



TECHNISCHE
UNIVERSITÄT
WIEN

ÖAW

ÖSTERREICHISCHE
AKADEMIE DER
WISSENSCHAFTEN

sdk Π

Doktoratskolleg
Particles and Interactions



INSTITUT FÜR HOCHENERGIEPHYSIK

DISSERTATION

Data-driven background modelling and trigger algorithms for compressed supersymmetry searches with the CMS experiment at the LHC

Ausgeführt zum Zwecke der Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

eingereicht an der Fakultät für Physik der Technischen Universität Wien

von

Mateusz Zarucki MSci.

Matrikelnummer 0927578

ausgeführt am Institut für Hochenergiephysik (HEPHY)
der Österreichischen Akademie der Wissenschaften (ÖAW)
und am Atominstitut der Technischen Universität Wien (E141)

unter der Leitung von

Priv.-Doz. Mag. phil. Mag. Dr. rer. nat. Manfred Jeitler

und

Dipl.-Ing. Dr. Ivan Mikulec

als verantwortlich mitwirkender Assistent

Unterstützt vom Österreichischen Wissenschaftsfonds (FWF) im Rahmen des Doktoratsprogramms
Particles and Interactions (W1252-N27).

Wien, am 24. April, 2023

In memory of my beloved father,

Ryszard Zarucki

Jestem z Ciebie
dumny

Abstract

Supersymmetry (SUSY) is a highly motivated theory that can provide solutions to central issues and open questions of the standard model (SM) of particle physics. However, the absence of discoveries of exotic particles at the Large Hadron Collider (LHC) puts it under significant pressure. Nonetheless, some corners of the parameter space remain less explored, owing more to their experimental difficulty rather than a lack of theoretical motivation. SUSY scenarios with a *compressed* mass spectrum, where the mass difference between the produced superpartners and the lightest supersymmetric particle (LSP) is relatively small, offering little detectable energy, are very well motivated by naturalness considerations and dark matter relic constraints.

A compressed SUSY search focusing on models with top squark (stop) pair-production in the single-lepton channel is presented. Due to the limited available energy and resulting low momentum decay products, such a search is made viable by requiring a boost from an initial-state radiation (ISR) jet. The analysis results are based on collision data from year 2016 of LHC *Run 2*, recorded with the Compact Muon Solenoid (CMS) detector at $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 35.9 fb^{-1} . Several data-driven background modelling methods are implemented for the estimation of sub-leading nonprompt lepton backgrounds, using Monte Carlo (MC) simulation together with normalisation to data. Dedicated studies are conducted for the identification of low momentum electrons as well as improvements in signal modelling.

Based on the completion of the analysis with 2016 data, another focus of the thesis are improvements of the trigger strategy to extend the acceptance of the signal, also towards less-accessible compressed *electroweakino* (EWKino) models. Dedicated trigger algorithms are developed, exploiting the typical ISR-boosted signature, in order to achieve lower missing momentum thresholds. The algorithms were used online during year 2018 of LHC *Run 2* and are currently collecting data during the ongoing LHC *Run 3*. A strategy for incorporating the new trigger in the analysis is presented, evaluating the potential expected gains in sensitivity. Possible future improvements in expanding the search capabilities in previously unexplored directions, including recent developments, such as machine learning (ML) or the identification of displaced signatures from long-lived particles (LLP), are also discussed.

Deutsche Kurzfassung

Die Supersymmetrie (SUSY) ist eine gut motivierte Theorie, die Lösungen für zentrale Probleme und offene Fragen des Standardmodells (SM) der Teilchenphysik bietet. Da jedoch bisher keine exotischen Teilchen am Großen Hadronen-Speicherring (LHC) entdeckt wurden, steht sie unter erheblichem Druck. Nichtsdestotrotz bleiben einige Bereiche des Parameterraums weniger erforscht, was eher auf hohe experimentelle Herausforderungen zurückzuführen ist als auf einen Mangel an theoretischer Motivation. SUSY-Szenarien mit einem komprimierten (*compressed*) Massenspektrum, bei denen der Massenunterschied zwischen den erzeugten Superpartnern und dem leichtesten supersymmetrischen Teilchen (LSP) relativ klein ist, und daher nur wenig detektierbare Energie bietet, sind sehr gut durch Natürlichkeitsüberlegungen und Beschränkungen der gemessenen Dichte der dunklen Materie motiviert.

In dieser Arbeit wird eine Suche nach *compressed* SUSY vorgestellt, die sich auf Modelle mit Top-Squark (Stop) Paarproduktion im Einzelleptonkanal konzentriert. In diesem Modell haben die Zerfallsprodukte nur einen geringen Impuls. Eine solche Suche wird allerdings durch einen Boost der SUSY Teilchen mittels eines abgestrahlten Jets im Anfangszustand (ISR) ermöglicht. Die Analyseergebnisse basieren auf Kollisionsdaten aus dem Jahr 2016 des LHC *Run 2*, die mit dem Compact Muon Solenoid (CMS) Detektor bei $\sqrt{s} = 13$ TeV aufgenommen wurden. Die Daten entsprechen einer integrierten Luminosität von 35.9 fb^{-1} . Mehrere Methoden zur Abschätzung der Hintergrundprozesse basierend auf Monte Carlo (MC) Simulationen und experimentellen Daten werden implementiert. Spezielle Studien zur Verbesserung der Identifikation von niederimpulsigen Elektronen sowie der Modellierung der Signale werden durchgeführt.

Nach Abschluss der Analyse mit Daten aus dem Jahr 2016 liegt ein weiterer Schwerpunkt der Arbeit auf der Verbesserung der Triggerstrategie, um die Signalakzeptanz auch auf weniger zugängliche komprimierte *electroweakinos* (EWKinos) auszuweiten. Es wurden spezielle Trigger-Algorithmen entwickelt, die die typische ISR-geboostete Signatur ausnutzen, um niedrigere Schwellenwerte für fehlenden Impuls zu erreichen. Die Algorithmen wurden im Jahr 2018 in *Run 2* online eingesetzt und sammeln derzeit Daten im laufenden *Run 3*. Es wird eine Strategie für die Einbeziehung des neuen Triggers in die Analyse vorgestellt und der potenzielle erwartete Gewinn an Sensitivität bewertet. Mögliche zukünftige Verbesserungen werden ebenfalls erörtert, einschließlich aktuelle Entwicklungen, wie maschinelles Lernen (ML) oder die Identifizierung von langlebigen Teilchen (LLP) mit versetztem Zerfallsvertex, um die Suchmöglichkeiten in bisher unerforschte Richtungen zu erweitern.

Contents

1	Introduction	1
I	Theoretical and Experimental Introduction	5
2	The Standard Model and Beyond	7
2.1	The Standard Model	7
2.1.1	Theoretical Introduction	10
2.1.2	Measurement	18
2.1.3	Shortcomings	21
2.2	Supersymmetry	25
2.2.1	Minimal Supersymmetric Standard Model	25
2.2.2	Light Stop Window and Compressed SUSY	34
3	Experiment	39
3.1	CERN	39
3.1.1	Accelerator Complex	39
3.1.2	Large Hadron Collider	41
3.2	CMS Experiment	47
3.2.1	Scientific Programme	47
3.3	CMS Detector	48
3.3.1	Geometry	49
3.3.2	Magnet	52
3.3.3	Tracker	53
3.3.4	Calorimeters	56

3.3.5	Muon Detectors	64
3.3.6	Trigger and Data Acquisition	70

4	Event Reconstruction	91
4.1	Global Event Description	91
4.2	Object Definitions	107
4.3	Simulation	114

II Compressed Stop Search **119**

5	Models and Signature	121
5.1	SUSY Searches	121
5.1.1	Simplified Signal Models	121
5.1.2	Stop Production and Cross Sections	122
5.2	Compressed SUSY Signature	131

6	Event Selection	133
6.1	Analysis Timeline	133
6.2	Analysis Framework	134
6.3	Data and Simulated Samples	134
6.3.1	Signal Generation	135
6.3.2	Simulation Corrections	140
6.4	Event Selection	144
6.4.1	Cut & Count Analysis	145
6.4.2	Trigger	146
6.4.3	Preselection and Backgrounds	148
6.4.4	Signal Regions	153
6.4.5	Summary	157

7	Background Estimation Methods	159
7.1	Prompt	159
7.2	Nonprompt	163
7.2.1	Emulation of Z-Boson Decays to Neutrinos	164

7.2.2	QCD Estimation with ABCD Method	171
7.2.3	Tight-to-Loose Method	176
8	Results	183
8.1	Systematics	183
8.2	Results	190
III	Soft Trigger	195
9	Additional Compressed Models	197
10	Soft Trigger Algorithms	201
10.1	Introduction	201
10.2	Trigger Design	201
10.2.1	Signal Acceptance and Background Rejection	202
10.2.2	Rates	205
10.2.3	Efficiency	207
10.2.4	Other Considerations	209
10.3	Trigger Development	211
10.3.1	Cross-Trigger	211
10.3.2	Rate Estimation	212
10.3.3	LIT Seed and HLT Path Development	213
10.3.4	Efficiency in Simulation	217
10.3.5	Final Trigger	219
10.4	Performance	221
10.4.1	Online Rate	221
10.4.2	Timing	221
10.4.3	Efficiency in Data	223
10.5	Sensitivity	226
10.5.1	Low-MET Region	226
10.5.2	Expected Limits	238
10.5.3	Soft Multileptons	243

II Outlook	245
II.1 Final Remarks	259
Acknowledgements	265
A Appendix	271
A Feynman Diagrams	271
A.1 Charm Decay	271
A.2 Stop Production	272
B Analysis	273
B.1 Samples	273
B.2 MVA Analysis Limits	274
C Soft Trigger	275
C.1 Signal Distributions	275
C.2 LIT Seed Efficiency in Simulation	276
C.3 Low-MET Regions	277
C.4 Sensitivity	279
C.5 Trigger Tools	280
D Status Quo of Compressed SUSY Searches	287

Introduction

The standard model (SM) of particle physics has seen remarkable success throughout the modern scientific era. It has been verified with extreme precision at different energy scales, becoming the most successful theory of the fundamental interactions of elementary particles. Collider experiments, such as those at the Large Hadron Collider (LHC), have played a significant role in measurements and discoveries that contributed to the building of the SM. The discovery of the Higgs boson at the LHC by the CMS [1] and ATLAS [2] experiments is considered one of the major missing puzzle pieces of the model.

Despite the success of the SM, there are several shortcomings that need to be accounted for, especially in the context of a grand unified theory (GUT) or even a theory of everything (TOE). Ultimately, this implies that there exists new physics beyond the SM (BSM) that has not yet been discovered. Supersymmetry (SUSY) is one of the most promising theories that could provide solutions to central open questions and problems in the SM. SUSY postulates that for every SM particle, there exists a corresponding superpartner state, differing by half-integer spin. The additional particle content predicted by SUSY would provide a solution to the hierarchy or naturalness problem, while minimal supersymmetric extensions of the SM (MSSM) could allow for the unification of strong and electroweak gauge couplings at the GUT scale. Furthermore, the lightest supersymmetric particle (LSP), which is weakly-interacting, massive and stable, is an excellent dark matter (DM) candidate. A focus of the physics programme

at the LHC is BSM physics, which includes searches for SUSY. Many SUSY models require R-parity conservation, meaning that the SUSY particles are pair-produced and are predicted to have decay chains to SM particles, which can then be detected by the multi-purpose detectors.

This thesis focuses on a particular SUSY model with a highly *compressed* mass spectrum, where the mass difference between the produced superpartners and the LSP is relatively small. Such scenarios are additionally motivated by DM relic constraints and can still provide a window to natural SUSY. This is especially relevant, considering the significant pressure that natural SUSY is under due to the lack of discoveries. There exists an underlying difficulty with identifying the resulting decay products due to the little available energy, making it a very challenging region of SUSY parameter space to probe. The *soft single-lepton analysis* described in this thesis focuses on top squark (stop) pair-production with a nearly degenerate neutralino $\tilde{\chi}_1^0$ LSP in the single-lepton channel.

As the author of this thesis, significant contributions were made to the published LHC *Run 2* analysis with 35.9 fb^{-1} [3, 4, 5], as well as a previous preliminary result with early data (12.9 fb^{-1}) [6]. One of the main contributions to the publication is the estimation of sub-leading nonprompt backgrounds using several data-driven background modelling methods (emulation of Z-boson decays to neutrinos, ABCD method for QCD multi-jet, tight-to-loose method) using Monte Carlo (MC) simulation together with normalisation to data. Furthermore, dedicated studies were performed for the identification of low momentum electrons for the addition of the electron channel, which was one of the major additions with respect to the *Run 1* version of the analysis [7]. Improvements in signal modelling have been studied and implemented in the form of the development of a dedicated generator filter, making mass SUSY signal sample production viable for compressed models. A number of other contributions were made, such as in the determination of lepton scale factors or the evaluation of various

systematic uncertainties relevant to the analysis (nonprompt and prompt estimation, W-polarisation, pileup, ...). The analysis was presented at a number of global conferences, followed by corresponding published proceedings [8, 9, 10].

Based on the completion of the analysis with 2016 data, another focus of the thesis are improvements of the trigger strategy to extend the acceptance of the signal, also towards less-accessible compressed *electroweakino* (EWKino) models. Dedicated trigger algorithms were developed, exploiting the typical ISR-boosted signature, in order to achieve lower missing momentum thresholds. The algorithms were used online during year 2018 of LHC *Run 2* and are currently collecting data during the ongoing LHC *Run 3*. The performance measurements of the trigger in 2018 data were made public under a detector performance note [11]. A strategy for incorporating the new trigger in the analysis is presented, evaluating the potential expected gains in sensitivity.

In parallel, I have carried out additional work for the CMS experiment on various fronts. I supported online data-taking operations regularly by being a data acquisition (DAQ) shifter responsible for controlling the DAQ system in the CMS control room and monitoring the collection of data. I also developed the software for the data quality monitoring (DQM) of the LIT Global Trigger Upgrade (μGT). This was followed by the Level-1 Trigger (L1) menu co-coordinator role (L3) responsible for delivering the full trigger menu that pre-selects the type of data that is collected, encapsulating the entire physics programme of CMS. Currently, I am the co-convenor (L2) of the Field Operations Group (FOG) of the Trigger Studies Group (TSG), responsible for CMS data-taking operations of the High Level Trigger (HLT) system during the LHC *Long Shutdown 2* (LS2) and the ongoing LHC *Run 3*. This role additionally includes coordinating the development of the trigger Rate Monitoring (RateMon) software.

Part I

Theoretical and Experimental Introduction

The Standard Model and Beyond

2.1 The Standard Model

Throughout the last century, global efforts in the theoretical and experimental fields of particle physics focused on attempting to construct a unified theoretical model for the description of our universe. The standard model (SM) of particle physics is currently our best description of the fundamental building blocks of the universe and their underlying interactions, combining the contrasting fields of quantum mechanics (QM) and relativity. Its predictions are verified with extreme precision up to the highest achievable energies, yielding it as the most successful theory of the fundamental interactions of elementary particles [12, 13, 14, 15].

Short Historical Context ¹

Even though the notion that all matter constitutes fundamental and indivisible particles has been considered since ancient times, Thomson's discovery of the electron [16, 17] demonstrated the existence of particles smaller than atoms, marking the beginning of a modern era of particle physics. Rutherford's discovery of the proton within

¹This section provides a very brief historical context, only covering a select number of experimental discoveries and theoretical developments that were vital in the construction of the standard model. A number of equally important experimental discoveries and measurements, as well as theoretical frameworks, are omitted in the interest of brevity, however, their omission should not undermine their importance.

a concentrated nucleus [18], along with Chadwick's discovery of the neutron [19], highlight the so-called *classical* era of particle physics, where the primary constituents of an atom were defined.

The birth of the field of quantum mechanics (QM) coincides with the discovery of the particle-like nature of the photon at the beginning of the 20th century [20, 21, 22]. The nonintuitive quantum-mechanical effects manifest themselves at the smallest physical scales, in which non-relativistic dynamical particle systems are described as quantised systems by the Schrödinger Equation. The relativistic quantum mechanical (RQM) formulation of QM extends the theory with Einstein's special relativity, significant at very high energies, yielding the relativistic Klein-Gordon [23, 24] and Dirac wave equations [25, 26]. The negative-energy solutions of the Dirac equation postulate the existence of antiparticle partners [27], which were subsequently discovered [28, 29, 30]. This led to the description of elementary particles as the excitations of quantised relativistic fields in the context of a quantum field theory (QFT). Quantum electrodynamics (QED) [31], advanced by Tomonaga, Schwinger and Feynman, is a QFT that describes the quantisation of electromagnetic fields, expressed by the well-established Maxwell's equations [32], where the quantised photon acts as a mediator of the electromagnetic force between charged fermions.

Yukawa's Theory postulated a short-range force within nuclei that manifests itself as the exchange of a meson (pion) with a mass intermediate between that of a nucleon and an electron, which was the first indication of a strong force that acts between hadrons [33]. Experimental confirmation of the existence of strongly-interacting particles was achieved in studies of cosmic rays using cloud chambers, which discovered that secondary cosmic rays are predominantly composed of pion mesons decaying into muon leptons [34, 35, 36, 37]. The positron was the first antiparticle to be discovered via cosmic ray observations [28], followed by later discoveries of the antiproton [29] and antineutron [30] in synchrotron fixed-target experiments. Cosmic ray measurements were also crucial in the discovery of neutral Kaon decays to pions, forming characteristic V-shaped particle tracks [38, 39], as well as decays of charged K-mesons

[40]. A similar signature is left by the decay of the Lambda into a proton and pion [41], which was classified into the heavier baryon family due to its mass. In parallel, the discovery of the neutrino (postulated much earlier by Pauli [42]) in inverse beta-decay experiments [43] confirmed vital properties of QM of spin and lepton number conservation.

Starting from the 1950s, the development of accelerator experiments enabled the production of particles in a controlled laboratory environment, ranging from Brookhaven [44], SLAC [45] and Tevatron [46] to the highest energy CERN accelerators (e.g. SPS [47], LEP [48], LHC [49, 50]). They have since played a central role in verifying theoretical models and predictions of the SM, with a number of historical discoveries of elementary particles, such as the J/ψ [51, 52], tau lepton [53], gluon [54, 55], top quark [56, 57], and the W, Z [58, 59, 60, 61] and Higgs bosons [1, 2].

With the increasing number of particle discoveries, there arose the necessity to classify them based on the underlying physics, such as the classification of hadrons into the meson and baryon families. The law of conservation of the baryon number [62, 63, 64] was already introduced earlier to account for the stability of the proton and was applicable to the newly-discovered particles. The relatively longer lifetimes of some newly-discovered hadrons led to the introduction of the property of *strangeness* [65, 66, 67]. The eightfold way [68, 69] classifies the particles based on geometrical combinations of strangeness and charge, yielding the baryon and meson octets and baryon decuplet. The baryon decuplet predicted the existence of the Ω^- , discovered at Brookhaven [70], cementing the model's success.

The quark model [71, 72, 73] describes every hadron as a colourless bound state of spin-half fermionic quarks, where baryons consist of three quarks and mesons of quark-antiquark pairs. Experimental evidence of quarks was revealed by observing corresponding partons in deep inelastic scattering (DIS) experiments [74, 75], which probed the internal structure of hadrons by high energy scattering of electrons off nuclei. The joint discoveries of the J/ψ meson at Brookhaven and SLAC [51, 52] and the heavy top quark by the CDF and $D\bar{0}$ experiments at Fermilab [56, 57] essentially confirmed the

mathematical model of quarks. The electroweak unification [76, 77, 78], advanced by Glashow, Weinberg and Salam, took a significant step towards a unified theory of all interactions. Weak neutral current interactions predicted by the theory were discovered in neutrino scattering in the Gargamelle bubble chamber [79], originating from the Proton Synchrotron (PS) [80, 81] beams at CERN. Furthermore, it allowed the prediction of the masses of the W and Z bosons, which were later discovered by the UA1/2 experiments [58, 59, 60, 61] at the larger Super Proton Synchrotron (SPS) accelerator [47]. The discovery of the Higgs boson at the LHC by the CMS [1] and ATLAS [2] experiments in 2012 is considered one of the major missing puzzle pieces of the SM.

