chapter 6

Event Reconstruction and Selection

6.1 Physics Object Reconstruction

Physics objects, such as electrons, muons, jets, and missing transverse momentum (p_T^{miss}), are reconstructed from the raw detector signals by the CMS particle-flow (PF) algorithm. This algorithm combines information from all subdetectors to identify and reconstruct individual particles in each event.

6.1.1 Leptons (Electrons and Muons)

Electrons are reconstructed by matching clusters of energy deposits in the electromagnetic calorimeter (ECAL) with tracks in the silicon tracker. Identification is based on shower shape variables, track-cluster matching, and track impact parameters. Muons are reconstructed by combining signals from the muon detection system with tracks found in the inner tracker. Muon identification employs quality requirements on the number of hits in both systems and on the quality of the global track fit.

6.1.2 Jets

Jets are clustered from particle-flow candidates using the anti-k_T algorithm with a distance parameter of 0.4 (AK4 jets). Jet energy corrections (JEC) are applied to account for the non-linear detector response and the contribution from pileup interactions. The jet energy resolution is also broadened in simulation to match the resolution measured in data.

6.1.3 Missing Transverse Momentum ($\boldsymbol{p_T^{\text{miss}}}$)

The missing transverse momentum, p_T^{miss}, is defined as the negative vector sum of the transverse momenta ($\vec{p}_T$) of all reconstructed particle-flow candidates in the event. It represents the momentum imbalance in the transverse plane and is a key signature for weakly interacting particles that escape detection, such as neutrinos or the lightest supersymmetric particle (LSP).

6.2 Object Identification and Isolation Criteria

To suppress backgrounds from misidentified objects and non-prompt leptons (e.g., from heavy-flavor decays or hadronic jets), strict identification (ID) and isolation criteria are applied to the reconstructed objects.

Leptons:

- Leptons must satisfy *tight* identification requirements as defined by the CMS Collaboration for Run 2 data.
- Leptons are required to be isolated. The relative isolation is defined as the scalar p_T sum of charged hadrons, neutral hadrons, and photons within a cone of size ΔR around the lepton track, divided by the lepton p_T. Contributions from pileup are subtracted from the neutral-hadron and photon components.
- Typical kinematic thresholds are applied, e.g., $p_T > 10$ GeV and $|\eta| < 2.4$ for electrons and muons.

Jets:

- Jets must pass jet identification criteria to reject fake jets from detector noise and pileup.
- Jets are selected with $p_T > 30$ GeV and $|\eta| < 2.4$ (for the 2016 data-taking period) or $|\eta| < 2.5$ (for the 2017 and 2018 periods).

Chapter 7

Search for supersymmetric particle pair production in final states with two oppositely charged leptons and large missing transverse momentum

7.1 Description of the Analysis Strategy

The analysis focuses on two distinct SUSY production mechanisms. A preliminary version of the analysis has been released by the CMS Collaboration for presentation at open conferences CMS-PAS-SUS-23-002.
Electroweak Production: Chargino/slepton pair production, studied across the full accessible mass range.
Strong Production: top squark pair production in the compressed mass scenario where the mass difference between the top squark ($\tilde{t}_1$) and the lightest neutralino ($\tilde{\chi}_1^0$) has a a value between the mass of the W boson and the mass of the top quark.

7.1.1 Baseline Event Selection

To distinguish a potential signal from the immense Standard Model background, the following selection criteria were applied.
Leptons: Exactly two isolated leptons (ee, $\mu\mu$, or eμ) with pseudorapidity $|\eta| < 2.4$ were selected, and the leading (trailing) lepton has a transverse momentum $p_T > 25(20)$ GeV.

Lepton Invariant Mass: The invariant dilepton mass, $m_{\ell\ell}$, was larger than 20 GeV to exclude low-mass resonances and non-prompt leptons. For same-flavour pairs (ee, $\mu\mu$), an additional requirement $|m_{\ell\ell} - m_Z| > 15$ GeV is applied to suppress $Z \to \ell\ell$ events.
High missing transverse momentum, $\vec{p}_T^{\,miss} > 160$ GeV.
Third Lepton Veto: events with a third lepton $p_T > 10$ GeV and $|\eta| < 2.4$ are vetoed, which is used to suppress background from WZ and ZZ production processes.

7.1.2 Discriminatory Variable: The stransverse mass (mT2)

The stransverse mass m_{T2} is defined for events with two decay chains, each yielding one visible particle and an invisible particle. It is calculated as:

$$m_{T2}(\ell\ell) = \min_{\vec{p}_T^{\,\mathrm{miss1}} + \vec{p}_T^{\,\mathrm{miss2}} = \vec{p}_T^{\,\mathrm{miss}}} \left[\max\left(m_T(\vec{p}_T^{\,\ell_1}, \vec{p}_T^{\,\mathrm{miss1}}),\ m_T(\vec{p}_T^{\,\ell_2}, \vec{p}_T^{\,\mathrm{miss2}}) \right) \right] \quad (7.1)$$

Where $\vec{p}_T^{\,\ell_1}$ and $\vec{p}_T^{\,\ell_2}$ are the transverse momenta of the two leptons, $\vec{p}_T^{\,\mathrm{miss}}$ is the total missing transverse momentum, and the minimization is performed over all possible partitions of $\vec{p}_T^{\,\mathrm{miss}}$ into two hypothetical invisible momenta, $\vec{p}_T^{\,\mathrm{miss1}}$ and $\vec{p}_T^{\,\mathrm{miss2}}$, assigned to each decay chain. Here, m_T denotes the standard transverse mass.
For the main SM backgrounds like $t\bar{t}$, tW, and WW production, where the missing momentum originates from the two neutrinos produced in the W bosons´ decays, the $m_{T2}(\ell\ell)$ distribution has a kinematic upper bound at m_W. In contrast, signal events include extra missing energy from neutralinos, causing their $m_{T2}(\ell\ell)$ spectrum to extend beyond this endpoint, thus providing a powerful discriminant against the dominant backgrounds.

7.1.3 Search Regions (SRs)

To maximize sensitivity, events are categorized into search regions (SRs) by separating the selected events into four mutually exclusive regions of increasing $\vec{p}_T^{\,\mathrm{miss}}$:

- **SR1:** Defined by $160 \leq p_T^{\mathrm{miss}} < 220$ GeV
- **SR2:** Characterized by $220 \leq p_T^{\mathrm{miss}} < 280$ GeV
- **SR3:** Characterized by $280 \leq p_T^{\mathrm{miss}} < 380$ GeV

- **SR4:** Characterized by $p_T^{\text{miss}} \geq 380$ GeV

This strategy utilizes the fact that signal events have heavier decaying particles, which lead to higher-p_T invisible particles and hence harder $\vec{p}_T^{\text{miss}}$ spectra. Therefore, while fewer overall events occur at higher-p_T^{miss} SRs, the signal-to-background ratio increases, making SR3 and SR4 particularly sensitive to very high-mass supersymmetry scenarios.
The sensitivity of the analysis is further enhanced by splitting the SRs according to the event content in multiplicity of jets and b-tagged jets, optimized for the signal production hypothesis under study.

Chargino and slepton pair production

The searches for the pair production of charginos and sleptons benefit from the fact that these processes typically do not involve b-quarks. The strategy is designed to suppress top-quark backgrounds.
Lepton pair flavour: events are categorized into same-flavour (SF) SRs and different-flavour (DF) SRs, depending on whether the two leptons have the same flavour or a different one. b-jet veto: the SRs require there to be no b-tagged jets.
Jet Multiplicity: In SR1 and SR2, lower $\vec{p}_T^{miss}$ bins, events are further split by jet multiplicity.
Control regions (CRs): Control regions (CRs) are defined like SRs but require ≥ 1 b-tagged jets to constrain the normalization of the $\mathrm{t\bar{t}}$ and tW backgrounds from the data. Each of the SRs and CRs here defined is split into $m_{T2}(\ell\ell)$ bins with the following edges: 0, 20, 40, 60, 80, 100, 160, 240, 370, in GeV. For SR2 and SR3, the last two $m_{T2}(\ell\ell)$ bins are merged, while for SR1, the last three bins are merged.

Search regions for top squark pair production

The search is optimized for compressed mass spectra, where the mass difference is $\Delta m = (m_{\tilde{t}_1} - m_{\tilde{\chi}_1^0})$. Signal regions (SRs) are defined by the number of b-tagged jets:

- **0 *b*-tag:** For small Δm (close to the W boson mass). b-jets are too soft to tag.
- **$\geq$1 *b*-tag:** For larger Δm, where b-jets are more energetic and easier to tag.

In SR3 and SR4, a high-p_{T} ISR jet is required to boost the system and increase missing transverse momentum, improving signal separation.
Each SR is split into **same-flavor (SF)** and **different-flavor (DF)** events, and further divided into $m_{\mathrm{T2}}(\ell\ell)$ bins: [0, 20, 40, 60, 80, 100, 160, ∞] GeV.

7.1.4 Statistical Analysis Method

A simultaneous maximum-likelihood (ML) fit to the $m_{T2}(\ell\ell)$ distribution in all the SRs and CRs is performed to extract the signal (as described in Section 7.1.3), with binning optimized separately for chargino/slepton and stop searches. Low m_{T2} bins constrain dominant Standard Model (SM) backgrounds ($t\bar{t}$, tW, WW), while higher bins enhance sensitivity to SUSY signals.

7.2 Background Estimation

The events that have been selected in the SRs comprise a potential signal coming from supersymmetric processes and a non-negligible contribution from SM background processes. Accurate estimation of these SM backgrounds is imperative for sensitivity in the search and correct interpretation of the results. The dominant backgrounds arise from the production of top quark pairs ($\mathrm{t\bar{t}}$), single top quarks in association with a W boson (tW), and diboson production (WW mainly). The normalization of these backgrounds is determined by the ML fit, as mentioned in Section 7.1.4. Their $m_{T2}(\ell\ell)$ distribution has a natural endpoint at the W boson mass, and events enter into the relevant region for signal extraction mainly because of detector resolution effects. For this reason, we study the modeling of the $m_{T2}(\ell\ell)$ distribution for these processes in dedicated CRs in data described in Section 7.2.1. Subdominant contributions arise from Z+X production processes such as WZ, ZZ, $\mathrm{t\bar{t}Z}$, and Drell-Yan (DY) production. These backgrounds are constrained using a combination of data-driven techniques and simulation-based predictions, and their normalizations are ultimately determined in the ML fit to data through dedicated CRs as described in Section7.1.4.

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

The reconstruction of the $m_{T2}(\ell\ell)$ observable in $t\bar{t}$, tW, and WW events is sensitive to various detector resolution effects that can drive its value beyond the W mass kinematic endpoint: p_T^{miss}-measurements and nonprompt

leptons.
A CR to test the modeling of $m_{T2}(\ell\ell)$ at high p_T^{miss} is built selecting events from WZ$\rightarrow 3\ell 1\nu$ production and emulating the shape of the $m_{T2}(\ell\ell)$ observable in top quark and WW backgrounds by taking the p_T of lepton from the decay of the Z boson with the same charge as the lepton from the decay of the W boson, and adding it vectorially to the p_T^{miss}, just as it would be if it were a neutrino instread of a charged lepton. Technically, events are selected by requiring exactly three leptons and no b-tagged jets. The leptons from the Z boson decay are identified by taking the lepton pair of the same flavour and opposite charge whose invariant mass falls closer to the Z boson mass.
The simulation is found to slightly underestimate the data in the tail of the $m_{T2}(\ell\ell)$ distribution, and a correction is derived by this trend and applied to the $m_{T2}(\ell\ell)$ shape in $t\bar{t}$, tW, and WW events.
The second potential source of mismodeling in the tails of the $m_{T2}(\ell\ell)$ distributions arises from nonprompt leptons originating, for instance, from semileptonic decays of b-hadrons in b jets or from hadronic jets accidentally passing the lepton selection.
In simulated events, nonprompt leptons contribute less than 1% of the expected background across the different SRs. However, at large values of $m_{T2}(\ell\ell)$ and p_T^{miss}, events with nonprompt leptons constitute up to 20% of the $t\bar{t}$ background. The modeling of the rate of nonprompt leptons in simulation is studied in events passing the same selection as for the SRs, but where the two leptons have the same sign. The largest contribution to this sample comes from $t\bar{t}$ events with a nonprompt lepton. Based on the observed agreement between data and expected background, correction factors are derived for the nonprompt lepton rate in simulation, with values ranging from 1.1 to 1.4 across the data-taking periods.

7.2.2 Backgrounds from Z+X production

The contributions of subleading backgrounds from Z+X production (such as ttZ, WZ, ZZ, and Drell–Yan events) are also tested in CRs, which are included in the ML fit to constrain their normalizations:

- A three-lepton CR, with a pair of same-flavour, opposite-sign leptons, and veto on b-tagged jets, is used to constrain the normalization of the WZ events and validate the description of their $m_{T2}(\ell\ell)$ as provided by the simulation.

- A four-lepton CR with no b-tagged jets is analogously used to estimate the ZZ $\rightarrow 2\ell 2\nu$ background by reconstructing it with ZZ $\rightarrow 4\ell$ events. The four leptons are required to form two pairs of leptons of the same

flavor, opposite signs, of which at least one has an invariant mass compatible with the Z boson mass ($|m_{\ell\ell} - m_Z| < 15$GeV). The p_T of the lepton pair with invariant mass closer to the mass of the Z boson is added vectorially to the p_T^{miss} of the event.