2.1.1 Theoretical Introduction

Particle Content

The matter content of the standard model (SM) of particle physics shown in Figure 2.1 can be summarised as comprising three types of elementary particles: leptons, quarks and mediators. For each particle, there exists a corresponding antiparticle with the same mass, but quantum numbers reversed. There are six flavours of both leptons and quarks, subdivided into three generations, which differ mainly in mass while keeping other properties the same. Lepton generations consist of flavour pairs of charged leptons (electrons, muons or taus) and uncharged neutrinos. Quarks have fractional electromagnetic charge, and generations consist of pairs which differ in rest mass and one unit electromagnetic charge. In the context of QCD, they are also attributed to three different types of colour charges. The phenomenon of colour confinement [82] requires that coloured particles cannot be observed as isolated particles. Therefore, they form colourless bound mass states of two quark-antiquark pairs ($q\bar{q}$), known as mesons (bosons), or three distinctly coloured quarks (qqq), known as baryons (fermions). Leptons and quarks are half-integer spin ($\text{spin}^{-1/2}$) fermions and are, therefore, subject to the Pauli exclusion principle (PEP) [83], which forbids that two indistinguishable particles are in the same quantum state.

The fundamental forces are transmitted via quantised mediators that are integer spin bosons (spin-1 vector bosons), which are not subject to the PEP. The electromagnetic force is mediated via massless photons, the strong force via eight coloured gluons and the weak force via massive weak intermediate vector bosons: charged W^\pm or neutral Z bosons. The massive Higgs boson is a spin-0 particle responsible for the generation of mass. The SM does not consider the gravitational force, described by general relativity, due to the complexity of a unified quantum gravity theory (postulating a spin-2 graviton) and is neglected due to its much weaker strength at the microscopic scale.

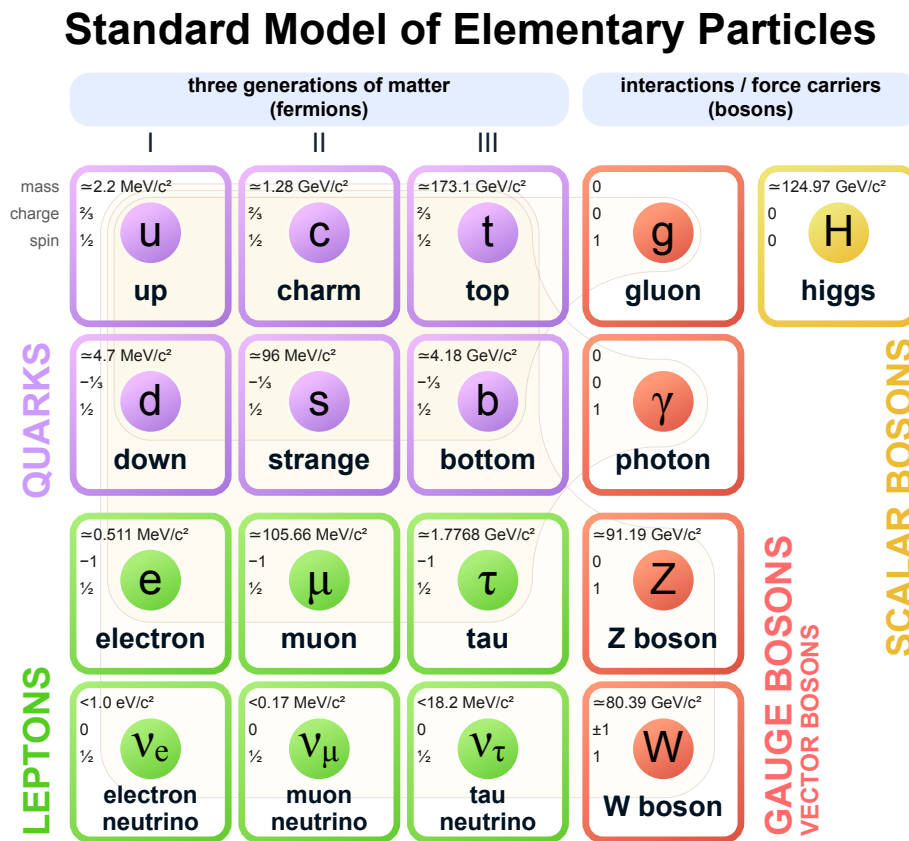


Figure 2.1: Particle content of the standard model (SM) of particle physics in terms of fundamental leptons, quarks and mediators [84].

Quantum Field, Lagrangian, Gauge Theories

Special relativity and quantum mechanics are consolidated within the Dirac equation, which is a relativistic wave equation for quantum-mechanical systems. Quantum field theory (QFT) describes elementary particles as excitations of quantised relativistic fields, allowing an arbitrary and changeable number of particles to exist. The SM is a QFT where the fundamental forces between elementary particles are expressed as quantised mediators.

In Lagrangian field theory, such systems of fields can be described in terms of Lagrangian densities, where the dynamics are dictated by the minimisation of *action* ($\mathcal{S} = \int \mathcal{L} dt$). Elementary matter particles described by the Dirac equation are expressed as Dirac fields, which interact with each other via the fundamental force fields. The influence of both fields on one another can be expressed by a combined Lagrangian density and can account for nearly all observed phenomena on different scales.

The SM Lagrangian uniquely describes the particle content and underlying interactions in the SM, and it is generally defined as:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{Yukawa}} + \mathcal{L}_{\text{Higgs}} \quad (2.1)$$

in terms of gauge, fermion, Yukawa couplings, as well as the Higgs interaction:

- $\mathcal{L}_{\text{gauge}}$ describes the gauge part of the Lagrangian density, which includes the kinetic terms of the different bosonic gauge fields.
- $\mathcal{L}_{\text{fermion}}$ comprises the dynamic fermionic terms that include interactions of fermions and gauge bosons as well as fermion-fermion interactions.
- $\mathcal{L}_{\text{Higgs}}$ describes the dynamic Higgs term, the Higgs potential, mass terms of the gauge bosons and interactions of gauge bosons with the Higgs field.
- $\mathcal{L}_{\text{Yukawa}}$ includes the fermionic mass terms and interactions of fermions with the Higgs field.

Ultimately, the exact form of interactions in the SM is exposed by the underlying redundancies in our mathematical description of nature. Nöther's theorem postulates that every symmetry in nature is associated with a conservation law [85]. The Lagrangian formulation manifests the underlying global symmetries of a system, which correspond to conserved quantities or currents.

The mathematical formalism for interactions is expressed in the context of group field theory, where group elements are represented by (special) unitary matrices that are defined by infinitesimal generator operators [86]. Symmetries of space-time are described by the Poincaré group, including Lorentz transformations and space-time translations. Gauge transformations transform the fields in such a way that the dynamics of the system are unchanged. According to the Coleman-Mandula theorem [87], the Poincaré group can be extended by a gauge group as a direct product, and can only combine in a trivial way. Imposing the requirement of invariance of the Lagrangian under local phase transformations (LGI) leads to Lie symmetry groups with a corresponding Lie algebra of group generators. These ultimately expose the mathematical symmetries and redundant degrees of freedom of the Lagrangian density of a physical system (gauge symmetries) and thereby determine the exact form of all interactions. Each generator of the symmetry group introduces a gauge field, which then can be associated with a vector boson field.

Non-commutative symmetry groups, such as $SU(3)$ of QCD and $SU(2)$ of weak interactions, are referred to as non-abelian. Non-abelian gauge theories are described by the Yang-Mills theorem [88], which describes interactions among vector fields. The gauge potentials are matrices that also transform as part of gauge transformations, leading to gauge fields that self-interact, corresponding to coloured gluons and weak vector bosons. Ultimately, QCD is characterised by an $SU(3)$ gauge symmetry, whereas electroweak (EWT) theory is expressed via a combined $SU(2) \times U(1)$ symmetry.

To summarise, the mathematical description of the SM is based on a QFT with the requirement of LGI, characterised by a non-abelian $SU(3) \times SU(2) \times U(1)$ gauge symmetry group. The number of independent generators in the group corresponds to the

number of gauge mediator fields, with the gauge bosons as the corresponding quanta. Therefore, in total, there are 12 gauge bosons: the photon, 3 weak bosons and 8 gluons, as indicated in Figure 2.1. The explicit form of the SM Lagrangian is:

$$\begin{aligned} \mathcal{L}_{\text{SM}} = & -\frac{1}{2}F_{\mu\nu}F^{\mu\nu} + \bar{\Psi}_L i\gamma^\mu D_\mu \Psi_L + (D_\mu \Phi)^\dagger D^\mu \Phi + \mu^2 \Phi^\dagger \Phi \\ & -\frac{1}{2}\lambda (\Phi^\dagger \Phi)^2 + \left(\frac{1}{2} \Psi_L^T C h \Phi \Psi_L + h.c. \right) \end{aligned} \quad (2.2)$$

where $F_{\mu\nu}$ is the field strength tensor, Ψ_L^T represents the fermion fields (with left-handed spinors), D_μ is the covariant derivative (gauge group acting on the fermions), and h is the matrix of Yukawa couplings. When the Higgs field acquires a VEV via the EBH mechanism, additional mass terms are generated in the form:

$$\frac{1}{2} \Psi_L^T C h \Phi \Psi_L = h_{uij} \bar{u}_{Ri} q_{Lj} \Phi + h_{dij} \bar{d}_{Ri} q_{Lj} \tilde{\Phi} + h_{eij} \bar{e}_{Ri} \ell_{Lj} \tilde{\Phi} + h_{nij} \bar{n}_{Ri} \ell_{Lj} \Phi \quad (2.3)$$

where $i, j = 1, 2, 3$ label the families. The first two terms are mass terms for the up-type and down-type quarks, respectively, and cannot be diagonalised simultaneously, giving rise to the flavour-violating CKM matrix. The last two terms reflect the mass generation of leptons and neutrino mixing.

Quantum Electrodynamics

In quantum electrodynamics (QED), the quantisation of the electromagnetic field, described by Maxwell's equations, results in a gauge boson mediator that is the photon γ . It is an abelian gauge theory, and the interaction exhibits a phase symmetry described by the symmetry group $U(1)$. Charge conservation is linked to the invariance of the electromagnetic four-potential under local phase transformations. All charged fermions (quarks and charged leptons) interact electromagnetically, with a coupling strength of $\alpha = e^2/4\pi\eta_0\hbar c \approx 1/137$ at low energy scales.

Quantum Chromodynamics

Quantum chromodynamics (QCD) is a non-abelian gauge theory and is characterised by an $SU(3)$ gauge symmetry relating to three colour charges (r, b, g). The attribution of 3 colours to the quarks, requiring bound states to be colourless, solved the apparent violation of the PEP, attributing a $SU(3)$ symmetry. The strong interaction is mediated by massless gluons that carry a unit of colour and anticolour charge. There are eight independent combinations of pure colour-anticolour states, corresponding to the number of interacting gluons.

The strong force is transmitted between coloured particles, namely, quarks and gluons. As described in Section 2.1.2, the strength of the interaction, described by the running coupling 'constant' α_s (≈ 1 at 1 GeV), decreases with increasing energy scale. Therefore, the interaction is strong at low energies, leading to colour confinement [82], whereas at very high energies, quarks behave as if outside of bound states described by the phenomenon of asymptotic freedom [89, 90]. Furthermore, this behaviour is also exhibited in a short range of the strong force, as manifested within the Yukawa interaction.

Quantum Flavourdynamics

Flavourdynamics describes weak interactions that are possible between all fundamental fermions, within a $SU(2)$ representation. They are mediated by the massive charged W-bosons (charged current interactions) and neutral Z-bosons (neutral current interactions), with masses of approximately 80.4 GeV and 91.2 GeV, respectively. The lifetime of the massive mediators is very short, leading to a very short range of the force. The relative strength of the force is expressed in terms of the coupling constants g_W and g_Z , respectively, which are of the order of 10^{-6} and thus much weaker than both the electromagnetic and strong couplings.

Weak couplings are dependent on the helicity of the fermion, which is defined as the projection of the spin in the momentum direction. Chirality (or *handedness*) is the Lorentz-invariant quantity, which corresponds to helicity for massless particles that travel at the speed of light. Left and right chiral states couple differently to gauge bosons, where W -bosons couple with left-handed particles only. Since neutrinos are considered massless in the SM, all neutrinos are left-handed, and all antineutrinos are right-handed [91], which ultimately leads to parity and CP-Violation [92, 93]. This fundamental difference with respect to the electromagnetic and strong forces is the fact that flavour is not conserved for weak interactions, and the particle nature can change during interactions. The charged currents change the flavour of particles, as one unit of charge is carried away by the mediator, and the mixing between the quark generations is quantified by the Cabibbo–Kobayashi–Maskawa (CKM) 3×3 unitary matrix [94, 93]. The CKM matrix and its elements are fundamental flavour-violating (and CP-violating) parameters of the SM. Ultimately, it reflects the mixing angles and phases of the quark sector. The angles are small, so the CKM matrix is almost diagonal, and the level of flavour-violation is relatively small.

Electroweak Theory

The inconsistencies in a $SU(2)$ representation of the weak interactions, such as the existence of massive gauge bosons and different strengths of weak charges, are resolved with the electroweak (EWK) unification theory (EWT). The Glashow-Weinberg-Salam (GWS) model [76, 77, 78] unifies the electromagnetic and weak forces and incorporates the mass of the vector bosons. They are expressed as mixtures of three EWK gauge fields of weak isospin, W_3 , which couples to left-handed (L) fermions only, and hypercharge Y , which couples to both left- and right-handed particles. Therefore, the physical W and Z bosons are mass eigenstates formed from a mixture of the EWK gauge fields (W_3 and Y). The level of mixing is defined by the weak mixing angle θ_W , which is a fundamental parameter of the SM, where $\cos \theta_W = m_W/m_Z$. Similarly, the photon γ

can also be expressed as a mixture. This results in a $SU(2)_L \times U(1)_Y$ symmetry, where $SU(2)_L$ is the isospin symmetry I_3 of the W_3 , coupling to left-handed (L) particles, and $U(1)_Y$ refers to the hypercharge Y .

The EWK sector is completed by including a Higgs interaction responsible for the mass generation of vector bosons and fermions, which is explained by the EBH mechanism² [95, 96, 97]. The Higgs field is a scalar field with two neutral and two electrically charged components that form a complex doublet of the weak isospin $SU(2)_L$ symmetry. The vacuum expectation value (VEV) of the Higgs potential is non-zero (246 GeV) at the minimum of the Higgs potential. For particles to acquire mass, the global weak isospin symmetry is spontaneously broken (SSB). The EBH mechanism explains how the massless Nambu-Goldstone bosons [98, 99], generated by SSB, are absorbed by the weak vector bosons to acquire their masses via the mechanism of electroweak symmetry breaking (EWSB). The neutral component couples to fermions, and its coupling strength is proportional to the mass of the coupled quarks and leptons. The Higgs-fermion coupling indicates that mass eigenstates do not correspond to the weak flavour eigenstates, and there is mixing between the three generations, as indicated by the CKM matrix. Following EWSB, the symmetry is reduced to the electromagnetic $U(1)$, and the remaining electrically neutral component manifests itself as an uncharged Higgs boson H , with zero spin and a mass of approximately 125 GeV.

²Or more accurately, the London-Anderson-Englert-Brout-Higgs-Guralnik-Hagen-Kibble-Weinberg mechanism.

2.1.2 Measurement

Ultimately, the SM has 19 free parameters, which need to be measured experimentally: elementary particle masses, gauge coupling strengths, CKM mixing angles and CP-violation phase, and finally, the Higgs VEV. These measurements are performed by a number of global experiments, especially those at particle accelerators, such as the LHC.

Conservation Laws

Processes can be classified between decays and scattering processes, which are mediated by the fundamental forces. Massive particles spontaneously decay into lighter states, with a given probability depending on the properties of the fundamental interaction. Relativistic scattering processes involve fundamental interactions between particles and can be induced in laboratory conditions, such as in particle colliders like the LHC.

The totalitarian principle postulates that any process that is not forbidden must occur. Therefore, any process that is possible but not observed must be forbidden by conservation laws or highly suppressed below our experimental reach. There exist absolute conservation laws, such as the classical kinematic conservation of energy and (angular) momentum, including intrinsic spin, which are connected to translations (rotations) in time and space. Dynamical conservation laws include the conservation of baryon number (B) and lepton number (L) in all relevant interactions. The conservation of electric charge, colour charge and weak isospin is related to the gauge invariance of the corresponding symmetry Lie groups.

The weak interaction violates some of the symmetries that are present in the strong and electromagnetic interactions. Flavour (strangeness) is conserved in strong and electromagnetic but not in weak interactions. Discrete symmetries such as charge-conjugation (C), parity (P) or time-reversal (T) [100] and their combinations relate to quantum numbers that are conserved in certain interactions. Electromagnetic and

strong interactions respect parity and charge conjugation symmetries, whereas weak interactions do not, including the combined CP-violation [92, 93], which can give insights into the matter-antimatter asymmetry in the universe. W -bosons couple with left-handed particles only and since neutrinos are considered massless in the SM, all neutrinos are left-handed, and all antineutrinos are right-handed.

Measurable quantities include interaction rates, such as decay rates or widths, or (differential) cross sections, quantifying the probability of a scattering or absorption event. The interaction rates depend on the particle energies, the phase-space and probability amplitudes. The quantum-mechanical probability amplitudes of interactions are expressed as matrix elements (ME), which depend on the Lagrangian density, determined by the nature of the particles, underlying interaction and phase-space. In scattering processes, the probability amplitudes are described by the scattering matrix (S -matrix). Conservation of probability is reflected in the unitarity requirement of the S -matrix, which is a central feature of QFTs. ME calculations result in non-linear systems that typically cannot be solved analytically. However, they can be solved by perturbative techniques, such as Feynman diagrams (e.g. Figure 6.8), where each diagram represents a term in a multiplicative perturbation series for relativistic field theories. Feynman vertices depend on the Lagrangian interaction term, with the interaction's coupling constant forming part of the vertex factor. Internal elements of Feynman diagrams are referred to as *virtual* particles, which are *off-shell* and not subject to relativistic energy-momentum conservation.

Renormalisation Group Evolution

The actual coupling strength of the forces is dictated by the summation of all the possible Feynman diagrams from perturbation theory, which include higher-order radiative quantum corrections, such as loop diagrams (e.g. Figure 2.4). An integration over all loop momenta of virtual particles leads to infrared and ultraviolet divergences. Regularisation of the integral treats infrared divergences and introduces an ultraviolet

cutoff Λ_{UV} , which marks the validity of a given QFT. It involves the cancellation of divergences in the limit $\Lambda_{UV} \rightarrow \infty$, often chosen at the level of the very large Planck scale $m_P \approx 10^{19}$ GeV, where quantum gravity effects become significant for proper treatment.

Expressions for physical and observable quantities, such as masses and coupling constants, that are independent of Λ_{UV} , are provided with the process of renormalisation. It redefines the parameters of a regulated theory at a given renormalisation scale and enforces mass and wavefunction normalisation. The renormalisation group equations (RGE) evolve from a high energy scale in order to be able to describe physically meaningful quantities near the electroweak scale. This effectively leads to the so-called *running* of coupling constants as a function of the energy scale.