- A CR with at least three leptons and at least a b-tagged jet is selected to constrain the normalization of ttZ production. Events with four leptons are used to validate the $m_{T2}(\ell\ell)$ shape for this process.
- The modeling of the $m_{T2}(\ell\ell)$ observable is tested in events with exactly two leptons of the same color, opposite signs, with invariant mass within 15 GeV of the Z boson mass. This CR is not included in the final ML fit since it is not sufficiently pure.

7.3 Systematic Uncertainties

The fit is constrained by systematic uncertainties that impact the normalization and shape of the distributions of signal and background processes. Each source of uncertainty is individually addressed and propagated to the final signal and background yields by varying the respective parameter across its uncertainty range and propagating the variation through the full analysis chain. They are together known as experimental and theoretical uncertainties.

7.3.1 Experimental Uncertainties

Experimental uncertainties are concerned with the measurement and reconstruction of physical objects:

- **Integrated Luminosity:** The uncertainty on the measured integrated luminosity is 1.2–2.5%, depending on the data-taking year.
- **Lepton Identification and Isolation:** Scale factors applied to correct simulated lepton efficiencies have an associated uncertainty, which is typically less than 3% per lepton.
- **Jet Energy Scale (JES) and Resolution (JER):** The reconstructed jet energy scale and resolution are varied within their uncertainties, influencing the $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation and event classification.
- ***b*-tagging Efficiency:** The b-tagging efficiency to recognize jets from bottom quarks is corrected using data-to-simulation scale factors, which have uncertainties of 1–6%.

- **Unclustered Energy and $p_{\mathrm{T}}^{\mathrm{miss}}$ Modeling:** The energy scale of unclustered low-p_{T} particles is varied, propagating an uncertainty to the $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation.

- **Trigger Efficiency:** An uncertainty of 2% is assigned to the efficiency measured for the triggers used in selecting events.

7.3.2 Theoretical and Modeling Uncertainties

The theoretical uncertainties are associated with the simulation of physical processes:

- **Renormalization and Factorization Scales:** The renormalization and factorization scales used in the matrix element calculation are alternatively doubled or halved to probe the impact on the acceptance.

- **Parton Distribution Functions (PDFs):** The proton's incomplete knowledge of structure is probed according to the standard NNPDF3.1 error sets.

- **Background Normalization:** As described in Section 7.2, dominant backgrounds are constrained in the fit under their normalizations. A 50% prior uncertainties is placed for minor backgrounds.

- **Top quark p_{T} Reweighting:** Simulated $t\bar{t}$ events are reweighted to match them to the top quark p_{T} spectrum observed in data, and the process uncertainty is added.

- **Initial State Radiation (ISR):** In signal models, ISR is improved through a reweighting procedure, and an uncertainty is derived from closure tests.

7.3.3 Impact Assessment

The impact of these systematic uncertainties on the SM background predictions and a representative signal model is quantified in Table 7.1 and Table 7.2, respectively. The values represent the range of their effect on the total yield across the analysis regions and their maximum effect on the shape of the $m_{\mathrm{T2}}(\ell\ell)$ distribution.

Source of uncertainty	**Change in yields**	**Change in shape**
Integrated luminosity	1–3%	–
Trigger efficiency	2%	<1%
Pileup	≤2%	2–20%
Jet energy scale	3–8%	2–10%
Jet energy resolution	1–2%	2–8%
Unclustered energy	1–2%	2–13%
Lepton identification	2–4%	≤15%
b-tagging efficiency	≤5%	≤6%
b-tagging mistag rate	<1%	≤3%
Simulated sample statistics	≤3%	4–37%
Renorm./fact. scales	2–23%	1–15%
PDFs	≤2%	≤9%

Table 7.1: **Sizes of systematic uncertainties for SM background processes.** The first column shows the range of the effect on the global background normalization across different signal regions. The second column quantifies the maximum effect on the $m_{\mathrm{T2}}(\ell\ell)$ shape. Adapted from [[1]].

Source of uncertainty	**Change in yields**	**Change in shape**
Integrated luminosity	1–3%	–
Jet energy scale	≤5%	≤9%
Jet energy resolution	≤3%	≤14%
Lepton identification	2–7%	≤3%
Simulated sample statistics	3–15%	6–22%
Renorm./fact. scales	≤1%	<1%
ISR reweighting	≤2%	≤8%
$p_{\mathrm{T}}^{\mathrm{miss}}$ (FastSim)	≤11%	≤24%

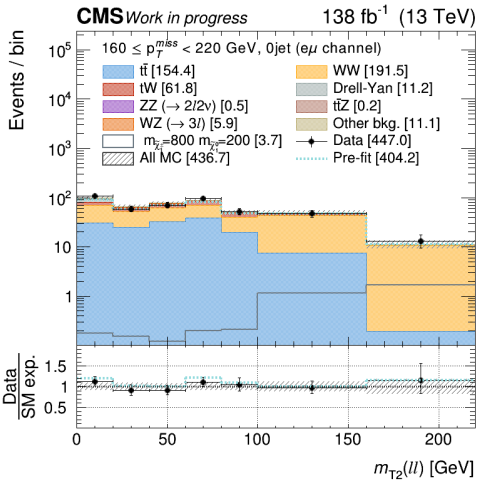
Table 7.2: **Sizes of systematic uncertainties for a representative signal model** ($m_{\tilde{\chi}_1^\pm} = 800$ GeV, $m_{\tilde{\chi}_1^0} = 200$ GeV, decaying via sleptons). Adapted from [[1]].

The uncertainties are treated as correlated or uncorrelated across years and analysis regions following the recommendations of the physics objects groups of the CMS Collaboration, ensuring a robust and accurate limit extraction.

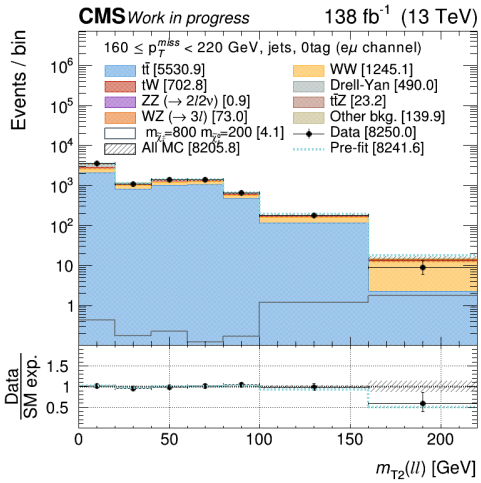
7.4 Results and Interpretation

7.4.1 Post-Fit Validation and Observed Yields

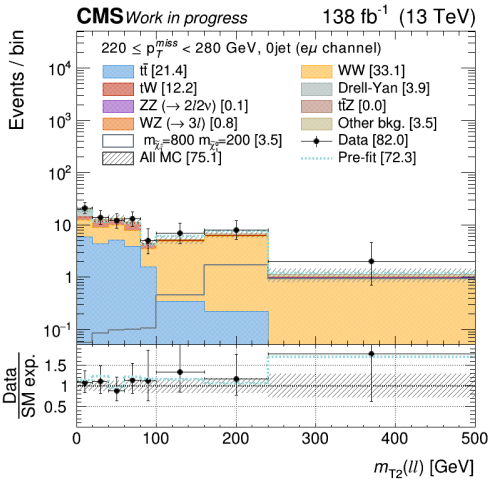
Before presenting the final limits, the agreement between the observed data and the fitted background model is validated. Figure 7.1 shows the post-fit distributions of the final discriminant, $m_{\mathrm{T2}}(\ell\ell)$, for events of different flaviour (DF) in some representative search regions, categorized by missing transverse momentum ($p_{\mathrm{T}}^{\mathrm{miss}}$) and jet multiplicity.
In all regions, the data (black points) are in good agreement with the post-fit SM background prediction (stacked histograms). The expected distribution for a reference signal point (gray histogram) is shown for comparison. The ratio panels at the bottom confirm that the agreement is consistent within the post-fit uncertainties (hatched band) throughout the $m_{\mathrm{T2}}(\ell\ell)$ spectrum. This demonstrates the robustness of the background modeling and the fit procedure.



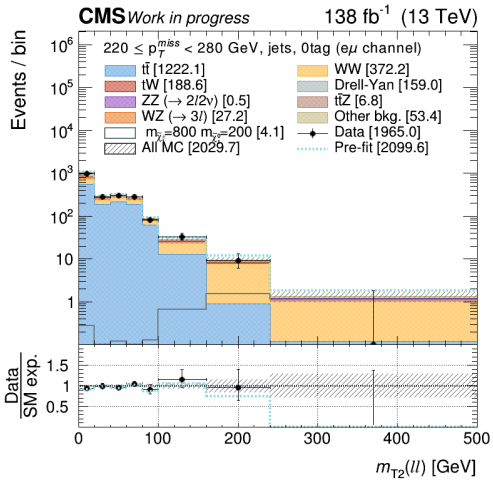
(a) SR1$^{0j}_{0b\text{tag}}$: $160 \leq p_{\mathrm{T}}^{\mathrm{miss}} < 220$ GeV, 0 jets.



(b) SR1$^{\geq 1j}_{0b\text{tag}}$: $160 \leq p_{\mathrm{T}}^{\mathrm{miss}} < 220$ GeV, $\geq$1 jet.



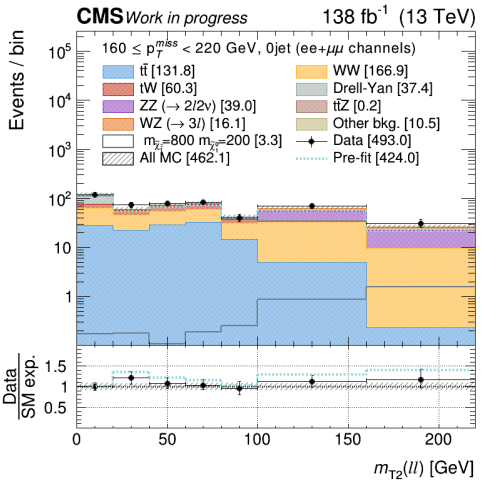
(c) SR2$^{0j}_{0b\text{tag}}$: $220 \leq p_{\mathrm{T}}^{\mathrm{miss}} < 280$ GeV, 0 jets.



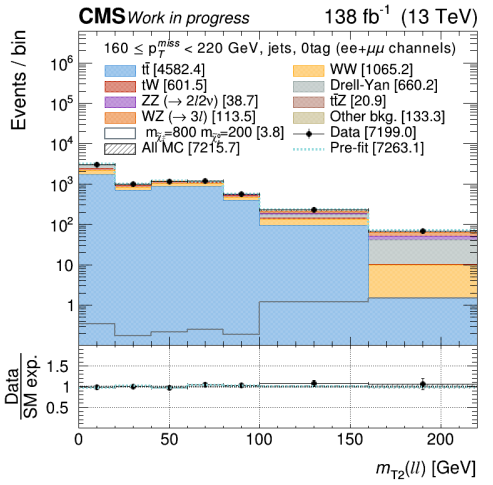
(d) SR2$^{\geq 1j}_{0b\text{tag}}$: $220 \leq p_{\mathrm{T}}^{\mathrm{miss}} < 280$ GeV, $\geq$1 jet.

Figure 7.1: **Post-fit $m_{\mathrm{T2}}(\ell\ell)$ distributions for Different-Flavor (DF) events.**

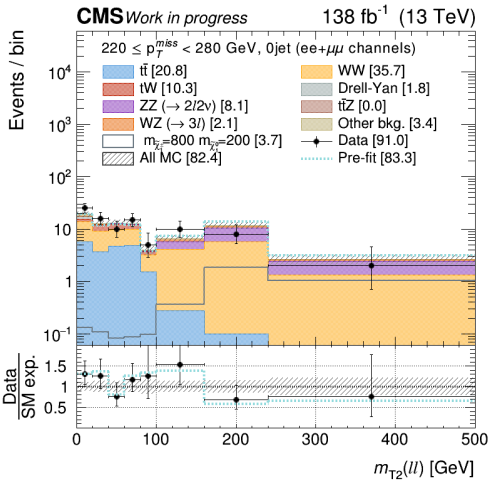
The data (points) are compared to the post-fit SM background prediction (stacked histograms) and the pre-fit total SM expectation (cyan dashed line). The grey histogram shows the expected distribution for a benchmark signal model of chargino pair production ($m_{\tilde{\chi}_1^{\pm}} = 800$ GeV, $m_{\tilde{\chi}_1^0} = 200$ GeV). The lower panels show the ratio of data to the post-fit background prediction, with the hatched band representing the total post-fit uncertainty.