The evolution of the gauge coupling strengths is shown in Figure 2.8. For the electromagnetic and weak forces, particles at low energy scales are essentially subject to a screening effect from a sea of virtual particles, and the coupling strength increases with energy, which unify in strength at the electroweak scale, as predicted by EWT. In the context of the QCD, the strong coupling α_S has a large dependence on energy due to self-interactions spanning from its non-abelian nature: the coupling strength decreases with increasing energy, providing an explanation for asymptotic freedom and quark confinement. Ultimately, the SM, which is a renormalisable QFT, can be considered a relatively low-energy approximation, allowing precise predictions to be made at energy scales within reach of our experiments.

Cross Section Measurements

The predictions of the SM have been verified with extreme precision at different energy scales by a number of global experiments, underlining its remarkable success. Collider experiments, such as those at the LHC, have played a significant role in testing the SM. Precision measurements of the cross sections for significant SM processes, including associated production of the recently-discovered Higgs boson, performed by the CMS

experiment are shown in Figure 2.2; an agreement over almost ten orders of magnitude is observed.

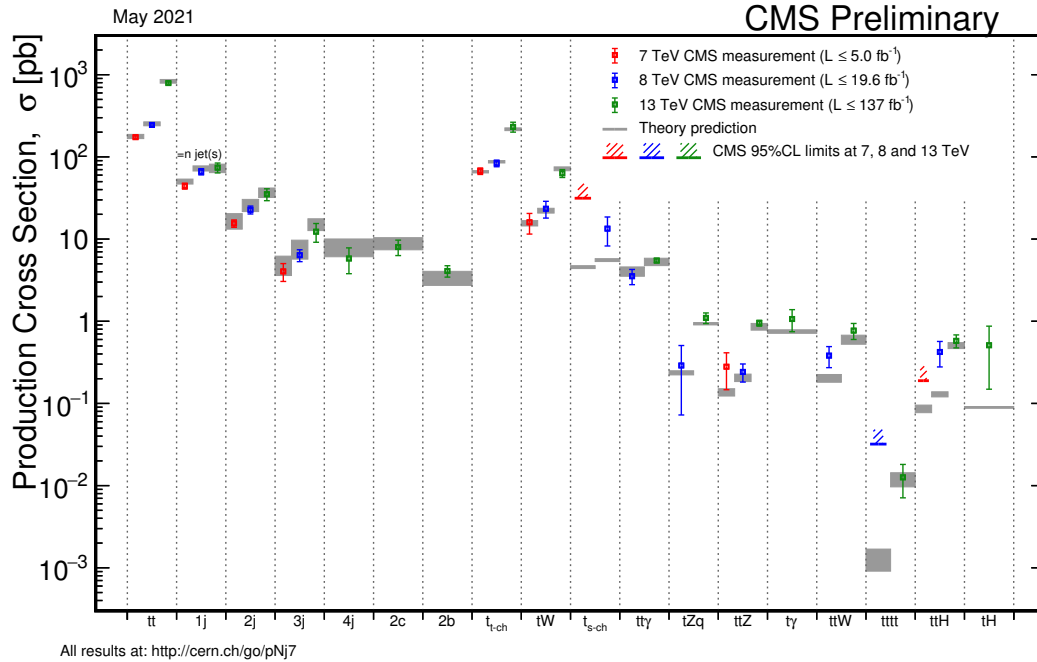


Figure 2.2: Precision measurements of cross sections for various SM processes performed by CMS at $\sqrt{s} = 7, 8$ and 13 TeV [101].

2.1.3 Shortcomings

Despite the success of the SM, whose properties have been experimentally verified with extreme precision over the past decades, there are several shortcomings that need to be accounted for, especially in the context of a grand unified theory (GUT) or even a theory of everything (TOE). Ultimately, this implies that there exists new physics beyond the SM (BSM) that has not yet been discovered.

Neutrino Masses

Oscillations between neutrino flavours are postulated [102, 103, 104] as a solution to the solar neutrino problem [105]. For oscillations to be possible, it requires neutrinos to have mass, which contradicts the massless assumption of the SM. Neutrino oscillations have been confirmed experimentally [106, 107], confirming their massive nature,

with upper limits in the sub-eV range [108]. The flavour mixing of neutrinos is characterised by the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix [109], which can be incorporated into the SM as a lepton-equivalent of the CKM matrix.

If neutrino masses are generated after EWSB, then the Yukawa-like couplings must be very small. Therefore, neutrino mass-generation mechanisms other than coupling to the Higgs field are considered. A prominent model is the see-saw mechanism, which postulates that massive neutrinos have a Majorana nature and are their own antiparticle, which involves right-handed neutrinos. The Majorana nature of neutrinos, together with CP-violation, could be used to explain the initial asymmetry between matter and antimatter in our universe. Evidence of this is being sought for in various experiments, such as neutrinoless double-beta decay experiments (e.g. [110]), which could additionally provide an indirect measurement of the neutrino mass by measuring the lifetime of the decay.

Dark Matter

Astrophysical observations provide strong evidence for the existence of dark matter (DM), a non-luminous weakly-interacting form of matter. These include rotation curves of galaxies [111, 112], gravitational microlensing effects or latest observations of the bullet cluster [113, 114], claiming direct empirical proof of the existence of dark matter. The latest measurements of the cosmic microwave background (CMB) anisotropies from the space telescope Planck provides further proof [115]. In the context of the lambda cold dark matter (Λ CDM) model of cosmology and Hubble constant measurements, it is estimated that dark matter and dark energy constitute $\approx 95\%$ of the universe and the predicted dark matter density is measured as $\Omega_c h^2 = 0.120 \pm 0.001$. Despite the strong evidence for its existence, there is still no information on its exact nature via direct or indirect detection or production in colliders (illustrated in Figure 2.3). Hypothetical candidates for non-baryonic dark matter include weakly interacting massive particles (WIMPs), axions, sterile neutrinos or geons [116, 117].

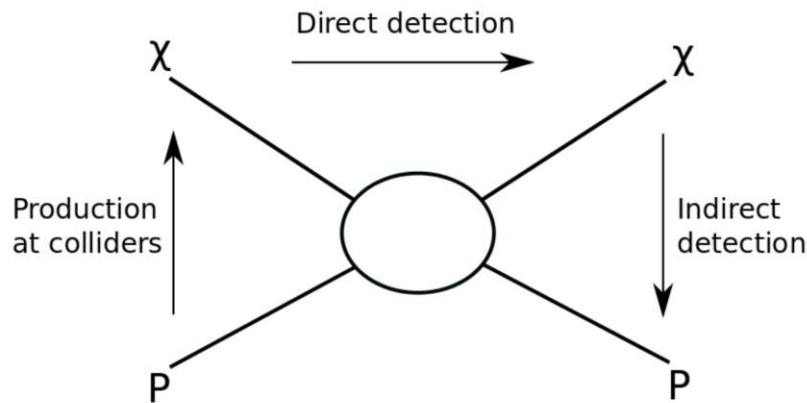


Figure 2.3: Illustration of possible dark matter (DM) detection modes: direct, indirect or production in colliders. An ordinary matter particle is denoted as P, whereas, a DM particle is denoted by χ . [118]

Grand Unification and Gravitational Force

Extrapolations of the energy scale dependence of the coupling constants (discussed above in Section 2.1.2) indicate that at a grand unification energy scale $\Lambda_{GUT} \approx 10^{16}$ GeV, the three fundamental forces in the SM become roughly equal in strength, as shown in Figure 2.8. The grand unified theory (GUT) postulates that all forces can be described by a single gauge field. Such a theory could be described by a larger Lie symmetry group (e.g. $SU(5)$), which is spontaneously broken at the GUT scale. Nevertheless, an *exact* unification of the strong and electroweak forces is not achieved within the SM itself and requires some form of new physics. These energies are beyond the reach of any conceivable collider experiments, so only indirect measurements of particles predicted by GUT models are possible.

The SM does not consider the gravitational force due to the complexity of a unified quantum gravity theory and is neglected due to its much weaker strength at the microscopic scale. However, at such large energy scales approaching the Planck scale ($m_P = 10^{19}$ GeV), quantum gravitational effects begin to become important. Therefore, a complete unification would have to account for an additional gravitational gauge field, postulating the existence of a spin-2 graviton, taking a step closer towards a theory of everything (TOE). The recent direct detection of gravitational waves [119],

predicted from the classical general relativity theory, provides further hints towards the existence of gravitons.

Hierarchy Problem

The sole fact that the GUT scale is so much larger than the EWK scale (or equivalently, gravity is so much weaker than the electroweak force) is an indication of BSM physics. The *hierarchy* (or *naturalness*) problem spans directly from the nature of the Higgs sector and the explicit scale of the EWSB mechanism. The question can be rephrased as to why the discovered Higgs boson is so much lighter than the GUT unification energy (or Planck mass).

The Higgs mass receives radiative quantum corrections from the virtual effects of every massive particle that couples, directly or indirectly, to the Higgs field. The most considerable contributions come from the heaviest fermions (t, b, τ) due to their strong Yukawa couplings with the Higgs field (Figure 2.5). Possible one-loop corrections from fermions and scalars are shown in Figure 2.4.

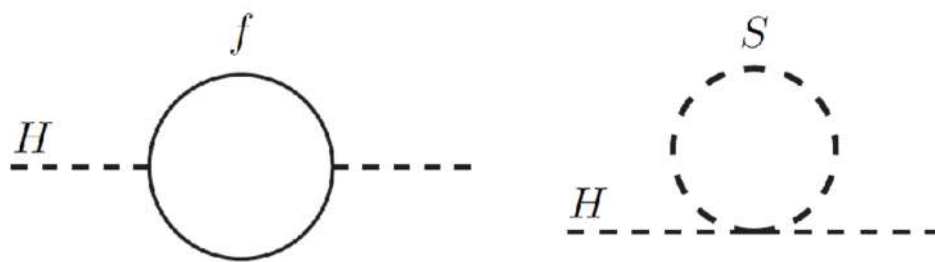


Figure 2.4: One-loop quantum corrections to the Higgs squared mass parameter m_H^2 , due to a (left) Dirac fermion (left) and a scalar (right) [120].

A light mass of the Higgs boson that was observed ($m_H \approx 125$ GeV) is not explicitly forbidden, however, for this to be possible, there must exist substantial negative quadratic vacuum corrections to the Higgs scalar field of the order of $\approx (100 \text{ GeV})^2$:

$$\Delta m_H^2 = -\frac{|\lambda_f^2|}{8\pi^2} \Lambda_{UV}^2 + \dots \quad (2.4)$$

This level of extreme fine-tuning of the Higgs mass is undesirable, especially in the context of any future extensions of the SM or unified theories, in which the Higgs boson mass is expected to be calculable.

2.2 Supersymmetry

Supersymmetry (SUSY) is one of the most promising theories that could provide solutions to central open questions and problems in the SM. The systematic cancellation of the large quantum corrections in the mass of the Higgs boson could be cured with the introduction of an underlying symmetry relating fermions and bosons. A SUSY transformation converts fermionic states into bosonic ones and vice-versa. Therefore, given a supersymmetric generator operator \hat{Q} :

$$\hat{Q} |\text{Fermion}\rangle = |\text{Boson}\rangle \quad \text{and} \quad \hat{Q} |\text{Boson}\rangle = |\text{Fermion}\rangle \quad (2.5)$$

Supermultiplets are irreducible representations of the SUSY algebra, containing both fermionic and bosonic states that are each other's superpartners. They contain the same number of bosonic and fermionic degrees of freedom. Effectively, it postulates that a corresponding superpartner state exists for every SM particle, differing by half-integer spin, with all other quantum numbers conserved.

2.2.1 Minimal Supersymmetric Standard Model

The Minimal Supersymmetric Standard Model (MSSM) [121, 122] is an extension of the SM that adds a minimal set of new parameters in order to realise SUSY. It contains 105 new free parameters (masses, couplings, phases and mixing angles) relating to the extra particle content, which is summarised in Table 2.1 in terms of gauge eigenstates or superfields. Fermionic superpartners of gauge bosons are called *gauginos* and partners of the Higgs boson are called *higgsinos*. Bosonic superpartners of fermions are called *sfermions*, categorised into squarks (\tilde{q}) and *sleptons* $\tilde{\ell}$ (*selectrons* \tilde{e} , *smuons* $\tilde{\mu}$, *staus* $\tilde{\tau}$). The helicity refers to that of the superpartners, and the gauge interactions mirror those of the SM, such as the left-handed coupling of massless *sneutrinos* ($\tilde{\nu}$).

	Particle	Spin	Superpartner	Spin
Quarks	quark q	$1/2$	squark \tilde{q}	0
	$Q : q = u, d, c, s, b, t$		$\tilde{q} = \tilde{u}, \tilde{d}, \tilde{c}, \tilde{s}, \tilde{b}, \tilde{t}$	
Leptons	lepton $\ell : L = (\ell, \nu)$	$1/2$	slepton $\tilde{\ell} : L = (\tilde{\ell}, \tilde{\nu})$	0
	e, μ, τ		$\tilde{e}, \tilde{\mu}, \tilde{\tau}$	
Bosons	gluon g	1	gluino \tilde{g}	$1/2$
	W boson	1	wino \tilde{W}	$1/2$
	W^\pm, W^0		$\tilde{W}^\pm, \tilde{W}^0$	
B boson B^0	1	bino \tilde{B}	$1/2$	
Higgs	Higgs boson H $H_u = (H_u^+, H_u^0)$ $H_d = (H_d^0, H_d^-)$	0	higgsino \tilde{H} $\tilde{H}_u = (\tilde{H}_u^+, \tilde{H}_u^0)$ $\tilde{H}_d = (\tilde{H}_d^0, \tilde{H}_d^-)$	$1/2$

Table 2.1: Extra particle content in the MSSM with particles and their corresponding superpartners, in terms of gauge eigenstates or superfields [120].

Naturalness and Soft SUSY Breaking

Given that no supersymmetric partners of light SM particles have been discovered yet, it indicates that SUSY has to be a symmetry that is broken in the vacuum state. It is expected that this symmetry is spontaneously broken and thereby hidden at low energies. There are many models of spontaneous symmetry breaking (SSB) that extend the MSSM to include new particles and interactions at very high mass scales. If SUSY is to provide a solution to the naturalness problem [123, 124], a realistic phenomenological model introduces *soft* SUSY breaking \mathcal{L}_{soft} where the effective MSSM Lagrangian can be written as:

$$\mathcal{L}_{MSSM} = \mathcal{L}_{SUSY} + \mathcal{L}_{soft} \quad (2.6)$$

where \mathcal{L}_{SUSY} contains all gauge and Yukawa interactions, and \mathcal{L}_{soft} violates SUSY but only contains positive mass terms and coupling parameters.

The Lagrangian describing the soft SUSY breaking in terms of the MSSM gauge eigenstates can be explicitly expressed as:

$$\mathcal{L}_{soft} = -\frac{1}{2}(M_1\tilde{B}\tilde{B} + M_2\tilde{W}\tilde{W} + M_3\tilde{g}\tilde{g} + c.c) - \dots \quad (2.7)$$

where M_1 , M_2 and M_3 are *binos*, *winos*, and *gluinos* mass terms, respectively, while the remaining terms are Yukawa coupling terms. The terms in \mathcal{L}_{soft} include scalars and gauginos but not their respective superpartners and thus break supersymmetry. Therefore, the mass splittings between the SM particles and their partners are determined by the soft SUSY breaking mass terms, dictated by a mass scale chosen by nature.

The additional corrections to the Higgs scalar squared mass (Equation 2.4) are consequently defined as:

$$\Delta m_H^2 = m_{soft}^2 \left[\frac{\lambda}{16\pi^2} \ln(\Lambda_{UV}/m_{soft}) + \dots \right] \quad (2.8)$$

where m_{soft} is the largest SUSY mass scale associated with the soft terms and λ includes various dimensionless couplings. Naturalness favours superpartner masses (especially stops, higgsinos and gluinos) that are not too large, with the fine-tuning corrections to the Higgs scalar mass defined by the SUSY mass scale.

SUSY Interactions

The MSSM superpotential is expressed as:

$$W_{MSSM} = \bar{u}\mathbf{y}_uQH_u - \bar{d}\mathbf{y}_dQH_d - \bar{e}\mathbf{y}_eLH_d + \mu H_uH_d \quad (2.9)$$

where the terms are the chiral superfields corresponding to the chiral supermultiplets (Q , L , H) in Table 2.1 and the $\mathbf{y}_{u,d,e}$ are Yukawa coupling matrices. The μ term is the higgsino mass parameter, and it is the supersymmetric version of the Higgs boson mass in the SM, giving rise to the Yukawa couplings and masses of all quarks and leptons.

The Yukawa matrices determine the masses and CKM mixing angles of ordinary quarks and leptons.

The heaviest fermions (t, b, τ) give the largest contributions to the Yukawa couplings. The top (s)quark Yukawa couplings in the MSSM are shown in Figure 2.5, where the right- and left-handed components couple to the neutral Higgs(ino) fields with the same coupling strength y_t .

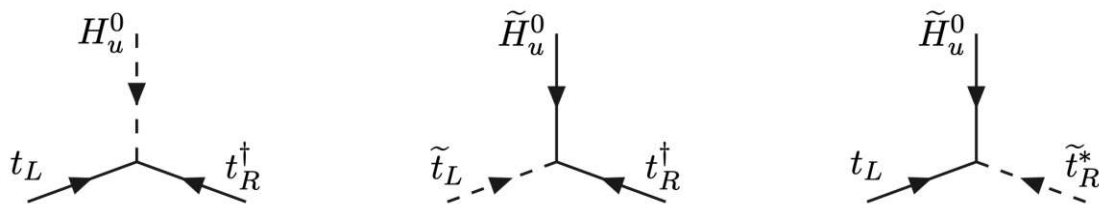


Figure 2.5: Top quark Yukawa couplings and their supersymmetrised versions in the MSSM. The right- and left-handed components couple to the neutral Higgs(ino) fields with the same coupling strength y_t . [120]

Although the significant Yukawa couplings are those of third-family fermions, the production and decay processes in the MSSM are typically dominated by SUSY interactions of the gauge-coupling strength. The couplings of the SM gauge bosons are determined entirely by LGI. The gauginos (gluino \tilde{g} , wino \tilde{W} , bino \tilde{B}) also couple to (s)quark, (s)lepton and Higgs(ino) pairs, as shown in Figure 2.6, with coupling strengths proportional to the electroweak gauge couplings g and g' . Winos only couple to left-handed squarks and sleptons, and the bino coupling is also proportional to the weak hypercharge Y . The gluino \tilde{g} is a colour octet fermion and is thus unique in that it does not mix with any other MSSM particles.

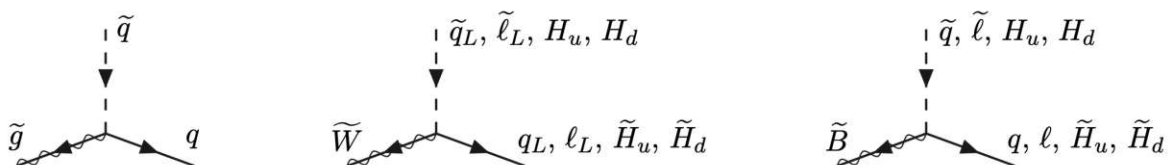


Figure 2.6: Couplings of gauginos (gluinos \tilde{g} , winos \tilde{W} , binos \tilde{B}) in the MSSM [120].

Higgsinos and EWSB

For the Higgs scalar boson, there must be two Higgs chiral supermultiplets with two copies of the Higgs field to cancel the gauge anomalies of the electroweak gauge symmetry. The doublet complex scalar fields are H_u and H_d , with weak hypercharge $Y = \pm 1/2$. The weak isospin components of the Higgs isodoublets have positive and neutral electrical charges with $T_3 = \pm 1/2$, as shown in Table 2.1 (H_u^+ , H_u^0 and H_d^0 , H_d^+). The fermionic superpartners of the Higgs scalar fields are \widetilde{H}_u and \widetilde{H}_d , are called higgsinos and similarly are broken down into superpartner states of the corresponding weak isospin components.