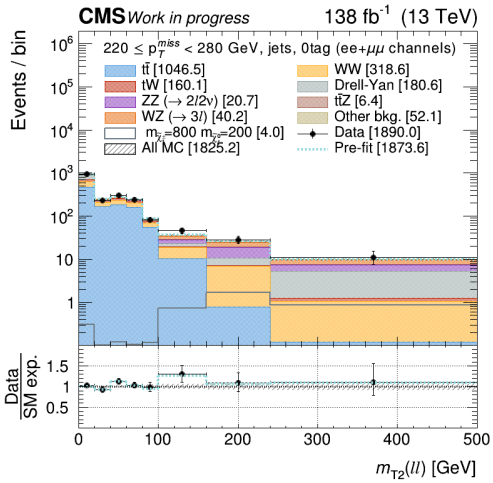
(a) $SR1^{0j}_{0btag}$: SF events



(b) $SR1^{\geq 1j}_{0btag}$: SF events



(c) $SR2^{0j}_{0btag}$: SF events



(d) $SR2^{\geq 1j}_{0btag}$: SF events

Figure 7.2: **Post-fit $m_{T2}(\ell\ell)$ distributions for Same-Flavor (SF) events.** The structure of this figure is identical to Fig. 7.1 but for the SF final state.

The same good agreement between data and the background-only prediction is observed for the same-flavor (SF) final states, as shown in Figure 7.2

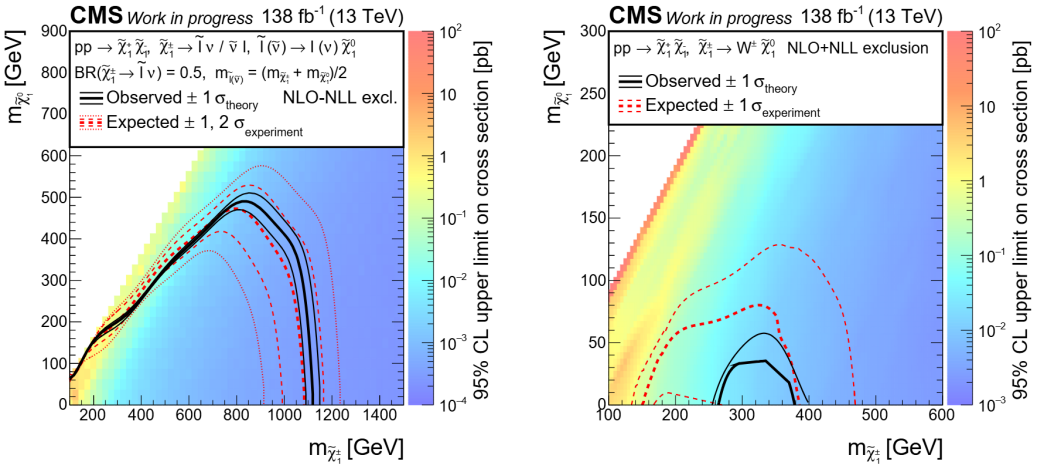
7.5 Limit Setting Procedure

Since no new physics signal is visible, upper limits are set on the production cross sections of supersymmetric particles. The limits are derived from a

simultaneous binned maximum-likelihood fit to the $m_{\mathrm{T2}}(\ell\ell)$ distribution in all signal regions (SRs) and dedicated control regions (CRs). This approach allows the normalizations of the major backgrounds ($\mathrm{t\bar{t}}$, tW, WW, WZ, ZZ, $\mathrm{t\bar{t}Z}$) to be constrained directly from data.
The CL_s method with a profile likelihood ratio is used as the test statistic. All systematic uncertainties are incorporated as nuisance parameters in the likelihood fit. The observed and expected 95% confidence level (CL) limits are computed using the asymptotic approximation.

7.6 Exclusion Limits in Simplified Models

The results are interpreted in the framework of simplified SUSY models. The observed and expected 95% CL exclusion contours are derived by comparing the upper limits on the cross section with the theoretical NLO+NLL production cross sections.



(a) Chargino pair production with cascade decay via sleptons ($\tilde{\chi}_1^{\pm} \to \ell\nu\tilde{\chi}_1^0$). Excludes chargino masses up to 1100 GeV for neutralino LSP masses up to 480 GeV.

(b) Chargino pair production with direct decay to W bosons ($\tilde{\chi}_1^{\pm} \to W\tilde{\chi}_1^0$). The exclusion is less stringent due to the smaller leptonic branching fraction of the W boson.

Figure 7.3: Exclusion limits at 95% CL in the $(m_{\tilde{\chi}_1^{\pm}}, m_{\tilde{\chi}_1^0})$ plane for chargino pair production. Panel (a) shows the cascade decay via sleptons, while panel (b) shows the direct decay to W bosons. The thick black line indicates the observed exclusion region, and the red dashed line indicates the expected exclusion.

The most stringent limits obtained in this search are summarized below (see Fig. 7.3):

- For chargino pair production with decay via sleptons ($\widetilde{\chi}_1^{\pm} \to \ell\nu\widetilde{\chi}_1^0$), chargino masses up to **1100 GeV** are excluded for neutralino LSP masses up to **480 GeV**.
- For slepton pair production ($\widetilde{\ell} \to \ell\widetilde{\chi}_1^0$), slepton masses are excluded up to **700 GeV** for LSP masses up to **360 GeV**.
- For chargino pair production with direct decay to W bosons ($\widetilde{\chi}_1^{\pm} \to W\widetilde{\chi}_1^0$), the exclusion is less stringent due to the smaller leptonic branching fraction of the W boson.

Chapter 8

Global pMSSM Interpretation and Implications

The analysis presented in the previous chapters looked for SUSY in simplified model scenarios. To finalize what the entire dataset tells us about the full SUSY theory, a global statistical interpretation is now performed within the more realistic 19-parameter phenomenological MSSM (pMSSM) [2].

8.1 The pMSSM Framework and Analysis Strategy

The phenomenological MSSM (pMSSM) is a realistic version of supersymmetry that reduces the (100) free parameters of the general MSSM to a manageable set of 19 defined at the SUSY scale,

$$Q_{\text{SUSY}} = \sqrt{m_{\tilde{t}_1} m_{\tilde{t}_2}},$$

Where $m_{\tilde{t}_1}$ and $m_{\tilde{t}_2}$ are the masses of the two stop eigenstates. This framework captures realistic SUSY features, including:

- A broad and varied SUSY particle mass spectrum,
- Complex particle mixings (e.g. neutralino and chargino sectors),
- And direct connections between collider observables and dark matter phenomenology.

A global Bayesian fit is carried out, combining the results of multiple CMS searches. The impact of the CMS data is evaluated by using a combined Monte Carlo simulation, from which posterior probability densities, survival

probabilities, and Bayesian significances are inferred. This makes it possible to determine which models are **ruled out** and which **survive**, gaining a complete picture of where SUSY could still be hiding.

8.2 Global Impact on the Supersymmetric Parameter Space

The combined CMS data set places strong constraints across the pMSSM parameter space. The main findings are as follows.

- **Electroweak Particles:** Charginos ($\tilde{\chi}_1^{\pm}$) and next-to-lightest neutralinos ($\tilde{\chi}_2^0$) lighter than ∼200 GeV are largely excluded. However, a Higgsino-like LSP around 350 GeV remains consistent with all constraints and is mildly preferred in the posterior distributions.
- **Colored Particles:** Gluinos and squarks are pushed to very high mass values. Models with colored states lighter than ∼1 TeV are almost entirely excluded, and the most probable gluino mass lies around 2.3 TeV.

8.3 Survival probability

The survival probability measures which pMSSM models remain consistent with the data. Figure 8.1 illustrates these results:

- **Top row:** $\Delta m(\tilde{\chi}_1^{\pm}, \tilde{\chi}_1^0)$ vs $m(\tilde{\chi}_1^0)$,
- **Bottom row:** $\Delta m(\tilde{\chi}_2^{\pm}, \tilde{\chi}_1^0)$ vs $m(\tilde{\chi}_1^0)$.

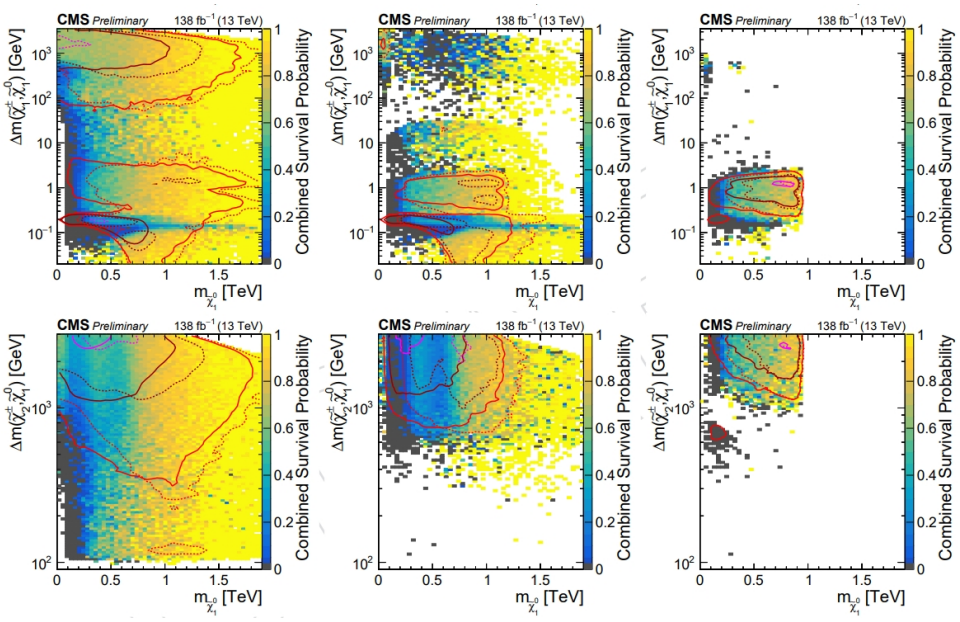


Figure 8.1: Survival probabilities for representative pMSSM parameters in the plane of $\Delta m(\chi_1^{\pm}, \chi_1^0)$ vs $m(\chi_1^0)$ masses (top row) and $\Delta m(\chi_2^{\pm}, \chi_1^0)$ vs $m(\chi_1^0)$ masses (bottom row). The probabilities on the left columns are derived from a global Bayesian fit that combines multiple CMS SUSY searches at $\sqrt{s} = 13$ TeV. On the middle column, experimental constaints from DM experiments are included. On the right column, thoeretical assumptions are also considered. The color scale ranges from **yellow** (higher survival fraction) to **blue/black** (lower/excluded regions).

Figure 8.1 illustrates the global impact of CMS searches on the pMSSM parameter space. **Top row:** The heatmaps show the posterior probability densities in the plane of $\Delta m(\chi_1^{\pm}, \chi_1^0)$ vs $m(\chi_1^0)$ masses. Yellow regions correspond to parameter points more consistent with the data, while blue indicates regions that are strongly constrained. The first column shows the result of the fit based on CMS search results; the second one shows the same when experimental constraints from DM experiments are added; the third one shows the same when theoretical assumptions on DM are also considered. **Bottom row:** the same as for the top row but for the plane of $\Delta m(\chi_2^{\pm}, \chi_1^0)$ vs $m(\chi_1^0)$ masses.

8.4 Conclusion of the Global Interpretation

The global pMSSM interpretation leads to the following conclusions:

- Large regions of SUSY parameter space are excluded, especially for light electroweak states.
- Surviving models typically feature heavy particles or compressed mass spectra, confirming the importance of advanced search strategies.
- The most consistent scenarios involve Higgsino-like dark matter candidates.

This overall analysis confirms that the searches for supersymmetry in the future will need to focus on challenging cases with heavy masses, compressed spectra, and complex particle decays. This will determine the way forward for future studies at the LHC and beyond.

Chapter 9

Conversion of Analysis Inputs to pMSSM Framework

The scope of this project is to develop an interface between the results of the search for supersymmetric particle production presented in Chapter7 and the framework used for the pMSSM effort of the CMS Collaboration presented in Chapter8, to allow for a reinterpretation of the former in the context of the pMSSM. In this chapter, we present the technical implementation of the developed interface.
In the pMSSM interpretation, a scan of the 19 fundamental parameters is done, both assuming prompt decays of the supersymmetric particles and in the context of long-lived scenarios. Simulated signal events are produced for each choice of parameter values (each choice is referred to as a parameter point in the following). Given the huge number of parameter points considered (about 500000), the expected yields of signal events are not stored in histograms (`TH1`/`TH2`/`TH3` objects in `ROOT`) but rather in `THnSparse` objects, i.e. multidimensional histograms that allocate memory only for bins with no zero content.
The objective of the interface is to classify the events into the search regions (SRs) of the simplified model analysis (as in 7), apply the appropriate kinematic and tagging selections and the event weight corrections, and finally fill the `THnSparses` with the appropriate parameter point coordinates for subsequent statistical analysis in the pMSSM framework. The expected contribution from SM processes to the SRs is taken directly from the results of the original analysis.

9.1 Parameter Point Identifiers

Each parameter point of the pMSSM is labeled by two integer-valued identifiers (IDs) that uniquely identify it. The pMSSM identifiers are used to map each simulated signal event to the set of pMSSM parameters it was created for, and completely determine the underlying physics.
The `THnSparse` object used in this study is therefore set up with three dimensions. The first two dimensions are dedicated to the pMSSM IDs, while the third dimension designates the analysis bin. The first pMSSM ID takes on values between 1 and 600, the second pMSSM ID takes on values between 1 and 144855. The analysis bin ranges between 1 and 174, where bin 1 is for the total number of events before any selection, and bins 2 to 174 are associated with the different SRs (and $m_{T2}\ell\ell$ bins) and CRs described in Sections7.1 and 7.2.