Following electroweak symmetry breaking (EWSB) the charged components of the Higgs scalars cannot have vacuum expectation values (VEVs), so H_u^+ and H_d^- are set to zero, while the neutral components H_u^0 and H_d^0 acquire non-zero VEVs. Their ratio is defined by $\tan \beta \equiv v_u/v_d \equiv \langle H_u^0 \rangle / \langle H_d^0 \rangle$, which, together with the Yukawa couplings, determines the masses and CKM mixing angles of the quarks and leptons in the MSSM.

The Higgs gauge-eigenstate fields can be expressed in terms of the mass eigenstates as:

$$\begin{pmatrix} H_u^0 \\ H_d^0 \end{pmatrix} = \begin{pmatrix} v_u \\ v_d \end{pmatrix} + \frac{1}{\sqrt{2}} R_\alpha \begin{pmatrix} h^0 \\ H^0 \end{pmatrix} + \frac{i}{\sqrt{2}} R_{\beta_0} \begin{pmatrix} G^0 \\ A^0 \end{pmatrix} \quad (2.10)$$

where G^0 , G^\pm are Nambu-Goldstone bosons which become longitudinal modes of the Z^0 and W^\pm vector bosons. The remaining five scalars are CP-even (h^0 , H^0), CP-odd (A^0) and two charged scalars (H^\pm). The masses of A^0 , H^0 and H^\pm can be arbitrarily large, while the mass of h^0 is bounded by an upper value and is considered the lightest Higgs boson of the MSSM.

Incorporating the remaining loop (including two- and three-loop) effects of radiative corrections to the Higgs scalar mass (discussed above), it was estimated prior to the Higgs discovery that $m_{h^0} < 135$ GeV in the MSSM, making it compatible with the Higgs boson of mass 125 GeV discovered at the LHC.

Electroweakinos

The wino \widetilde{W} and bino \widetilde{B} states are not mass eigenstates and mix with higgsinos \widetilde{H} due to EWSB. *Neutralinos* $\widetilde{\chi}_i^0$ are mixtures of neutral electroweak gauginos ($\widetilde{W}^0, \widetilde{B}$) and the neutral higgsinos ($\widetilde{H}_u^0, \widetilde{H}_d^0$) forming four different mass eigenstates ($i = 1, 2, 3, 4$). The winos \widetilde{W}^\pm and charged higgsinos (\widetilde{H}_u^\pm and \widetilde{H}_d^\pm) combine to form two *charginos* $\widetilde{\chi}_i^\pm$ ($i = 1, 2$). The lightest neutralino $\widetilde{\chi}_1^0$ is often assumed to be the lightest supersymmetric particle (LSP) in R-parity conserving models (discussed in Section 5.1.2).

The neutralino mass part of the MSSM Lagrangian is governed by the bino M_1 , wino M_2 (Equation 2.7) and higgsino μ (Equation 2.9) mass terms as well as $\tan \beta$. A common assumption made in phenomenological models is that $M_1 \approx M_2/2$ at the electroweak scale, motivated by gaugino unification at the GUT scale predicted by renormalisation group evolution (RGE) shown in Figure 2.10. In that scenario, the neutralino masses and mixing angles only depend on three unknown parameters.

Depending on the mass parameters, neutralino mass eigenstates can be predominantly bino-like, wino-like and higgsino-like. Many models assume $M_1 < M_2 \ll |\mu|$, emerging from mSUGRA [125] supergravity boundary conditions required for EWSB, leading to a bino-like neutralino LSP $\widetilde{\chi}_1^0$. In the approximation $m_Z \ll |\mu \pm M_{1,2}|$, the lightest neutralino $\widetilde{\chi}_1^0$ is bino-like ($\approx \widetilde{B}$), $\widetilde{\chi}_2^0$ is wino-like ($\approx \widetilde{W}^0$) and, $\widetilde{\chi}_3^0$ and $\widetilde{\chi}_4^0$ are higgsino-like ($\approx \widetilde{H}^0$). In this approximation, the chargino $\widetilde{\chi}_1^\pm$ is mass degenerate with $\widetilde{\chi}_2^0$, and the higgsino-like $\widetilde{\chi}_3^0, \widetilde{\chi}_4^0$ and $\widetilde{\chi}_2^\pm$ have masses of order $|\mu|$. The scenario of $M_1 \approx M_2/2 \ll |\mu|$ often serves as a benchmark in a number of phenomenological models.

A prediction coming from mSUGRA [125] or GMSB [126, 127] boundary conditions is an approximate ratio between the electroweak and gluino mass parameters of

$$M_3 : M_2 : M_1 \approx 6 : 2 : 1$$

near the TeV scale. This would imply that the gluino is considerably heavier than the

lighter EWKinos. However, one must take into account that with the gluino being a strongly interacting particle, M_3 is a running mass parameter, strongly depending on the energy scale, as shown in Figure 2.10.

Sfermions: Squarks and Sleptons

Mass eigenstates of squarks and sleptons are acquired from the squared-mass matrices of up-type squarks, down-type squarks, charged leptons and sneutrinos. The majority of the mixing angles are predicted to be very small. The third family squarks (\tilde{t} , \tilde{b} , $\tilde{\tau}$) can have different masses with respect to the other families due to the effects of the large Yukawa and soft couplings in the RGE. They are also expected to have significant mixing between their left- and right-handed components. The first and second family squarks and sleptons have negligible Yukawa couplings, resulting in nearly degenerate unmixed pairs.

For top squarks (stops), there are two mass eigenstates (\tilde{t}_1 , \tilde{t}_2), and the mixing of the states is defined by the stop mixing angle $\theta_{\tilde{t}}$ (with $0 < \theta_{\tilde{t}} < \pi/2$) [128]:

$$\begin{pmatrix} \tilde{t}_L \\ \tilde{t}_R \end{pmatrix} = \begin{pmatrix} \sin \theta_{\tilde{t}} & \cos \theta_{\tilde{t}} \\ -\cos \theta_{\tilde{t}} & \sin \theta_{\tilde{t}} \end{pmatrix} \begin{pmatrix} \tilde{t}_1 \\ \tilde{t}_2 \end{pmatrix} \quad (2.11)$$

Mass Spectrum

Ultimately, in the MSSM there are 32 distinct masses, excluding the gravitino \tilde{G} . The supersymmetric mass eigenstates formed from the gauge eigenstates are summarised in Table 2.2.

Specific models consider minimal forms with a subset of free parameters, such as unmeasured parameters in mSUGRA or GMSB models. Different assumptions on the various model parameters, including eigenstate mixing, result in a large number of different possible mass spectra and, consequently, possible decay chains.

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

Table 2.2: Supersymmetric mass eigenstates formed from the gauge eigenstates in the MSSM. The sfermion (squark and slepton) mixing for the first two families is assumed to be negligible. [120]

A generic illustration of a potential SUSY mass spectrum motivated by naturalness and flavour considerations is shown in Figure 2.7.

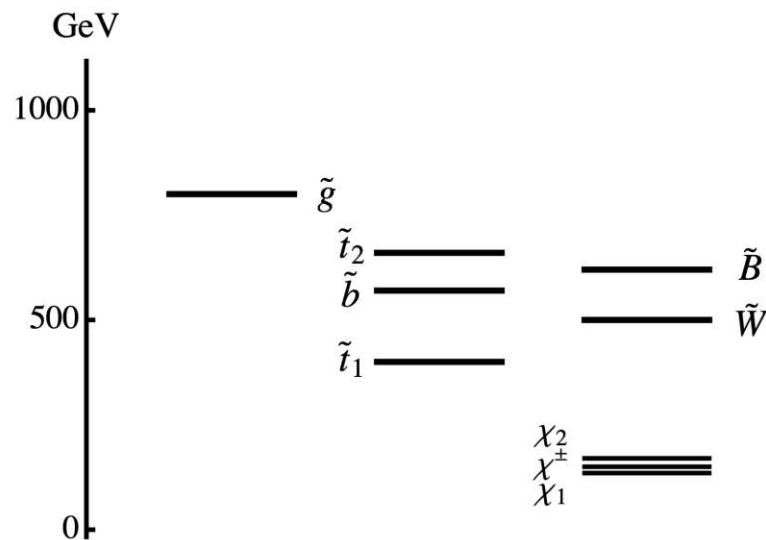


Figure 2.7: Generic natural SUSY mass spectrum with masses below 1 TeV. Lightest fermions are mostly higgsino-like with $m_{\tilde{\chi}_1^0} < m_{\tilde{\chi}_1^\pm} < m_{\tilde{\chi}_2^0}$. [124]

All other SUSY particle masses can be much heavier without violating naturalness. Therefore, they can be considered decoupled from the theory and phenomenologically irrelevant for LHC searches.

Grand Unification

RGE evolves the MSSM gauge couplings, superpotential parameters and soft terms from a high to a low energy scale. The RGE evolution of the fundamental gauge couplings α in the SM and the MSSM is shown in Figure 2.8. An illustrative example of the RGE evolution of scalar and gaugino mass parameters for a specific parameter configuration and boundary conditions is shown in Figure 2.10 below.

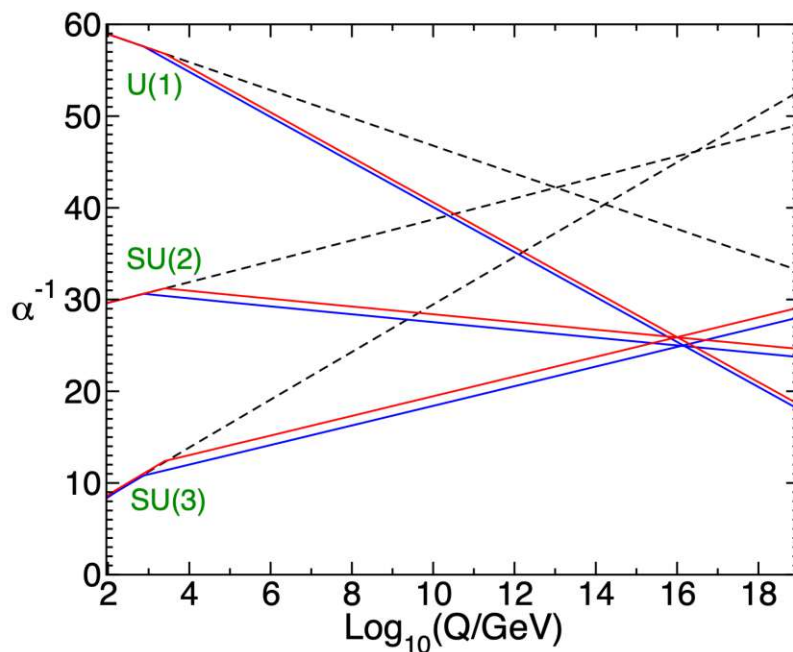


Figure 2.8: Two-loop renormalisation group evolution (RGE) of the inverse gauge couplings $\alpha^{-1}(Q)$ in the SM (dashed lines) and the MSSM (solid lines), for the electromagnetic (SU(1)), weak (SU(2)) and strong (SU(3)) forces. In the MSSM case, the sparticle masses are treated as a common threshold varied between 750 GeV and 2.5 TeV, and $\alpha_3(m_Z)$ is varied between 0.117 and 0.120. [120]

The additional sparticle content could allow for the unification of strong and electroweak gauge couplings at a GUT scale ($M_U \approx 1.5 \times 10^{16}$ GeV), which is not achieved in the SM. One must note that the unification is not entirely perfect, with $\alpha_3(M_U) <$

$\alpha_1(M_U) = \alpha_2(M_U)$, however, this difference can be accounted for depending on the new particle content at this scale. This is another of the strong motivations for SUSY in the context of a grand unified theory (GUT) or superstring models.

2.2.2 Light Stop Window and Compressed SUSY

In order to scale down in the realm of possibilities, one should consider phenomenologically viable models motivated by theoretical considerations and constrained by experimental measurements. A specific configuration is considered, imposing constraints by naturalness and flavour considerations as well as additional motivations from dark matter relic density predictions. The relevant phenomenological properties are described in terms of a few physical parameters, which allows a simple characterisation of its main features, useful to define a strategy for its discovery by LHC searches. Further phenomenological motivations include the uncovered parameter space due to the difficulty of performing such an analysis in terms of experimental limitations and background separation.

Light Stop

Similarly to the SM, the Higgs scalar mass is subject to significant quantum corrections $\Delta(m_H^2)$ in the MSSM, introduced in Equation 2.8, indicating that naturalness favours sparticle masses that are not too large. The largest contributions to the radiative corrections come from the heaviest fermions with the strongest Yukawa couplings with the Higgs fields (Figure 2.5), and the most significant are the top and stop loops shown in Figure 2.9.

The size of the radiative corrections is dependent on the stop masses $m_{\tilde{t}_1}, m_{\tilde{t}_2}$ and the mixing of the states, defined in Equation 2.11. In order to achieve the observed Higgs boson mass, there is a trade-off between either large stop mixing or heavy stop masses. Considering the natural case of light stops, the maximum h^0 mass with the smallest radiative corrections occurs for rather large stop mixing, referred to as the

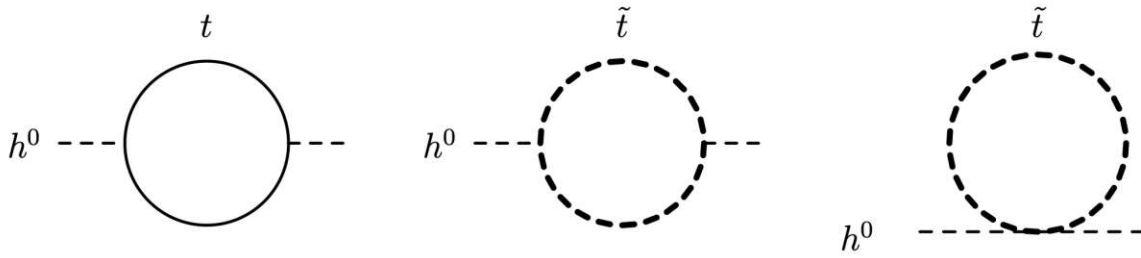


Figure 2.9: Contributions to the lightest Higgs h^0 squared mass from top quark and squark one-loop diagrams [120].

maximal mixing scenario. Ultimately, naturalness favours light stops, and in many natural SUSY models the lightest top squark \tilde{t}_1 is expected to be the lightest particle within the squark sector.

In order to have light stops, the existence of a substantial stop trilinear A_t term ($|A_t/m_t| \gtrsim 1$) is required. The lightest average stop mass between the two stops $m_s \equiv \sqrt{m_{\tilde{t}_1} m_{\tilde{t}_2}}$ that can lead to the observed Higgs boson mass is obtained for when $m_s \approx 500$ GeV. If there exists a large splitting between $m_{\tilde{t}_1}$ and $m_{\tilde{t}_2}$ (as indicated in Figure 2.7), this pushes the mass of the lighter stop \tilde{t}_1 lower and within the potential reach of LHC searches. The RGE evolution in the context of third-generation squark masses, shown in Figure 2.10, can lead to a large splitting between the right- and left-handed stop components where $m_{\tilde{t}_R} \ll m_{\tilde{t}_L}$ at the weak scale. This further motivates a light stop \tilde{t}_1 that is mostly right-handed.

Therefore, one can consider the specific case where the lightest stop \tilde{t}_1 is mostly right-handed with a mass below a TeV, and thus accessible by the LHC. The heavier stop \tilde{t}_2 , which is mostly left-handed, can be considered to have a much larger mass in the 1-2 TeV range. The masses are correlated with one another such that the geometric average $m_s \approx 500$ -600 GeV.

Additionally, in the case of light gauginos or higgsinos, a light stop can have an impact on low-energy flavour-physics observables. This is in the case that the light stop is mostly right-handed and thus aligned with the top quark, with flavour violation described by the CKM matrix. One can impose flavour constraints on the deviations

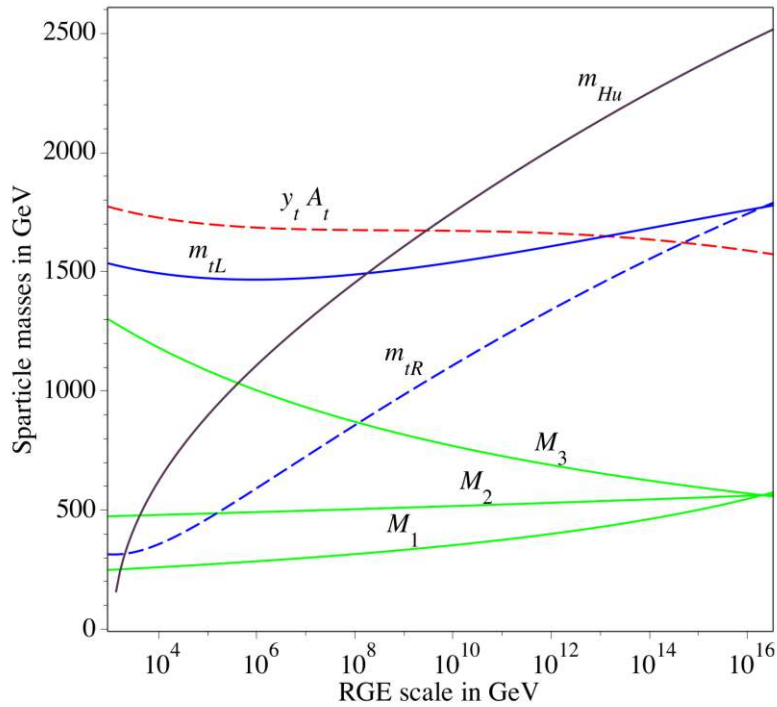


Figure 2.10: Illustrative example of renormalisation group evolution (RGE) from the unification scale to the weak scale of gaugino masses M_1 , M_2 , M_3 (green curves), of the stop mass parameters $m_{\tilde{t}_L}$ and $m_{\tilde{t}_R}$ (full and dashed blue curves, respectively), $y_t A_t$ (red dashed curve), m_{H_u} (black curve). [129]

from the SM of the $BR(B \rightarrow X_s \gamma)$ decay, which rule out a light left-handed stop but are consistent with a light right-handed stop, further motivating a large mass splitting between the two.

Dark Matter and Compressed Mass Spectra

In R-parity conserving models, the lightest supersymmetric particle (LSP) is massive and stable. If it is electrically neutral, then it will weakly-interact with ordinary matter. Consequently, it is an excellent non-baryonic dark matter (DM) candidate, which is another strong motivation for SUSY. In many models, the lightest neutralino $\tilde{\chi}_1^0$ (typically bino-like) is often assumed to be the LSP.

In generic SUSY models, the measured cosmological dark matter relic density (discussed above in Section 2.1.3) is excessive for a bino LSP and too small for a higgsino or wino LSP. A light stop opens up a window to reproduce the correct cosmological dark matter relic density via the process of co-annihilation of light stops (or charginos $\tilde{\chi}_1^\pm$) and LSPs (neutralinos $\tilde{\chi}_1^0$) [130, 131, 132], which was depleted in the early universe. This can occur if the mass difference Δm between the stop and the LSP is relatively small (≈ 30 GeV) or almost *degenerate*, as shown in Figure 2.11. In this case, one considers the mass spectrum or model as *compressed*.

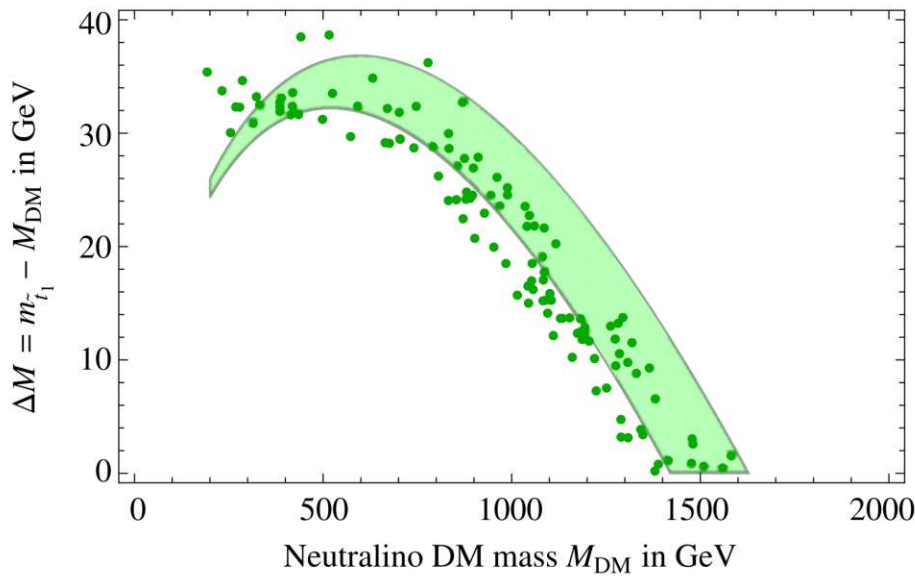


Figure 2.11: SUSY parameter space that leads to the correct cosmological dark matter relic density at small mass splittings Δm between light stops and LSPs [129].

Experiment

3.1 CERN

CERN, the European Council for Nuclear Research (*Conseil Européen pour la Recherche Nucléaire*), is Europe's largest research facility focusing on fundamental research in the area of particle physics. It consists of accelerator facilities that provide particle beams to a number of experiments, ranging from colliding beam detectors to fixed-target experiments.

3.1.1 Accelerator Complex

The Large Hadron Collider (LHC) at CERN is the largest and most energetic particle accelerator and collider in the world [49, 50]. It forms the final element of a large accelerator complex, which sequentially boosts the energy of a beam of protons or heavy ions. A simplified layout of the accelerating structures can be seen in Figure 3.1.

- The protons originate from a Hydrogen bottle where H gas atoms are injected into a duoplasmatron cylinder to have their electrons stripped off in an electric field.
- The beam of protons is then focused and accelerated in a radio frequency quadrupole (RFQ) cavity.

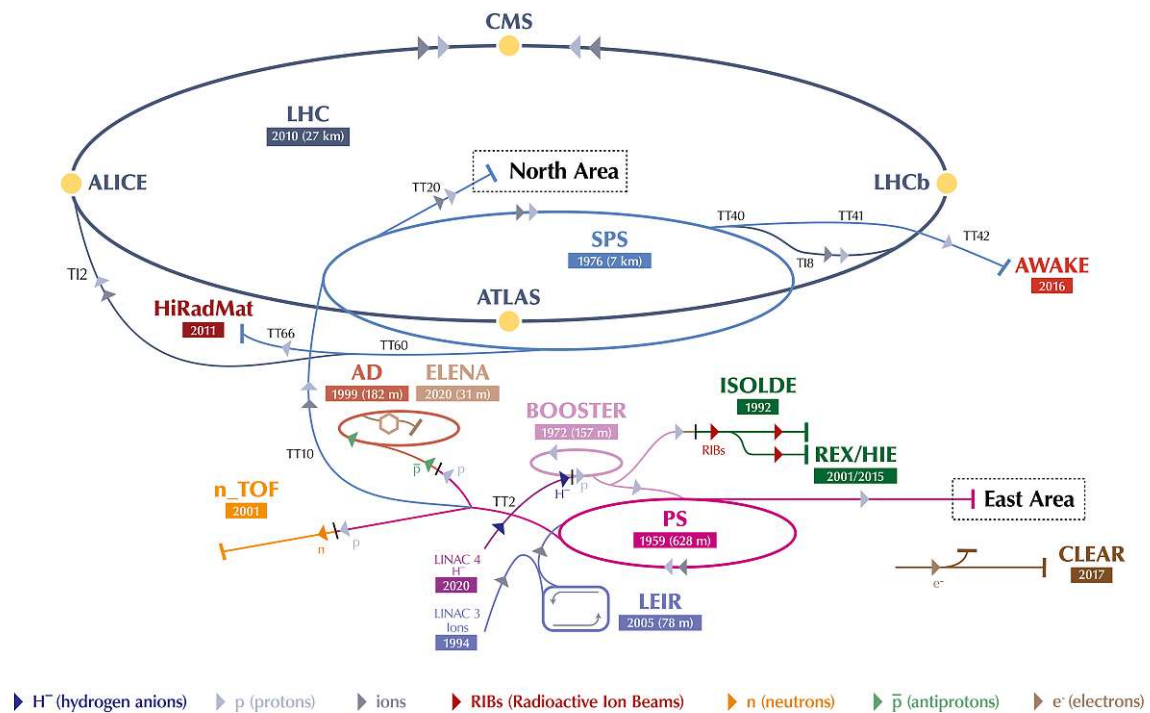


Figure 3.1: Schematic of the accelerator complex at CERN, with the LHC as the last ring [133].

- The linear accelerator *Linac 2* uses radio-frequency cavities to charge cylindrical conductors in an alternating manner, with which the beam is accelerated to an energy of 50 MeV.
- The next step is the Proton Synchrotron Booster (PSB) [134], which is made from four synchrotron rings that use strong electromagnets to accelerate the beam to 1.4 GeV in a circular orbit.
- These are then passed onto a larger synchrotron machine, the Proton Synchrotron (PS) [80, 81] where the protons reach an energy of 25 GeV.
- The larger version of the PS is the Super Proton Synchrotron (SPS) [47], which boosts the energy up to 450 GeV.
- Subsequently, the beam is split and injected into the LHC in two separate beam pipes travelling in opposite directions, where each is ramped up to a maximum energy of 6.8 TeV, corresponding to a centre-of-mass energy $\sqrt{s} = 13.6$ TeV.

A separate physics programme also focuses on accelerating heavy ion beams at lower energies, which are injected into the accelerating chain from the Low Energy Ion Ring (LEIR) [135], resulting in ion-ion or proton-ion collisions (e.g. lead-lead). The data is used to primarily study quark-gluon plasma (QGP), recreating conditions similar to that of the early universe.

3.1.2 Large Hadron Collider

The Large Hadron Collider (LHC) is located in a 26.7 km circular tunnel that is up to 175 m underground. It consists of two adjacent beam pipe rings in vacuum, surrounded by thousands of superconducting electromagnets. The electromagnets are cooled by liquid Helium to temperatures down to 1.9 K and generate magnetic fields up to 8.3 T. They are arranged in a lattice along the path of the charged particle beam to stabilise, align and bend it accordingly. There are 1232 niobium-titanium dipole magnets, each 14.3 m long, that bend the beams, and 392 quadrupole magnets that focus the beams. The particle beams are made to collide at four interaction points around the accelerator ring, where the large detector experiments are located:

- CMS: Compact Muon Solenoid [136]
- ATLAS: A Toroidal LHC ApparatuS [137]
- ALICE: A Large Ion Collider Experiment [138]
- LHCb: LHC Beauty Experiment [139]

At these four points, insertion magnets, such as an inner triplet consisting of three quadrupoles, squeeze the beam down to $\approx 16 \mu\text{m}$ across to increase the collision probability. Finally, the opposing beams are directed to cross and collide, allowing the experiments to detect the outgoing particles. CMS and ATLAS are general-purpose detectors with physics programmes ranging from searches for new physics (e.g. SUSY or exotic particles) to measurements of the SM properties, such as those of the Higgs boson. The LHCb experiment focuses on interactions of b-hadrons, including the measurement of the parameters of CP violation, while the ALICE experiment is optimised

to study quark-gluon plasma production in heavy-ion collisions. There also exist five complementary experiments (TOTEM [140], LHCf [141], FASER [141], MoEDAL-MAPP [142], SND@LHC [143]) with dedicated purposes such as measurements of forward particles from elastic collisions, magnetic monopoles and weakly interacting particles. The CMS-TOTEM Precision Proton Spectrometer (CTPPS) [144] was recently integrated into CMS to perform combined measurements. It is located in the very forward region on both sides of CMS at about 200 m from the interaction point.

LHC Performance

The performance of colliders is described by the quantity of instantaneous luminosity, \mathcal{L} , which is a process-independent measure of the number of interactions per unit area and time [49]. It is defined in Equation 3.1, where dN/dt is the reaction rate and σ is the cross section, typically expressed in barns ($1 b = 10^{24} \text{ cm}^2$). It can be redefined in terms of the specific beam parameters, such as the beam widths in the x and y directions, σ_x and σ_y , respectively, and the number of protons per bunch, N . It can also be parametrised in terms of the beam emittance, ϵ , and the size of the beam at the interaction point, β^* , which describe the LHC optics.

$$\mathcal{L} = \frac{1}{\sigma} \frac{dN}{dt} = f_{\text{BX}} \frac{N^2}{4\pi\sigma_x\sigma_y} F = f_{\text{BX}} \frac{N^2}{4\epsilon\beta^*} F \quad (3.1)$$

A geometric factor, F , is added to reflect the luminosity reduction due to the crossing angle θ_c , bunch length σ_z (RMS) and transverse beam size σ^* (RMS) at the interaction point, defined as:

$$F = \frac{1}{\sqrt{1 + \left(\frac{\theta_c\sigma_z}{\sigma^*}\right)^2}} \quad (3.2)$$

The LHC was designed to operate at a $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ [49, 50]. At $\sqrt{s} = 13 \text{ TeV}$, the protons bunches travel almost at the speed of light c , circling the LHC ring with an orbit frequency $f_{\text{LHC}} = 11\,245 \text{ Hz}$. This is quantified by a bunch spacing of 25 ns, with

a minimum of 7.5 m between bunches, or equivalently, a bunch crossing frequency (or revolution frequency) of $f_{\text{BX}} = 40$ MHz. The nominal filling scheme comprises $n_{\text{bb}} = 2808$ colliding bunches with approximately $N = 1.1 \times 10^{11}$ protons each. The bunches typically come in trains of 72 with empty spaces in between, where 80 % of the possible positions are filled, resulting in an effective bunch crossing rate of $\overline{f_{\text{BX}}} = f_{\text{LHC}} * n_{\text{bb}} \approx 30$ MHz.

Elastic (or *soft*) collisions result in the protons scattering off each other without interacting much and changing their form. In more interesting inelastic (or *hard*) collisions, the quarks and gluons of protons strongly interact and result in a large number of new particles due to the conversion of the relativistic energy into mass. The proton-proton QCD interaction has a total (*minimum bias*) cross section of approximately 100 mb, of which the branching fraction to hard inelastic collisions is ≈ 60 %. With the nominal beam currents, this results in an expected reaction rate of approximately 600 million hard collisions per second. Additionally, there are an order of 20-80 inelastic proton-proton interactions during the same or neighbouring bunch crossings, defined as pileup (PU). The PU is dependent on the exact beam parameters and evolves together with the \mathcal{L} . Over time, the beam suffers a reduction in its \mathcal{L} , due to the decreased number of protons after interaction, however, this can be recovered by tuning the beam parameters with levelling techniques. The typical beam lifetime is of the order of 10 hours, until the \mathcal{L} goes below a given threshold for the beam to be *dumped*.

The integrated luminosity, $L = \int \mathcal{L} dt$, is a measure of the total data delivered by the LHC and collected by the experiments and is often expressed in terms of inverse femtobarns fb^{-1} . The timeline of the operational running of LHC has been split into so-called *Runs*, which are divided by *Long Shutdown* periods where the machine and the detectors undergo technical upgrades. Furthermore, within runs, there are *Technical Stops* where still relatively large upgrades can be implemented. For example, the CMS pixel detector underwent a *Phase I* upgrade [145, 146] during the *Technical Stop* at the end of 2016 to handle the harsher conditions. Therefore, the running conditions

of the LHC have varied significantly over the years of data-taking and ultimately surpassed the initial design parameters. The experiments work very closely together with the LHC to ensure that they can adapt to the improved, but at the same time harsher, conditions for the detectors.

The LHC *Run 1* lasted from 2009 until 2013, and LHC delivered a total of 6.1 fb^{-1} of data at $\sqrt{s} = 7 \text{ TeV}$ and 23.3 fb^{-1} data at $\sqrt{s} = 8 \text{ TeV}$ with a mean PU of 21. Data collected in the first *Run* of the LHC led to the discovery of the Higgs boson. For *Run 2*, which lasted from 2015 until 2018, the LHC upgrades allowed the increase of the centre-of-mass energy to $\sqrt{s} = 13 \text{ TeV}$. In the later stages of *Run 2*, LHC achieved double the design luminosity (Figure 3.3), with a mean PU of 32 and tails going up to 80 (Figure 3.4), amounting to $\approx 163 \text{ fb}^{-1}$ of delivered data (Figure 3.2) [147]. The increase in LHC luminosity, and consequently PU, was motivated to increase the probability of hard collisions of interest, to ultimately provide a larger dataset for searches of new physics and studies of rare interactions. LHC *Run 3* started in 2022 with a peak luminosity up to $\mathcal{L} \approx 2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ and peak PU ≈ 60 , collecting $\approx 42 \text{ fb}^{-1}$ at a slightly higher energy $\sqrt{s} = 13.6 \text{ TeV}$. The run is currently ongoing, with the potential to go up to $\mathcal{L} = 2.4 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ and PU ≈ 70 .

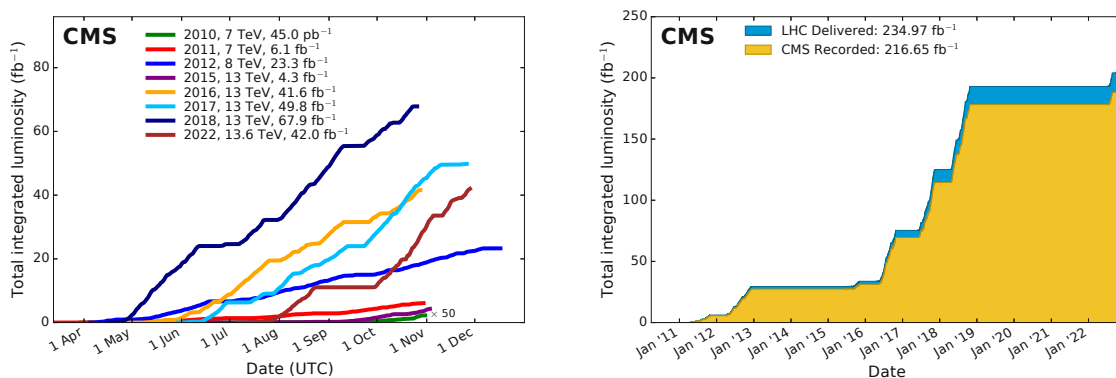


Figure 3.2: Cumulative total integrated luminosity L delivered by LHC versus day (left) and delivered and recorded by CMS versus date (right), for all years during stable beams for p-p collisions [147].

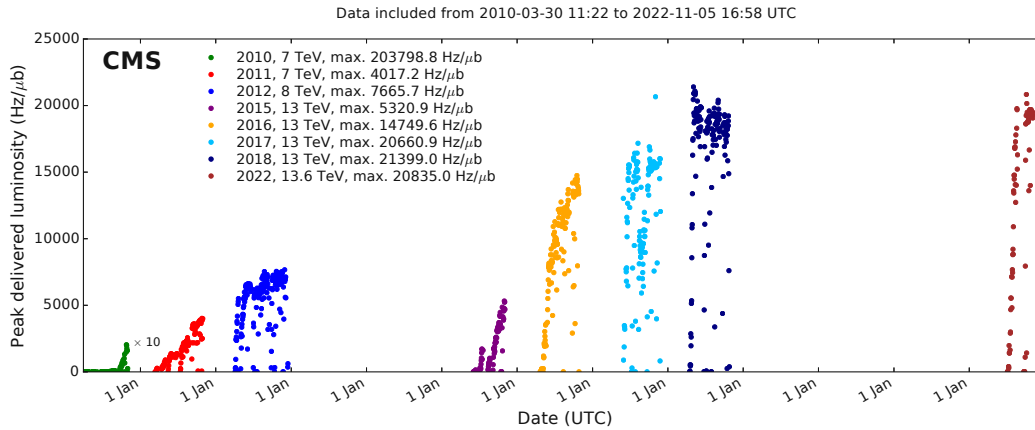


Figure 3.3: Peak instantaneous luminosity \mathcal{L} (left) and pileup PU (right) over all years during stable beams for p-p collisions [147].

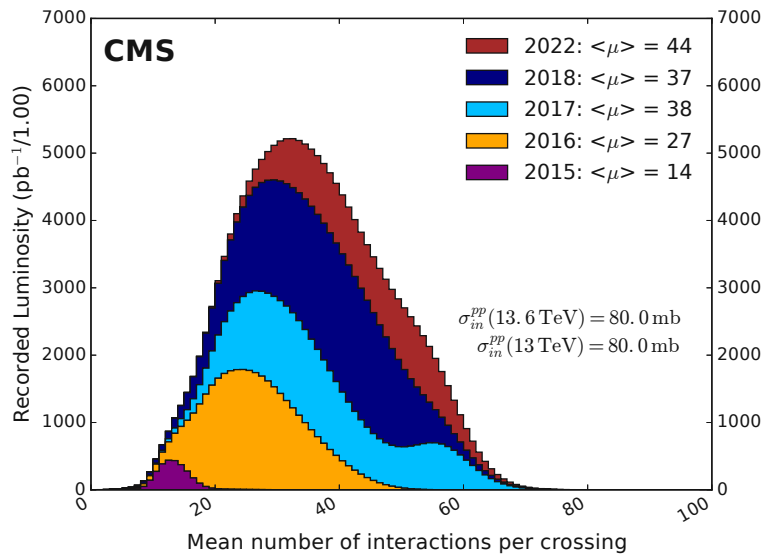


Figure 3.4: Mean pileup (PU) for all years in LHC *Run 2* and *Run 3* during stable beams for p-p collisions at $\sqrt{s} = 13$ TeV, assuming a minimum bias cross section of 80 mb [147].

Beyond *Run 3*, during the *LS3* from 2025 until 2027, the LHC will undergo a major upgrade known as the High Luminosity LHC (HL-LHC) [148] to achieve instantaneous luminosities up to ten times larger than the nominal design value, with PU of at least 140. Figure 3.5 shows the event display of a special high PU fill recorded during LHC *Run 2*, underlining the difficulties that will need to be overcome. This will allow experiments to enlarge their collected data by one order of magnitude compared with

the LHC baseline programme in order to increase the potential for new discoveries.

In the context of the Future Circular Collider (FCC) study for hadron colliders, there are concepts of a High Energy LHC (HE-LHC) [149] in the LHC tunnel, with a centre-of-mass energy $\sqrt{s} = 27$ TeV and target luminosity four times that of HL-LHC. There are also prospects of a much larger accelerator located in a 100 km tunnel which could collide hadrons (FCC-hh) or electrons and positrons e^-e^+ (FCC-ee) at 100 TeV [150]. Other possible future e^+e^- colliders include linear colliders such as the ILC [151] or CLIC [152], or the circular CEPC [153] accelerator.