9.2 Event Selection

We initially check whether each event contains exactly two leptons (for the SRs) or more than two leptons (for the CRs), and then retrieve the main observables from physics: missing transverse momentum (p_T^{miss}), number of jets, event weight for b-tagging, and stransverse mass ($m_{T2}(\ell\ell)$). This sets the stage to identify the SR bin of the `THnSparse` object to which each event belongs.

9.3 Definition of the Search Region

Search regions are defined for events based on the number of leptons (and their flavour), the missing transverse momentum and the multiplicity of jets and b-tagged jets. Events with exactly two leptons are assigned to one of the SRs described in Section 7.1.3, while events with more than two leptons are assigned to the CRs described in Section 7.2.2. In particular, the SRs for the chargino and slepton pair production (described in Section 7.1.3) are used in this study, as they are designed with a more general scope.

9.4 $m_{T2}(\ell\ell)$ Binning

The $m_{T2}(\ell\ell)$ observable is used to further subdivide the SRs, as described at the end of Section 7.1.3: for SR1, 20 GeV wide bins from 0 to 100 GeV, and

two further bins above 100 GeV ($100 < m_{T2}(\ell\ell) < 160$ GeV and $m_{T2}(\ell\ell) > 160$ GeV) are used; for SR2, SR3, and SR4, the $m_{T2}(\ell\ell) > 160$ GeV bin is split in two bins ($160 < m_{T2}(\ell\ell) < 240$ GeV and $m_{T2}(\ell\ell) > 240$ GeV); for SR4, the $m_{T2}(\ell\ell) > 240$ GeV bin is further split into two other bins ($240 < m_{T2}(\ell\ell) < 370$ GeV and $m_{T2}(\ell\ell) > 370$ GeV). Each of the $m_{T2}(\ell\ell)$ bins defines an analysis bin of the `THnSparse` object.

9.5 Histogram Filling

Finally, the `THnSparse` multidimensional histogram is filled with coordinates the (`pMSSMid1`, `pMSSMid2`, `srbin`). Each event enters the histograms weighted by the same data/MC efficiency scale factors and correction weights (such as the pileup reweighting described in Section **??**) used in the original analysis but the signal cross section and integrated luminosity, which are handled directly by the pMSSM interpretation framework.

9.6 Impact on the Analysis

This coding scheme assigns each simulated signal event to the appropriate SR and m_{T2} bin. The generated THnSparse is the main input to the statistical interpretation, allowing for limits to be imposed on the pMSSM parameter space.

Chapter 10

Summary of Key Findings

Conclusion and Outlook

This thesis presents a reinterpretation of results from a previous supersymmetry search conducted at IFCA. While the original analysis used simplified models, here the results are interpreted within the more general and realistic 19-parameter phenomenological Minimal Supersymmetric Standard Model (pMSSM). The search focused on supersymmetry signal production in final states with two oppositely charged leptons and significant missing transverse momentum, using 138 fb^{-1} of proton–proton collision data at $\sqrt{s} = 13$ TeV, collected by the CMS experiment.

The search set stringent exclusion limits on supersymmetric particle production within the simplified model framework. In the case of pair production of charginos with intermediate decays, masses of charginos up to 1100 GeV are excluded at the 95% confidence level for a massless lightest neutralino. For direct slepton pair production, sleptons with masses up to 700 GeV are excluded. These results extend the exclusion limits of previous searches and currently provide the strongest constraints on these processes.

The results of this search are to be combined with those from other CMS studies and interpreted within the more general 19-parameter phenomenological MSSM (pMSSM) framework. During my master's study, I developed a software interface to analyze the expected pMSSM signal yield within the framework of the original simplified model search. This work provided the crucial input for reinterpreting the search results as constraints on the pMSSM parameter space. The global fit for the pMSSM reinterpretation, combining also the results from other CMS analyses, is now being carried out by the CMS Collaboration. The findings of this work will have a significant impact on the future of the field, indicating which regions of the pMSSM

parameter space are excluded and which ones should instead be the focus of further searches.

Bibliography

[1] CMS Collaboration. Search for supersymmetry in events with two leptons and missing transverse momentum in proton-proton collisions at $\sqrt{s} =$ 13 TeV. Technical Report CMS-PAS-SUS-23-002, CERN, 2023. URL `https://cds.cern.ch/record/2852371`.

[2] CMS Collaboration. Phenomenological MSSM interpretation of CMS searches in pp collisions at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2024. URL `https://cds.cern.ch/record/2897876`.

Appendix A

Code Snippets and THnSparse Configuration

This appendix collects the key Python scripts used in the analysis. The code is written in Python using ROOT libraries and NumPy.

A.1 Event Selection and Signal Region Binning

A.1.1 Event Selection

```
if chain.nLepton == 2:
    ptmiss = chain.ptmiss
    njets = chain.nCleanJet
    btagweight = getattr(chain, btagwp)
    nobtagweight = 1 - btagweight

    # Extract pMSSM model IDs
    pmssid1 = chain.pMSSMid1
    pmssid2 = chain.pMSSMid2

    # Lepton flavors
    pdg0 = chain.Lepton_pdgId[0]
    pdg1 = chain.Lepton_pdgId[1]
    isDF = (abs(pdg0) != abs(pdg1))
```

A.1.2 Signal Region Definition

```
if 160 <= ptmiss < 220:
    region = "SR10jet" if njets == 0 else "SR1jets"
elif 220 <= ptmiss < 280:
    region = "SR20jet" if njets == 0 else "SR2jets"
elif 280 <= ptmiss < 380:
    region = "SR30tag"
elif ptmiss >= 380:
    region = "SR40tag"
else:
    region = "other"
```

A.1.3 Region-to-Bin Mapping

```
def get_srbin(region, isDF, mt2ll):
    srbase = region_to_srbin.get(region, -1)
    if srbase == -1:
        return -1

    srbin = 2 * srbase
    if isDF:
        srbin += 1

    if mt2ll >= 0:
        mt2llbin = mt2bin(mt2ll, region)
        srbin = get_srmt2bin(region, mt2llbin, srbin)

    return srbin
```

A.1.4 mT2 Binning

```
def mt2bin(mt2ll, region):
    if mt2ll < 20: return 0
    elif mt2ll < 40: return 1
    elif mt2ll < 60: return 2
    elif mt2ll < 80: return 3
    elif mt2ll < 100: return 4
    elif mt2ll < 160: return 5
    elif "SR1" in region: return 6

    # SR2 and SR3 bins
```

```
    if 160 <= mt2ll < 300:
        return 6
    elif "SR2" in region or "SR3" in region:
        return 7

    # SR4 bins
    if 240 <= mt2ll < 370:
        return 7
    elif mt2ll >= 370:
        return 8

    return -1
```