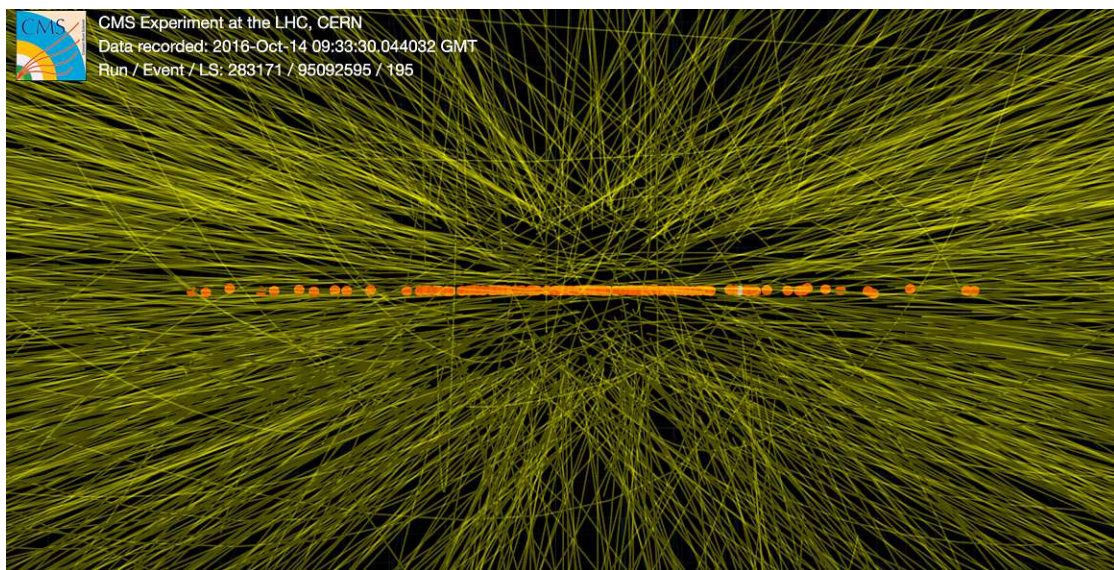


Figure 3.5: Proton-proton collision at a centre-of-mass energy of 13 TeV recorded during a special high pileup fill of LHC Run 2. The events are from isolated bunches with an average pileup (PU) of around 100. [154]

3.2 CMS Experiment

3.2.1 Scientific Programme

The scientific programme of the Compact Muon Solenoid (CMS) experiment [136] covers a broad spectrum of high-energy physics phenomena with diverse final state signatures. One of the main focuses are searches for new physics, such as SUSY or exotic particles, which are motivated by the shortcomings of the SM (covered in Section 2.1.3). Another important focus area is precision measurements of SM properties in the electroweak, top quark and QCD sectors, with special attention to b-quark physics. Furthermore, following the discovery of the Higgs boson, measuring and confirming its properties became a central point of the programme. Moreover, data from heavy-ion collisions are also analysed, expanding our insight into the dynamics of quark-gluon plasma (QGP).

In order to carry out the diverse and dynamic scientific programme, the corresponding particle detector is required to have the following performance [136]:

- Excellent muon identification and momentum resolution; charge identification of muons with $p_T < 1$ TeV; good dimuon mass resolution ($\approx 1\%$ at 100 GeV).
- Excellent charged particle momentum resolution and reconstruction efficiency, requiring highly efficient tracking for $p_T > 1$ GeV; efficient determination of the primary interaction vertex.
- Efficient triggering and offline tagging of tau leptons and jets originating from b-quarks (*b-jets*) with secondary decay vertices (SVs), requiring tracking pixel detectors close to the interaction region.
- Excellent electromagnetic energy resolution; measurement of the direction of photons; π^0 rejection and efficient photon and lepton isolation; good di-photon and dielectron mass resolution ($\approx 1\%$ at 100 GeV).
- Wide geometric coverage for the identification of muons, electrons and photons.

- Excellent missing transverse momentum and dijet mass resolution, requiring hadron calorimeters with a broad hermetic geometric coverage and fine segmentation.

3.3 CMS Detector

The CMS detector [155, 156, 136] is a large-scale general-purpose particle detector at one of the interaction points (Point 5) of the LHC, located in a cavern approximately 100 m underground. It has dimensions of 15 m in diameter and a length of 21 m, with a mass of around 14 000 t. Its design is based on the typical structure of multi-purpose detectors at colliders, which consists of cylindrical layers of specific subsystems built around the beam axis, each with a different purpose. The major subsystems of CMS are shown in Figure 3.6:

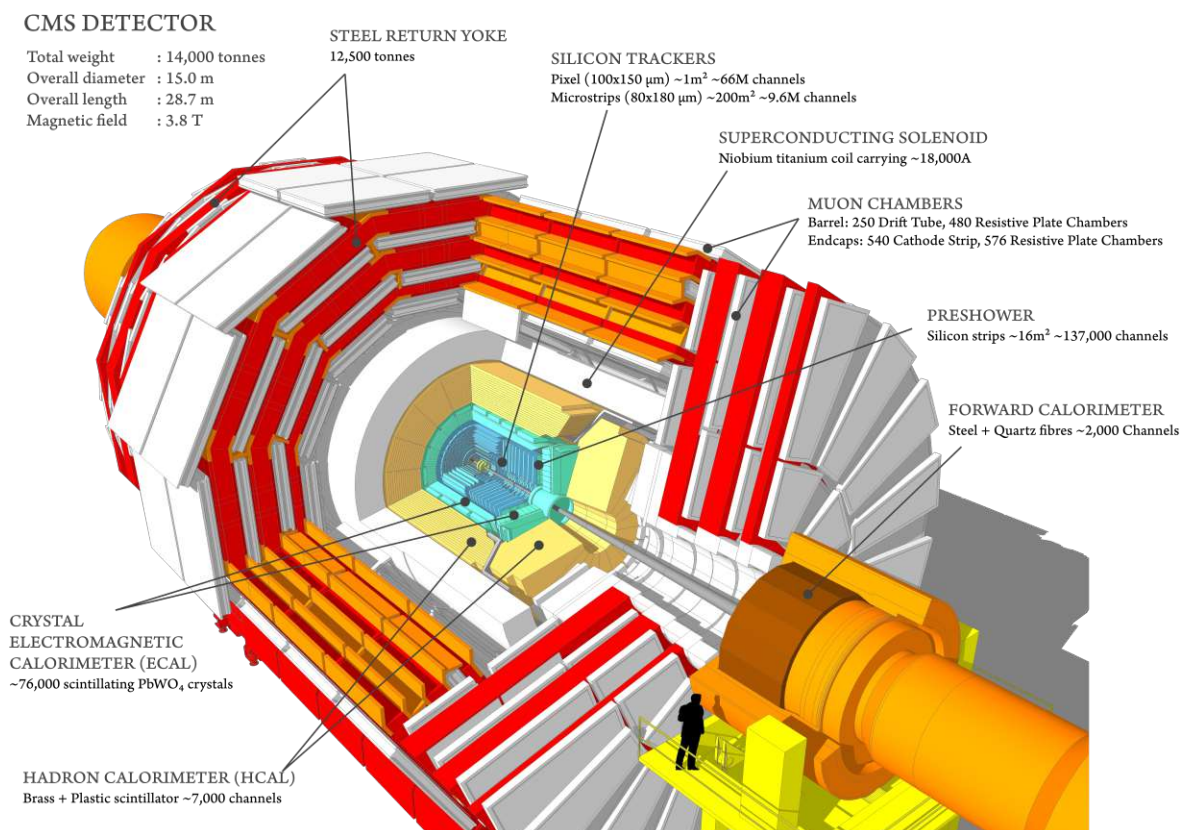


Figure 3.6: Schematic of the CMS detector indicating the most important subsystems and their scale [136].

- The innermost part of the detector that surrounds the beam interaction region consists of a **silicon pixel** and **micro-strip tracker**, which can accurately measure the trajectory and momentum of charged particles.
- The middle layer consists of **crystal electromagnetic** and **brass hadronic calorimeters** that provide precise energy measurements and locations of the corresponding deposits.
- The calorimeters and tracker are encapsulated by a **superconducting solenoid coil**, generating a magnetic field with a strength of 3.8 T, which bends the charged particle tracks and allows particle properties (trajectory, momentum, charge) to be deduced.
- A **steel return yoke** that surrounds the solenoid confines the magnetic field.
- Dedicated gas-ionisation **muon detectors** are placed on the outermost parts of the detector, in chambers embedded within the yoke, measuring their trajectory and momentum.
- The ECAL **silicon pre-shower** and the **steel forward calorimeter** are complementary detector systems in the forward regions, which target particles that travel in directions more parallel to the beam.

3.3.1 Geometry

CMS uses a right-handed coordinate system centred at the nominal *interaction point* (IP) [156, 136], as shown in Figure 3.7 (left): the x -axis points towards the centre of the LHC ring, the y -axis points vertically upwards, and the z -axis points along the anticlockwise beam direction. In order to consider the collisions in a symmetrical manner, it is common to use spherical coordinates where the azimuthal angle ϕ is used to indicate the direction in the x - y plane transverse to the beam, with $\phi = 0$ along the positive x -axis and $\phi = \pi/2$ along the positive y -axis. The polar angle θ is measured with respect to the beam axis z . Differences in rapidity are Lorentz invariant under boosts

along this axis, so the unit of pseudorapidity η is a preferable coordinate choice, which is defined as and shown in Figure 3.7 (right):

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) \quad (3.3)$$

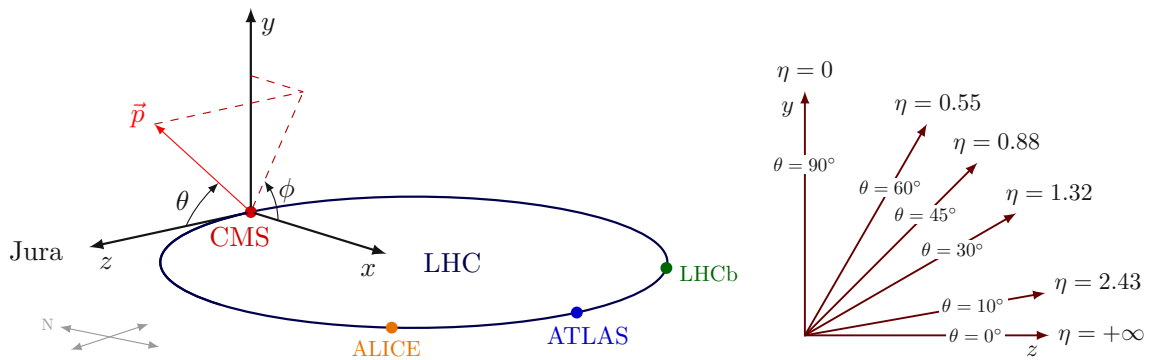


Figure 3.7: Diagram of the spherical coordinate system used by CMS, illustrating the corresponding azimuthal (ϕ) and polar (θ) angles for a particle with momentum \vec{p} , produced at the origin of CMS (left) and pseudorapidity η as defined in Equation 3.3 (right) [157].

Due to boost invariance and momentum conservation with respect to the z -direction, quantities are often expressed in terms of the projection of the four-vectors onto the x - y plane perpendicular to the beam, such as transverse momentum, p_T , or transverse energy, E_T . The CMS detector is sub-divided into a cylindrical *Barrel* region enclosed by two *Endcap* regions, providing an almost full coverage of the collision events. Correspondingly, each of the main detector subsystems is split into the two regions, as shown in Figure 3.8, which displays a detailed longitudinal view of one quadrant of CMS. The forward region is further away from the interaction point at very small η -angles. A summary of the pseudorapidity coverage of the different subsystems shown in Figure 3.8 is specified in Table 3.1.

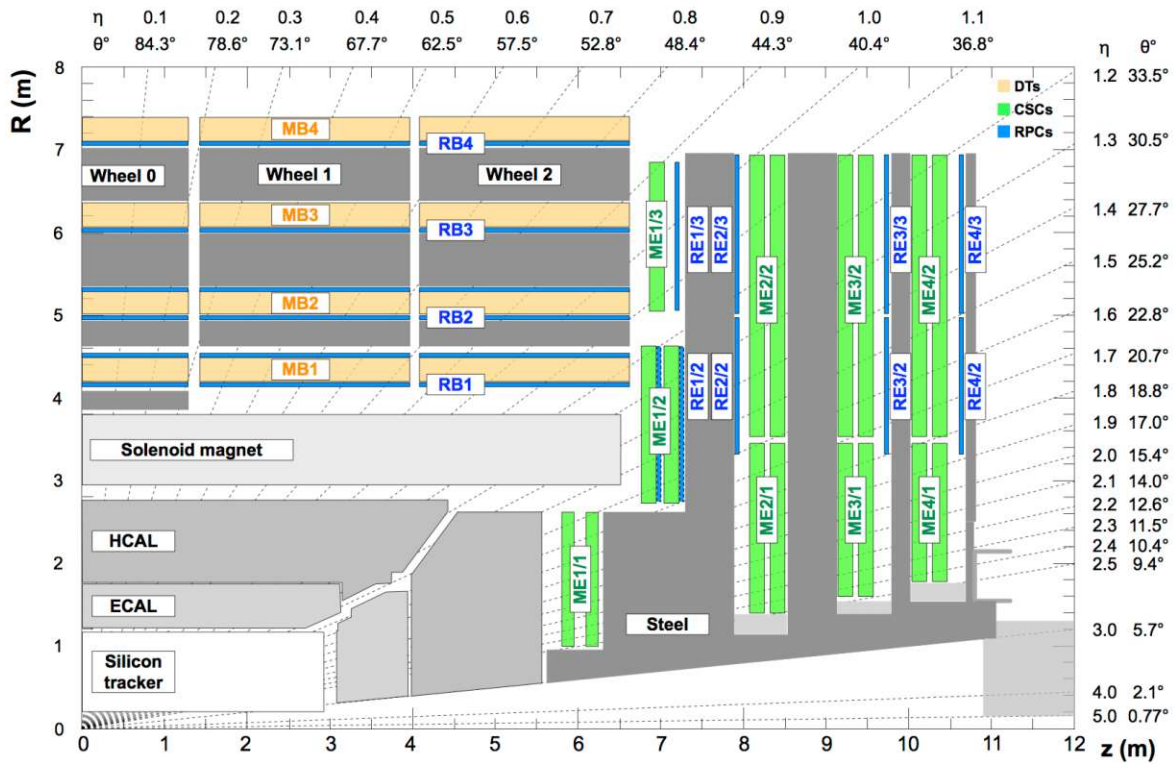


Figure 3.8: Longitudinal view of one quadrant of CMS displaying all the major subsystems: tracker, ECAL, HCAL, magnet, return yoke and muon detectors [158].

System	Coverage	Subsystem	Coverage
Tracker	$ \eta < 2.5$	Pixel	$ \eta < 2.5$
		Strips	$ \eta < 2.5$
ECAL	$ \eta < 3.0$	EB	$ \eta < 1.479$
		EE	$1.479 \leq \eta < 3.0$
		ES	$1.653 \leq \eta < 2.6$
HCAL	$ \eta < 5.2$	HB	$ \eta \leq 1.4$
		HE	$1.3 \leq \eta \leq 3.0$
		HO	$ \eta < 1.26$
		HF	$2.9 < \eta < 5.2$
Muon Detectors	$ \eta < 2.4$	DTs	$ \eta < 1.2$
		CSCs	$0.9 \leq \eta \leq 2.4$
		RPCs	$ \eta \leq 1.9$

Table 3.1: Pseudorapidity η segmentation of the main subsystems of CMS, including the tracker, ECAL, HCAL and muon detectors [136].

3.3.2 Magnet

One of the central elements of the CMS detector is the magnet, which is a large superconducting solenoid coil made from niobium and titanium, with a diameter of 6 m and a length of 12.5 m [159]. The magnetic coils are surrounded by a steel return yoke with a 14 m diameter and weighing 12 000 t, contributing to most of the detector's mass. The return yoke is composed of 5 three-layered dodecagonal wheels in the barrel and three disks at the endcaps. It has the role of making the field more homogenous in the tracker volume and confining the stray field on the outside. In addition, the steel plates play the role of absorber before the layers of muon chambers, allowing only muons through. Furthermore, the magnet structure serves as the detector's structural support, allowing it to withstand the forces of its own powerful magnetic field.

The purpose of the magnet is to bend the trajectories of charged particles (*tracks*), where the curvature is described by the Lorentz force. One can deduce different particle properties from the curvature of the tracks, such as momentum and charge, which are covered in more detail in the following Section 3.3.3. As in the case of high-energy colliders, the particles can be highly energetic ($\mathcal{O}(1 \text{ TeV})$ at a $\sqrt{s} = 13 \text{ TeV}$), and the magnetic field flux must be sufficiently strong to bend such trajectories within the detector's volume. The tracker and both electromagnetic and hadronic calorimeters are compact enough to fit inside the volume of the solenoid, minimising the amount of material in front of the calorimeters. A very strong magnetic field provides sufficient separation between tracks and calorimeter energy deposits left by charged and neutral particles, even at relatively low energies.

The modelled flux and field lines are shown in Figure 3.9. It is largely homogenous in the central detector volume with a central magnetic flux density of 3.8 T. The magnetic flux density can be measured in data precisely using muons originating from cosmic ray showers that pass through the detector in random locations, as opposed to originating from the interaction region [160]. This data is usually collected when there are no beams in the LHC. The measurements serve as important inputs for calibration, such as the alignment of the tracker or input to event reconstruction and simulations.

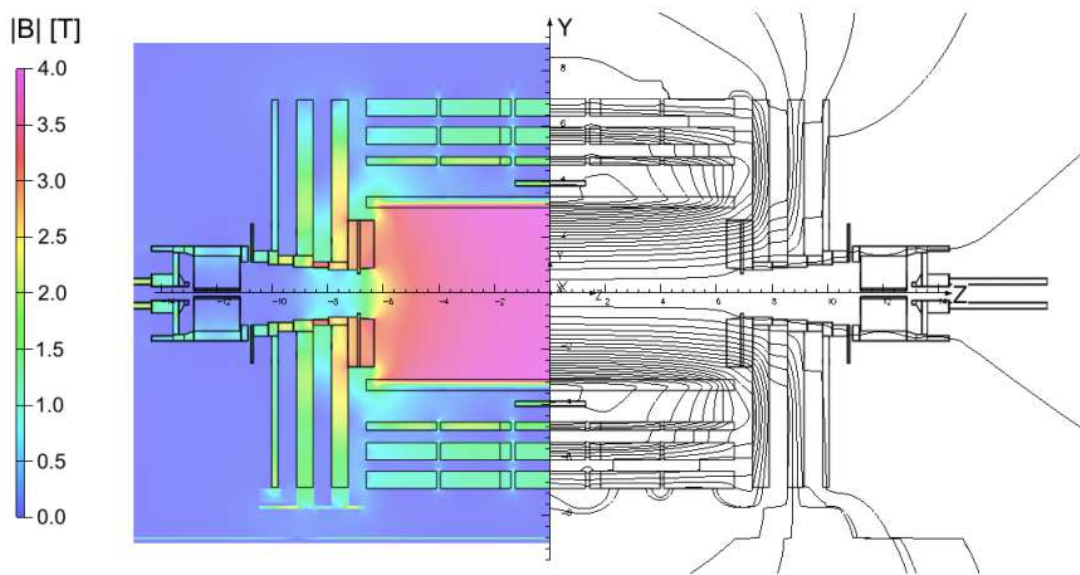


Figure 3.9: Magnetic flux density (left) and field lines (right) predicted on a longitudinal section of the CMS detector at a central magnetic flux density of 3.8 T. Each field line represents a magnetic flux increment of 6 Wb [160].

3.3.3 Tracker

The innermost part of the CMS detector is a silicon tracker [161, 162], consisting of an inner *pixel* detector surrounded by outer micro-strip sensors, both utilising semiconductor technologies. It is the largest all-silicon tracker ever built in the world. Its task is high-precision measurements of the trajectories, or *tracks*, of charged particles, which are curved when travelling through the strong magnetic field. Thereby, it also provides an independent measurement of the momenta of the particles with respect to the calorimeters. It is the closest detector element to the interaction point, between 4 cm and 1.2 m in radius and 5.6 m in length along the beam axis, where the outgoing particle flux and radiation are the highest. Therefore, its task is to reconstruct and disentangle all the tracks at the extremely high LHC luminosities and pileup (e.g. Figure 3.5), including displaced vertices of very short-lived particles. Correspondingly, it has to endure the high-radiation environment for longer time periods, motivating the choice of materials. Furthermore, the design is lightweight in terms of the materials used and the number of layers, in order to minimise the particle interaction with the

detector, such as multiple nuclear scattering, which can affect the original trajectories and energies of the particles.

The pixels and strips are singular semiconductor micro-sensors, which record the electrical current generated by electron-hole pairs in the silicon semiconductor material. This happens when a charged particle traverses the material and excites the electrons in the valence band to the conduction band. The micro-sensors are typically implanted into n-type semiconductor silicon bulk material with a p-type back side. Both sub-detectors are read out via a chain of analogue electronic and optical links that are able to transmit the absolute pulse height of the collected current for each sensor. Jointly mapping the positions in the sensors that the charged particle interacted with, known as *hits*, allows precise reconstruction of the tracks of the particles. Overall, in the barrel, there are 4 pixel layers and 10 micro-strip layers, while for each endcap, there are 3 layers of pixel detectors and 12 layers of strip detectors. The full tracker provides a minimum of 5 hits per track within a coverage up to $|\eta| < 2.5$. The objective of the tracker is to perform momentum, angle and position measurements with the highest resolutions, yielding a high reconstruction efficiency.

The schematic view of one quadrant of the tracker is shown in Figure 3.10.

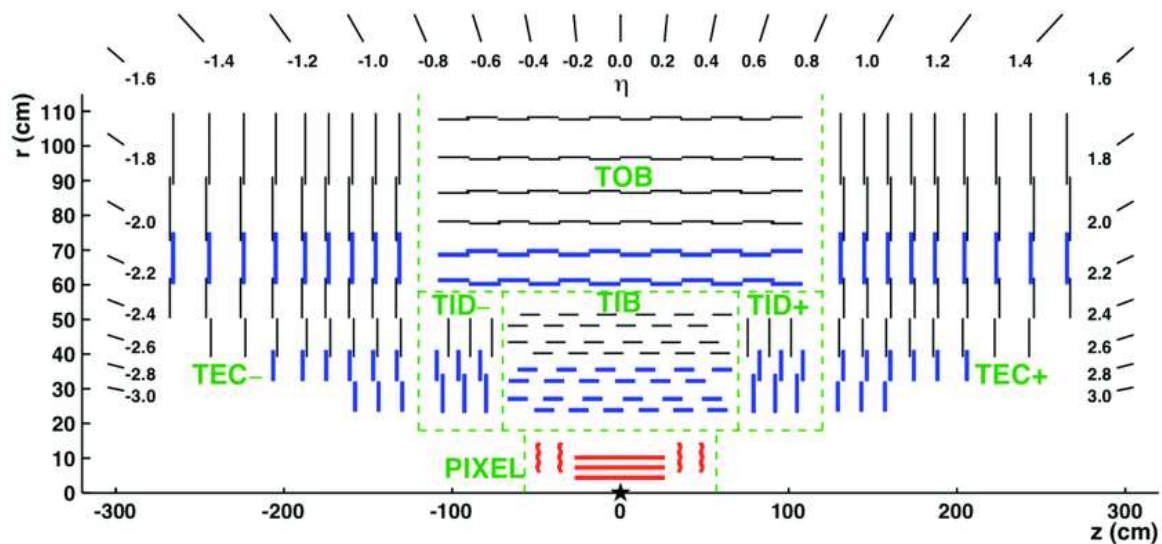


Figure 3.10: Schematic view of the silicon tracker in the r - z plane, where the pixel modules are located within the red area [163].

Pixel

The inner pixel detector is a silicon semiconductor device containing 124 million pixels arranged into 1 856 modules that are integrated into larger support structures. A module is made of a 285 μm thick n^+ -in-n type silicon sensor that comprises 66 560 pixels. Each pixel has a rectangular shape with an area of $100 \mu\text{m} \times 150 \mu\text{m}$, so each module covers an active area of $16.2 \times 64.8 \text{ mm}^2$, amounting to a total area of 1.9 m^2 of the detector. The modules are arranged in four barrel (BPIX) layers at 4.4 cm, 7.3 cm, 10.2 cm and 16 cm radius and three double-layers in the form of forward-backwards disks at the endcaps (FPIX). This arrangement is to ensure a four-hit coverage within an $|\eta| < 2.5$.

Due to their almost symmetric layout, the pixels provide two-dimensional spatial measurements. Therefore, three-dimensional tracks are reconstructed when taking into account the information from the different layers. The pixel detector is crucial in precise measurements of the track origin and, thus, the reconstruction of the primary (PVs) and secondary vertices (SVs) of long-lived particles.

Strips

The pixel detector is surrounded by the strip tracker consisting of a total of 9.6 million micro-strips covering approximately 206 m^2 total area. It comprises 15 148 modules containing one or two silicon strip sensors, each with 512 or 768 strips, distributed among the different detector areas, amounting to a total of 24 244 sensors. It has a much lower sensor density than the pixel, making it more cost-effective for the larger surface area that is covered, while maintaining high performance.

The micro-strips are narrow and long asymmetric elements extending across the full length of the sensors. The strips are 10-20 cm long and 80-180 μm wide. The silicon micro-strip modules are located in concentric layers around the pixel, with 10 layers in the barrel, of which 4 are inner double layers (TIB), and 6 are single outer layers

(TOB), reaching out to a radius of 1.3 m. Each endcap consists of 3 inner disks (TID) and 9 endcap disks (TEC), as shown in Figure 3.10. The strips are aligned parallel to the beam direction in the barrel and perpendicular in the endcaps. The asymmetric geometry of the strips provides only one-dimensional information on the location of the hit on the sensor. Nevertheless, the inner layers, as well as some of the outer ones, are equipped with stereo modules, which have two silicon sensors mounted back-to-back with their strips aligned at a relative angle of 100 mrad, allowing two-dimensional measurements.

3.3.4 Calorimeters

The combined CMS calorimeter system is designed to measure the total energy and direction of electromagnetically and strongly interacting particles, which deposit all their energy in the absorbers. The electromagnetic calorimeter (ECAL) is vital in the measurement of electron and photon energies. The hadronic calorimeter (HCAL) is necessary for the identification of quarks and gluons, which hadronise in the detector in the form of particle *jets*. Thus, it is important for the measurements of charged and neutral hadrons, such as pions, kaons, protons or neutrons. While the ECAL is also useful in the identification of charged particles in jets, the HCAL is also used in the identification of electrons, photons and muons. The calorimeters surround the inner tracker and are compact enough to fit inside the magnet volume. Therefore, the bending of trajectories of charged particles allows the energy deposits to have a defined separation, which is a very powerful feature in distinguishing different particles. The calorimeters are also particularly important for the measurement of apparent missing momentum induced by neutrinos or exotic particles which do not interact with the detector.

3.3.4.1 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) system comprises a central barrel (EB) and two endcaps (EE), with coverage up to $|\eta| < 1.479$ and $1.479 \leq |\eta| < 3.0$, respectively. The endcaps are complemented by a silicon strip preshower (ES) detector with higher

granularity to provide additional background rejection in the region $1.653 < |\eta| < 2.6$. A schematic of one quarter of the ECAL system is shown in Figure 3.11.

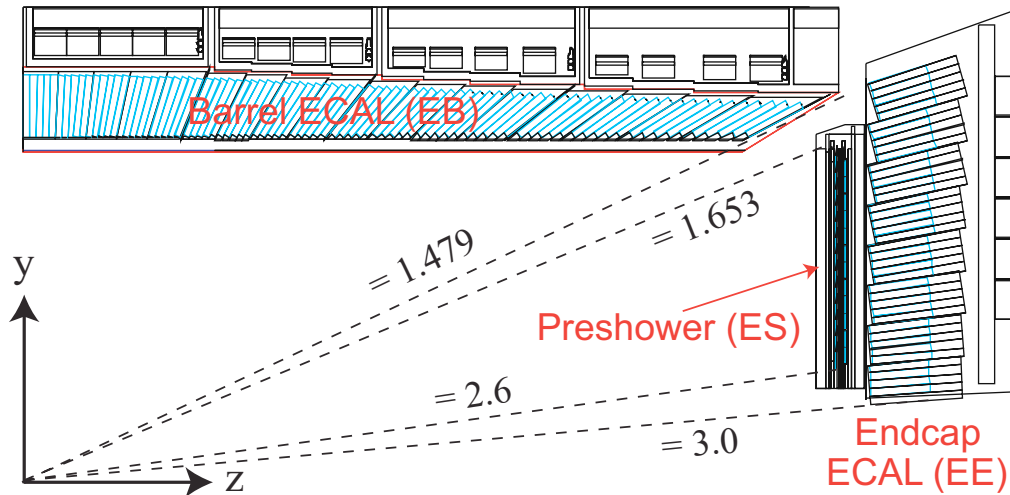


Figure 3.11: Longitudinal view of one quadrant of the electromagnetic calorimeter ECAL, indicating the positions of the central barrel (EB) and endcap (EE) and preshower (ES) [164].

Central ECAL

The central electromagnetic calorimeter (ECAL) [165] is a hermetic and homogenous lead tungstate (PbWO_4) crystal calorimeter. Its function is the precise measurement of the energy deposits of particles that interact via the electromagnetic interaction, primarily electrons and photons or charged particles in jets, such as charged hadrons. The ECAL comprises 75 848 crystals, providing a high granularity of measurements over $|\eta| < 3.0$. The central barrel region (EB) contains 61 200 trapezoidal crystals arranged into 5×2 submodules. They are assembled into larger 85×20 modules in 36 supermodules containing 1 700 crystals each, providing coverage up to $|\eta| < 1.479$. Both endcaps (EE) consist of two Dees with 3 662 crystals each, arranged into 138 5×5 supercrystals and 18 partial supercrystals, providing coverage between $1.479 \leq |\eta| < 3.0$.

Lead tungstate is an extremely dense but optically transparent material [166]. High energy electrons and photons traversing the crystals induce electromagnetic showers. These are cascades of particles where the electrons and photons undergo secondary emissions via bremsstrahlung and photon conversion (pair-production), respectively. The crystal material has a small Molière radius $R_M = 2.19$ cm, which characterises the induced electromagnetic shower size, and a short radiation length of $X_0 = 0.85$ cm. The dimensions of the crystals are $2.2 \times 2.2 \times 23$ cm in the EB and $2.86 \times 2.86 \times 22$ cm in the EE. Therefore, the R_M and X_0 are comparable to the crystal dimensions. The length of the crystals corresponds to $\approx 26 X_0$, which allows most of the electromagnetic shower energy to be deposited within the crystals. The properties of lead tungstate enable the ECAL to have excellent granularity as well as timing and energy resolutions. Interacting particles dissipate their energy in the crystals by exciting the atomic electrons, and scintillating light is re-emitted in the form of well-defined photon showers. The scintillation response is very rapid, where approximately 80 % of the light is emitted within 25 ns, which was an essential requirement respective to the high LHC bunch crossing frequency [166]. In the EB, the light pulses are read out by silicon Avalanche Photodiodes (APDs) connected to each crystal [165]. Inside the APDs, the light pulses are emitted as photoelectrons, where they are multiplied as an avalanche via the process of impact ionisation. The resulting current is proportional to the incident light pulses with an amplification factor, or *gain*, of ≈ 50 . The endcaps use Vacuum Phototriodes (VPTs) with a gain ≈ 10 , as they are more radiation resistant than APDs, which is necessary for harsher radiation at higher pseudorapidities.

Lead tungstate is intrinsically radiation hard, however, its response still changes with radiation due to the formation of colour centres that absorb the light and reduce the transparency [165]. Radiation damage can be reversed at room temperature via the process of thermal annealing, where the atoms regain their original structure. Typically, annealing takes several hours, and it occurs during breaks with no beam in the LHC cycles. During operation, the optical transmission is monitored via a laser monitoring system, which injects light pulses through each crystal via optical quartz fibres.

This gives a measure of the damage and recovery of the crystals, and the results are used to apply corrections both online in the trigger and offline analysis.

Electron and photon showers typically deposit their energy across several crystals, where approximately 94 % of the incident energy is contained in 3×3 crystal arrays and 97 % in 5×5 arrays [156]. Summing the energy measured in these fixed arrays gives the best performance for isolated photons or electrons. Before reaching the ECAL, electrons and photons can interact with tracker material, resulting in electromagnetic showers, in which the shower particles are spread along in the azimuthal direction ϕ by the strong magnetic field, thereby degrading the energy resolution.

ECAL Preshower

Each ECAL endcap (EE) calorimeter is preceded by a silicon strip preshower (ES) detector [165], which serves to provide additional background rejection. The principal target is neutral pions π^0 decaying into two closely-spaced lower energy photons, which can mimic a signal of high-energy photons. In the endcap regions, the angle between two photons from the π^0 decay is relatively small, so such backgrounds are more common. The ES is composed of 2 layers of lead absorbers interleaved by orthogonal silicon strip sensors and covers the pseudorapidity region $1.653 < |\eta| < 2.6$, shrouding most of the crystal endcaps, as demonstrated in the schematic layout of the ECAL in Figure 3.11. The two layers provide two measurements of the positions of the corresponding *hits*. High energy photons in the ECAL have their trajectories extrapolated to the ES, with the energy deposits summed over the two subsystems. The higher granularity with respect to the ECAL enables the separation of such signals. The background rejection is important in the context of the Higgs decay to two photons with high energies ($H \rightarrow \gamma\gamma$), which was one of the golden discovery channels. The ES also helps with the identification of electrons against minimum-ionising particles and improves the position determination of electrons and photons with high granularity.

3.3.4.2 Hadronic Calorimeter

The hadronic calorimeter (HCAL) [167] is a calorimeter with the purpose of identifying hadronising quarks and gluons by measuring the energy and direction of jets. Thus, it plays an essential role in performing precise energy measurements of charged and neutral hadrons, such as pions, kaons, protons or neutrons. It is also used in the identification of electrons, photons and muons in conjunction with the electromagnetic calorimeter and the muon systems. Furthermore, its hermeticity is vital in the determination of the missing momentum.

Since the HCAL is the outermost calorimeter that surrounds the ECAL, it is important to ensure that it is hermetic, so that all of the energy of interacting particles is deposited within its volume. As shown in Figure 3.12, the central hadron calorimeter is divided into barrel (HB) and endcap (HE) subsystems, complemented with an outer calorimeter (HO) located outside the solenoid coil, inside the return yoke. A separate forward calorimeter (HF) is located 6 m behind the HE to extend the hermeticity to higher pseudorapidities.

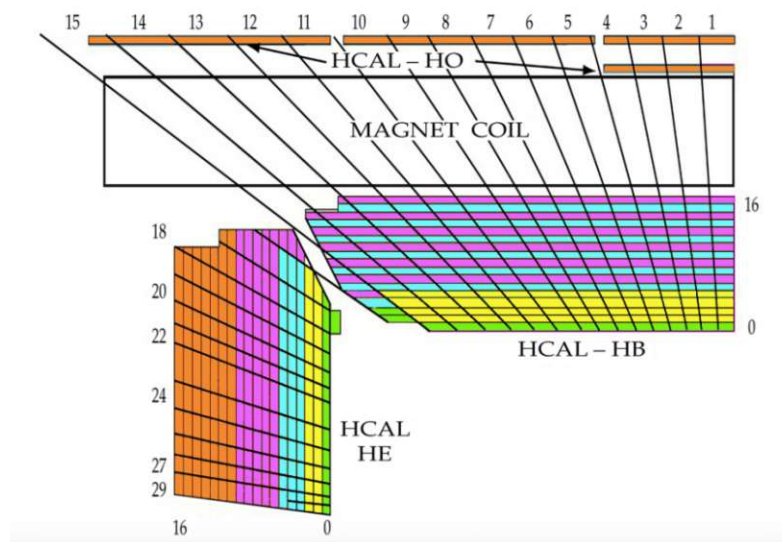


Figure 3.12: Schematic view of one quarter of the HCAL for the *Phase I* upgrade in the r - z plane showing the locations of the HB, HE, HO, and HF calorimeters. The numbers on top and left refer to segments in η , and the numbers on the right and bottom refer to scintillator layers. The colours indicate the combinations of layers that form the different depth segments, numbered sequentially from the IP. [168]

Central HCAL

The central calorimeters (HB, HE) [167] are located in the cryostat of the magnet and provide coverage up to $|\eta| < 3.0$: HB covers the pseudorapidity range $0 \leq |\eta| \leq 1.4$ and HE covers $1.3 \leq |\eta| \leq 3.0$, with an overlap region shared by both. They are sampling calorimeters made of alternating layers of cartridge brass plates (70 % Cu and 30 % Zn) as a dense absorber material, and tiles of plastic scintillator, as the active material. The absorber consists of a 40 mm thick front steel plate, followed by eight 50.5 mm thick brass plates, six 56.5 mm thick brass plates and a 75 mm thick steel back plate. The flat brass absorber plates are arranged into structures called *wedges* with the scintillators in between the plates. The radiation length of the absorber is $X_0 = 1.49$ cm, and the nuclear interaction length is $\lambda_I = 16.42$ cm, which is relatively short and motivated the choice of brass in order to comply with the compact design of CMS. The HCAL was designed to maximise the number of interaction lengths inside the magnetic coil. At $\eta = 0$, corresponding to 90° , the HB is ≈ 1 m thick and covers $5.82 \lambda_I$, while the ECAL adds $1.1 \lambda_I$; The effective thickness increases with polar angle and is $10.6 \lambda_I$ at $|\eta| = 1.3$, which contains more than 99 % of hadronic cascades.

The active material between the plates consists of $100 \times 100 \times 3.7$ mm³ plastic scintillator tiles made from a material named Kuraray (SCSN81), while the front and back steel plates are preceded with a 9 mm thick scintillator, called Bicron (BC408) [167]. The scintillator tiles are painted white along the narrow edges to reflect and contain the light. They are grouped together into single units called *megatiles*, with sizes dependent on their spatial location and orientation relative to the collision. The megatiles of a given azimuthal section and depth layer are grouped with the read-out fibres into a single trapezoidal-shaped scintillator unit, referred to as *trays*, that are inserted into the wedges.

The central (HB and HE) and forward (HF) calorimeters are each built of 36 identical azimuthal wedges, with every wedge covering 20° in ϕ . The HB is divided into two half-barrel sections (HB⁺ and HB⁻) which, along with the two endcaps (HE⁺ and HE⁻),

comprise 18 wedges each. The HB plates are oriented parallel to the beam axis, whereas the HE disks are placed perpendicular to the beam direction.

The HCAL is segmented into individual calorimeter cells in terms of η , ϕ and longitudinal depth layers that are represented by an integer. The HCAL is divided into 16 η sectors. In the barrel, every wedge is further split into four 5° megatile sectors in ϕ , whereas in the endcaps they are sub-divided into two 10° megatile sectors. This corresponds to an angular segmentation of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ in the HB and $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$ in the HE for $|\eta| \leq 1.6$. Ultimately, this defines the size of the megatiles and thus the granularity of the HCAL, which is chosen to match the size of the 5×5 ECAL crystals.

Most cells include several megatile layers grouped into longitudinal depth segments, represented by the different colours in Figure 3.12. The HB comprises 17 scintillator layers with 108 megatiles per layer, whereas both HE calorimeters consist of 19 layers amounting to a total of 1 368 megatiles. This amounts to a total of $\approx 70\,000$ and $20\,916$ scintillator tiles in the HB and HE, respectively. Towers correspond to optical read-outs from successive layers that are summed longitudinally across the depth segments. For example, in the HB calorimeter, most towers have a single longitudinal read-out, whereas the towers surrounding the endcap transition region are further segmented in depth with two or three longitudinal read-outs, motivated by the harsher radiation environment. The HB calorimeter has 2304 ($4 \times 16 \times 36$) towers, corresponding to the number of cells in a layer, and the total number of HCAL towers is $\approx 4\,300$.