A.1.5 Histogram Filling

```
if srbin >= 0:
    if njets == 0:
        coordinates = numpy.float64([pmssid1, pmssid2,
            srbin])
        thnsparse.Fill(coordinates, weight)
    else:
        # Case 1: no b-tag
        coordinates_jets = numpy.float64([pmssid1,
            pmssid2, srbin])
        thnsparse.Fill(coordinates_jets, weight *
            nobtagweight)

        # Case 2: with b-tag
        if "jets" in region:
            region_tag = region.replace("jets", "tags")
            srbin_tag = get_srbin(region_tag, isDF,
                mt2ll)
            if srbin_tag >= 0:
                coordinates = numpy.float64([pmssid1,
                    pmssid2, srbin_tag])
                thnsparse.Fill(coordinates, weight *
                    btagweight)
```

A.1.6 Control Regions

```
    elif opt.addcr and chain.nLepton>=3:

                if ptmiss_WZ>=160. or ptmiss_WZ>=160. or
                    ptmiss_WZ>=0. or ptmiss_ttZ>=0:

                        crbin = -1

                        if chain.nLepton==3 and
                           nTightLepton==3 and chain.
                           deltaMassZ_WZ<999. and
                           ptmiss_WZ>=160.:

                            crbin = get_CRbin(ptmiss_WZ,
                                njets)
                            if crbin>=0:
                                coordinates_cr = numpy.
                                   float64([pmssid1,
                                   pmssid2, sr_nbins+
                                   crbin+1])
                                thnsparse.Fill(
                                   coordinates_cr,
                                   weight * nobtagweight
                                   )

                        if chain.nLepton==4 and
                           nTightLepton>=3 and chain.
                           deltaMassZ_ZZ<15.  and
                           ptmiss_ZZ>=160.:

                            crbin = get_CRbin(ptmiss_ZZ,
                                njets)
                            if crbin>=0:
                                coordinates_cr = numpy.
                                   float64([pmssid1,
                                   pmssid2, sr_nbins+6+
                                   crbin+1])
                                thnsparse.Fill(
                                   coordinates_cr,
                                   weight * nobtagweight
                                   )
```

```
21                  if chain.nLepton>=3 and
                       nTightLepton>=3 and njets>=2
                       and (ptmiss_WZ>=0. or
                       ptmiss_ttZ>=0):
22
23                      crbin = -1
24                      btagweight_ttZ = btagweight
                           if chain.nLepton==4 else
                           getattr(chain,btagwp.
                           replace('_1tag_', '_2tag_
                           '))
25
26                      if chain.nLepton==4 and
                           chain.deltaMassZ_ttZ<15.
                           and chain.deltaMassZ_ttZ
                           >=0. and ptmiss_ttZ>160.:
27                          crbin = get_CRbin(
                               ptmiss_ttZ, -1, '2016
                               X' in opt.year)
28
29                      if chain.nLepton==3 and
                           chain.deltaMassZ_WZ<15.
                           and chain.deltaMassZ_WZ
                           >=0. and ptmiss_WZ>=0.:
30                          ptxGhost = ptmiss_WZ*
                               math.cos(
                               ptmiss_phi_WZ)
31                          ptyGhost = ptmiss_WZ*
                               math.sin(
                               ptmiss_phi_WZ)
32                          for ilep in [ chain.
                               lep0idx_WZ, chain.
                               lep1idx_WZ, chain.
                               lep2idx_WZ ]:
33                              if (chain.
                                   Lepton_pdgId[ilep
                                   ]*chain.
                                   Lepton_pdgId[
                                   chain.lep2idx_WZ
                                   ])<0 or ilep==
                                   chain.lep2idx_WZ:
34                                  ptxGhost +=
                                       chain.
```

```
                                        Lepton_pt[
                                        ilep]*math.
                                        cos(chain.
                                        Lepton_phi[
                                        ilep])
35                                  ptyGhost +=
                                        chain.
                                        Lepton_pt[
                                        ilep]*math.
                                        sin(chain.
                                        Lepton_phi[
                                        ilep])
36                              ptmiss_ttZ3Lep = math.
                                    sqrt(ptxGhost*
                                    ptxGhost + ptyGhost*
                                    ptyGhost)
37                              crbin = get_CRbin(
                                    ptmiss_ttZ3Lep, -1, '
                                    2016X' in opt.year)
38
39                          if crbin>=0:
40                              coordinates_cr = numpy.
                                    float64([pmssid1,
                                    pmssid2, sr_nbins+12+
                                    crbin+1])
41                              thnsparse.Fill(
                                    coordinates_cr,
                                    weight *
                                    btagweight_ttZ)
```

Appendix B

Event Selection Summary

This appendix summarizes the event selection criteria applied in the analysis. The selections are divided into baseline requirements, kinematic cuts, and signal region (SR) definitions. These ensure consistency between the data and Monte Carlo samples, and define the regions used in the statistical analysis.

B.1 Baseline Selection

- Exactly two opposite-sign isolated leptons (electrons or muons).
- Leptons required to pass identification and isolation criteria.
- Events must pass the relevant dilepton triggers.

B.2 Kinematic Requirements

- Missing transverse momentum: $p_{\mathrm{T}}^{\mathrm{miss}} > 160$ GeV.
- Jet multiplicity: events categorized by $n_{\mathrm{jets}} = 0$ or $n_{\mathrm{jets}} > 0$.
- b-tagging: separate treatment of events with and without tagged jets.

B.3 Signal Region Definitions

The signal regions are defined based on $p_{\mathrm{T}}^{\mathrm{miss}}$ and $m_{T2}(\ell\ell)$:

- **SR1:** Binned in $m_{T2}(\ell\ell)$ from 0 to 100 GeV in intervals of 20 GeV, plus two bins with $100 \leq m_{T2}(\ell\ell) < 160$ GeV and $m_{T2}(\ell\ell) > 160$ GeV.
- **SR2:** Same as SR1, but splitting the last bin into $160 \leq m_{T2}(\ell\ell) < 240$ GeV and $m_{T2}(\ell\ell) > 240$ GeV.

- **SR3:** Same as SR2.
- **SR4:** Same as SR2 and SR3, but splitting the last bin into $240 \leq m_{T2}(\ell\ell) < 370$ GeV and $m_{T2}(\ell\ell) > 370$ GeV.

Each region is further split into same-flavor (SF: ee, $\mu\mu$) and different-flavor (DF: $e\mu$) categories, which correspond to even and odd `srbin` indices, respectively. The overall bin index is then combined with the pMSSM model identifiers in the `THnSparse` object.