Hadronic particles interact with the layers of the absorber to induce hadronic showers, which are a series of inelastic hadronic interactions with the nuclei of the target material that produce cascades of secondary particles. Neutral mesons can decay into photons and initiate electromagnetic showers within the hadronic shower. The shower development scales with the nuclear absorption length λ_l , which necessitates a large thickness of the calorimeter. The shower particles lose their energy due to ionisation and excitation of atoms until the energy falls below a threshold for nuclear interactions to occur.

The resulting secondary particles interact with the active material layers to produce scintillation light. The scintillation light is extracted by wavelength-shifting (WLS) fibres embedded in the scintillator tiles and channelled to photodetectors via clear fibre-optic waveguides that fit into grooves cut into the individual tiles [167]. The *Phase I* upgrade of the HCAL [145, 168] during Run 2 involved an upgrade of the readout, replacing original Hybrid Photodiodes (HPDs) in favour of Silicon Photomultipliers (SiPMs). They have many advantages, such as improved tolerance to the strong magnetic field, ≈ 2.5 times the photon detection efficiency and 200 times the gain with respect to the HPDs. The SiPMs also provide a 350 % increase in the number of readout channels, allowing a finer longitudinal depth segmentation in the HE and HB as shown in Figure 3.12. This allows a better measurement of the shower-development, an improved calibration to correct for depth-dependent radiation damage as well as the mitigation of pileup effects.

Outer HCAL

Due to the limited space between the ECAL and solenoid, the central barrel HB calorimeter is not able to fully contain highly energetic hadronic showers. Therefore, it is complemented by an outer hadronic calorimeter (HO) [167, 169], or so-called tail-catcher, which is located outside the solenoid coil, as shown in Figure 3.12. With the magnet coil and cryostat, and the steel of the magnet return yoke as an absorber, this adds a sampling depth of approximately $3 \lambda l$ of thickness in the region $|\eta| < 1.26$, which amounts to $\approx 11 \lambda l$. This enables the identification and measurement of any leaks out of the back of the HB caused by highly energetic or late-starting showers, thereby improving the energy measurement of jets and missing transverse momentum.

Forward HCAL

In order to provide coverage up to higher pseudorapidities between $2.9 < |\eta| < 5.2$, two forward hadronic calorimeters (HF \pm) [167, 170] are located 6 m behind the HE \pm and 11.15 m from the interaction point. The HF calorimeter is a Cherenkov calorimeter consisting of a steel absorber and embedded acrylic-clad optical fibres with a quartz core as the active elements. Due to the increased amount of radiation dosage at high pseudorapidities, the HF had to be constructed from extremely radiation-hard materials compared to the other HCAL subsystems.

Similarly to the central calorimeters, the HF is built of 36 identical azimuthal wedges, each covering 20° in ϕ , with cylinders comprising 18 wedges on either side of the detector. The fibres run parallel to the beam, spaced at 5 mm, and are bundled to form towers with granularity $\Delta\eta \times \Delta\phi = 0.175 \times 0.175$, matching that of the HE. The HF calorimeter consists of two units with an active radius with fibres 1.4 m and 1.65 m long, optimised for electromagnetic and hadronic showers, respectively.

Cherenkov light is generated when charged hadronic shower particles with energies above the Cherenkov threshold traverse the active fibres. Therefore the HF is mostly sensitive to the electromagnetic component of the hadronic showers. The signals from the fibres in each η - ϕ tower are combined, and the light is guided into the photomultiplier tubes (PMTs).

3.3.5 Muon Detectors

Muons, which are ≈ 207 times more massive than electrons, emit less bremsstrahlung and lose a lower fraction of their energy when traversing matter. Therefore, they are much more penetrating than other types of particles, and the calorimeters are not able to stop them. Thus, special muon detectors are placed on the outermost layers of CMS with the purpose of measuring their momentum, location and charge [171].

The muon detectors are placed in independent modules called *chambers* that are embedded between the layers and return yoke outside the magnet coil, covering the re-

gion $|\eta| < 2.4$. They are assembled within four layers of muon detectors called *stations* around a fixed radius or z , in the barrel or endcaps, respectively. The trajectory of muons in the strong magnetic field is determined by measuring the position in multiple layers of each station. This provides a measurement of the muon momentum independent of the inner tracker, which combined result in a very efficient muon identification system.

The muon system uses three types of gaseous ionisation detectors: drift tubes (DTs), cathode strip chambers (CSCs) and resistive plate chambers (RPCs). In total, there are 1 400 muon chambers in CMS, of which 250 are DTs arranged in concentric cylinders in the barrel region ($|\eta| < 1.2$) and 540 are CSCs in the endcap disks ($0.9 \leq |\eta| \leq 2.4$). In addition, 610 RPCs are interspersed in both regions ($|\eta| < 1.9$) to provide redundant triggering signals. A schematic view of the muon system in the x - y plane can be seen in Figure 3.8 in Section 3.3.1. In the context of the muon system, three regions are referred to as the barrel ($|\eta| < 0.9$), endcap ($1.2 < |\eta| < 2.4$) and overlap ($0.9 \leq |\eta| < 1.2$), as indicated in Table 3.1, where, in the latter, muons are detected by both the DTs and CSCs,

The performance of the muon system can be summarised by a reconstructed hit efficiency of 94-99 %, a segment timing resolution < 3 ns, a segment efficiency of 97 % and a trigger bunch crossing identification > 99 % [173].

Drift Tubes

In the barrel region, in between the return yoke segments, the local magnetic field strength is relatively low and approximately uniform, as shown in Figure 3.9 in Section 3.3.2. Furthermore, the expected muon rates and neutron-induced backgrounds are low. This motivated the choice of drift chambers with standard rectangular drift cells as the muon detectors [171], as shown in Figure 3.13 (top left). Every tube cell has a cross section of 42×13 mm² with a stretched 50- μ m diameter gold-plated stainless-steel anode wire at the centre. The cell volume is filled with a gas mixture of 85 %

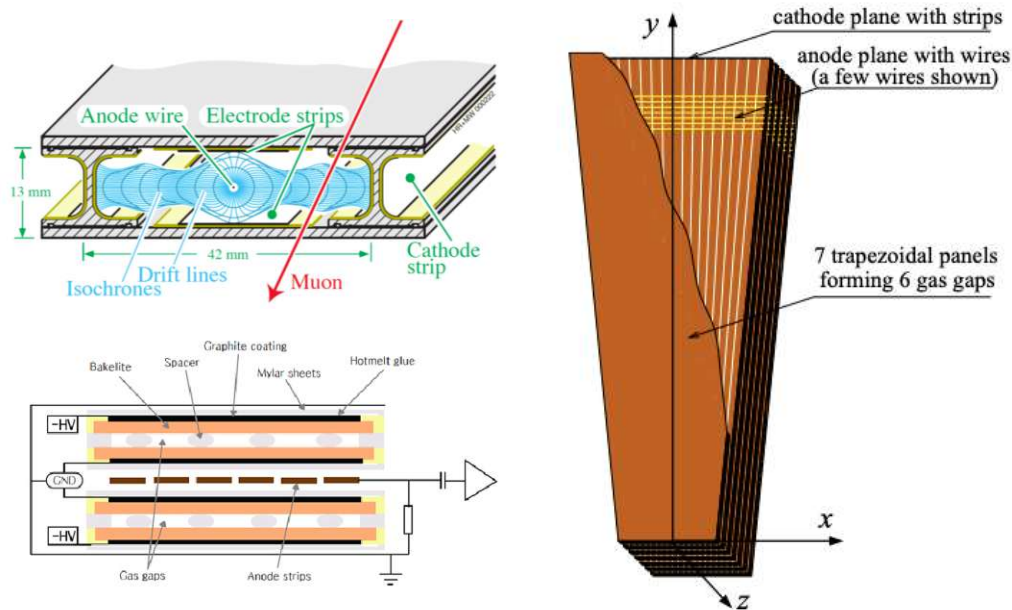


Figure 3.13: Schematic design of a (top left) Drift Tube (DT) [136], (bottom left) Resistive Plate Chamber (RPC) [172] and (right) Cathode Strip Chamber (CSC) [136].

Ar and 15 % CO_2 , which has quenching properties. Charged particles traversing the cell volume ionise the gaseous atoms, and the ionisation electrons are collected at the anode wire, resulting in a *hit* signal. The gas mixture saturates the drift velocity to about $55 \mu\text{m}/\text{ns}$, resulting in a maximum drift time of $\approx 400 \text{ ns}$. The positions of the hits can be deduced from timing measurements, assuming a constant drift velocity, providing a two-dimensional position measurement.

The DT chambers are interspersed among the layers of the return yoke plates, covering the region $|\eta| < 1.2$. They are organised into four concentric cylinders around the beam line corresponding to the stations *MB1-4* at different radii, as shown in Figures 3.14 and 3.8. The cylinders are sub-divided into 5 *wheels* ($0, \pm 1, \pm 2$) in the longitudinal direction along the beam line. Each wheel is organised into 12 ϕ -segments, corresponding to a segmentation of 30° , as shown in Figure 3.14.

The length of the chambers is defined by the longitudinal segmentation of $\approx 2.5 \text{ m}$, whereas their transverse width ranges from 1.9 m for *MB1* to 4.1 m for *MB4*. A super-

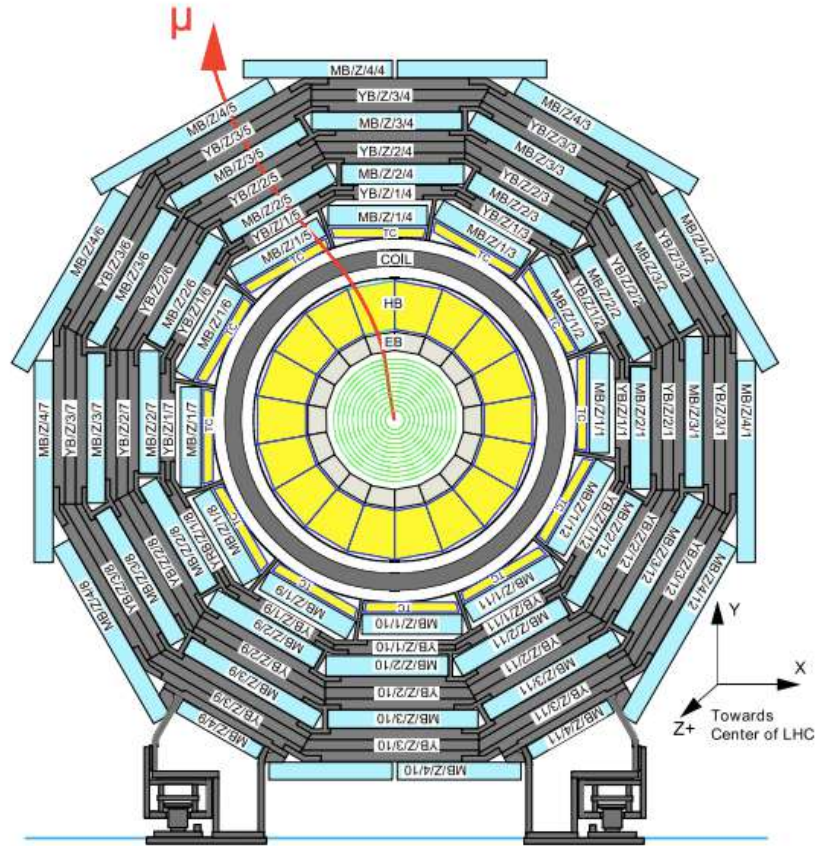


Figure 3.14: Layout of the CMS barrel muon DT chambers in one of the 5 wheels [136].

layer (SL) is formed from 4 staggered planes of drift cells. The chambers in the three innermost stations each contain a total of 3 SLs; Two SLs have the wires oriented parallel to the beam line measuring the r - ϕ coordinate in the transverse bending plane, while an inner SL has wires perpendicular to the beam providing an r - z or equivalently η measurement in the longitudinal direction. The chambers in the fourth station only have 2 SLs that perform r - ϕ measurements in the transverse direction. A SL provides excellent time-tagging capabilities, with a time resolution of a few nanoseconds, allowing for local, stand-alone and efficient bunch-crossing identification. The 3 inner stations have 60 DTs each, and the outer station has 70, amounting to 250 DTs with $\approx 172\,000$ sensitive wires, corresponding to the number of readout channels.

Cathode Strip Chambers

In the endcap regions, the radiation levels and muon rates are higher, while the magnetic field flux is non-uniform and stronger, as shown in Figure 3.9 in Section 3.3.2. This motivated the choice of cathode strip chambers (CSC), which can tolerate the non-uniformity of the magnetic field, have a fast response time due to a short drift path, provide finer segmentation and are radiation hard [171]. Apart from precise track measurements of muons, the fast response of CSCs additionally make them suitable for triggering.

CSCs are multi-wire proportional counters (MWPC) with a cathode strip readout that precisely measure the position at which a charged particle crosses the gas volume, as shown in Figure 3.13 (right). They comprise 6 anode wire planes interleaved among 7 cathode strips, with the wires and strips oriented approximately in an orthogonal fashion to each other. The volume of the strip chambers is filled with a gas mixture of 50 % CO₂, 40 % Ar, and 10 % CF₄. The layers of CSC measure the muon position in two coordinates: the wires run azimuthally and define a track's radial coordinate r , or equivalently η , while the strips are oriented radially to provide a measurement of the ϕ coordinate in the r - ϕ bending plane.

The CSC chambers are trapezoidal in shape and are mounted on the faces of the steel return yoke discs in the endcaps. They are oriented perpendicular to the beamline and cover the region $0.9 \leq |\eta| \leq 2.4$. A station in the endcaps is organised as a ring of chambers assembled between the layers of the return yoke at the same value of z along the beam line. There are 4 CSC stations in each endcap, labelled *ME1-ME4*, as shown in Figure 3.8. In the radial direction, stations are arranged in two or three rings of endcap CSCs. The inner rings of stations 2, 3, and 4 have CSCs that subtend a ϕ angle of 20° while all other CSCs subtend an angle of 10°. In total, there are 266 112 strip channels and 210 816 anode channels.

Resistive Plate Chambers

Resistive plate chambers (RPCs) are gaseous parallel-plate detectors that play the role of providing muon trigger signals parallel to the DTs and CSCs [171]. Similarly to the DTs and CSCs they measure the time and position of a muon hit with a good spatial resolution up to $30 \mu\text{m}$. Comparatively, they are high-speed detectors with an excellent time resolution of $\approx 2 \text{ ns}$, which is much less than the typical 25 ns LHC bunch spacing. Thus, they are capable of assigning a muon track to a particular bunch-crossing, even at very high luminosities. The trajectory is deduced from the pattern of hit strips, providing a measure of the muon momentum that is then used by the trigger.

RPCs consist of two parallel plates made from bakelite, a plastic laminate of a phenol resin, which has a very high resistivity of 10^{10} - $10^{11} \Omega\text{cm}$. They are covered with graphite paint to form the electrodes. The plates are separated by a 2 mm gap, and the volume is filled with a composition of 3 gases: 95.2% Freon ($\text{C}_2\text{H}_2\text{F}_4$) in order to enhance an ionisation, 4.5% Isobutane (iC_4H_{10}) used as a quencher and 0.3% SF_6 in order to control the background electrons. Incident charged particles ionise the gas, and the applied electric field causes an avalanche of electrons, which are collected by the anode. The avalanche mode allows the RPCs to operate for high rates of traversing particles. An RPC double-gap module consists of 2 gaps with common pick-up read-out strips in between, where the total induced signal is the sum of the two single-gap signals.

In the endcap region, there are 576 chambers, while in the barrel region, there are 480 chambers, amounting to a total of 1 056 RPCs. In the barrel, the RPC chambers form 6 concentric cylinders around the beam axis, arranged into 4 stations, denoted by *RBI-RB4*, as shown in Figure 3.8. The arrays of RPCs are located on the inner side of the DT stations, with the two first stations having additional external RPC layers. In the 4 endcap stations, each of the four CSCs stations has a corresponding RPC layer, labelled *RE1-RE4*. Each endcap RPC chamber consists of a double-gap structure enclosed in a flat trapezoidal-shaped box made of 2 aluminium honeycomb panels, as shown in Fig-

ure 3.13 (bottom left). The RPCs are oriented parallel to the DTs and CSCs to measure the coordinates in the bending plane. As part of the *Phase I* upgrade [145], during *LSI*, 144 additional High-Pressure Laminate (HPL) double-layer RPC chambers were added in the endcaps (*ME4*, *RE4*), extending the coverage to $|\eta| \leq 1.9$ and improving the efficiency. In total, there are 68 136 readout channels in the barrel and 55 296 channels in the endcaps.

3.3.6 Trigger and Data Acquisition

As discussed in Section 3.1.2, the LHC conditions correspond to a bunch crossing rate of 40 MHz and an instantaneous luminosity $\mathcal{L} \approx 2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. The DAQ and trigger systems are designed to analyse, filter and collect the collision data at these enormous rates. It is not feasible to read out and record every event due to hardware limitations, nor is it efficient in terms of the physics programme. Since the LHC is a proton-proton collider, the underlying interaction mechanisms are strong interactions, resulting in a significant fraction of events which are of lesser interest, such as soft elastic collisions or QCD multijet events from inelastic collisions. The scientific programme of CMS, covered in Section 3.2.1, focuses on more sophisticated signatures, such as those involving weakly-interacting particles. For this reason, high-energy experiments have dedicated trigger systems in place that serve as filters to select only the most interesting events for offline analysis.

CMS utilises a two-tiered trigger system, with the Level-1 Trigger (L1T) implemented in custom hardware, and the High Level Trigger (HLT) based on software [174, 164, 156, 175]. The L1T preselects the data and reduces the event rate to ≈ 100 kHz and subsequently, the HLT reduces the rate down to ≈ 2 kHz. The flexibility of the trigger systems is an important design consideration, as they need to adapt to the dynamic LHC conditions in order to maximise the efficiency of collecting data while being subjected to the imposed limits. The type of data that is preselected by a collection of filtering algorithms (*triggers*) inside dedicated trigger *menus* at both the L1T and HLT levels, ultimately encapsulating the entire physics programme of CMS.

The readout and filtering of the data are handled by the data-acquisition (DAQ) system [164] and sent off-site computing farm for event reconstruction, storage and analysis. In addition to collecting collision data, the trigger and DAQ systems record information for calibration and data quality monitoring (DQM) [176], performed during online data-taking.

3.3.6.1 Data Acquisition

All events that pass the LIT are sent to a computer farm which receives data from all the subsystems and builds and filters (via the HLT) the events. Therefore, the DAQ system must be able to sustain a given readout rate, known as the DAQ bandwidth or throughput:

$$B_{DAQ} = R_T^{max} \times S_E \quad (3.4)$$

which is determined by the trigger rate (R_T) and the event size (S_E). Therefore, the maximum allowed trigger rate (both at LIT and HLT) is defined by the capabilities of the DAQ system. For example, for an average event size of ≈ 1 MB, and an LIT (HLT) trigger rate of 100 kHz (1 kHz) results in a input (output) bandwidth of 100 GB/s (1 GB/s) that is transferred downstream [164]. The HLT output is still considerable and requires dedicated data handling methods for processing and storage.

The DAQ system underwent a full *Phase I* upgrade during *LS2* [145, 177] in order to accommodate the new detector readouts and handle the larger number of readout channels. The upgrade has extended the event size margin to over 2 MB and doubled the available aggregate throughput to 200 GB/s while increasing the amount of data that can be sent to storage even up to 30 GB/s for *Run 3*. The full schematic of the upgraded system used in *Run 2* can be seen in Figure 3.15.

An event enters the system as a set of fragments from the sub-detectors, which are read out in parallel into the Front-End Drivers (FEDs). There are ≈ 700 FED modules, each carrying between 1-8 kB of data per event. The data from the FEDs are sent via optical links (*Slink64/Slink-Express*) to the front-end readout link (FEROL) of the central DAQ system. The FEROLs convert the data stream to TCP/IP via optical links, provid-

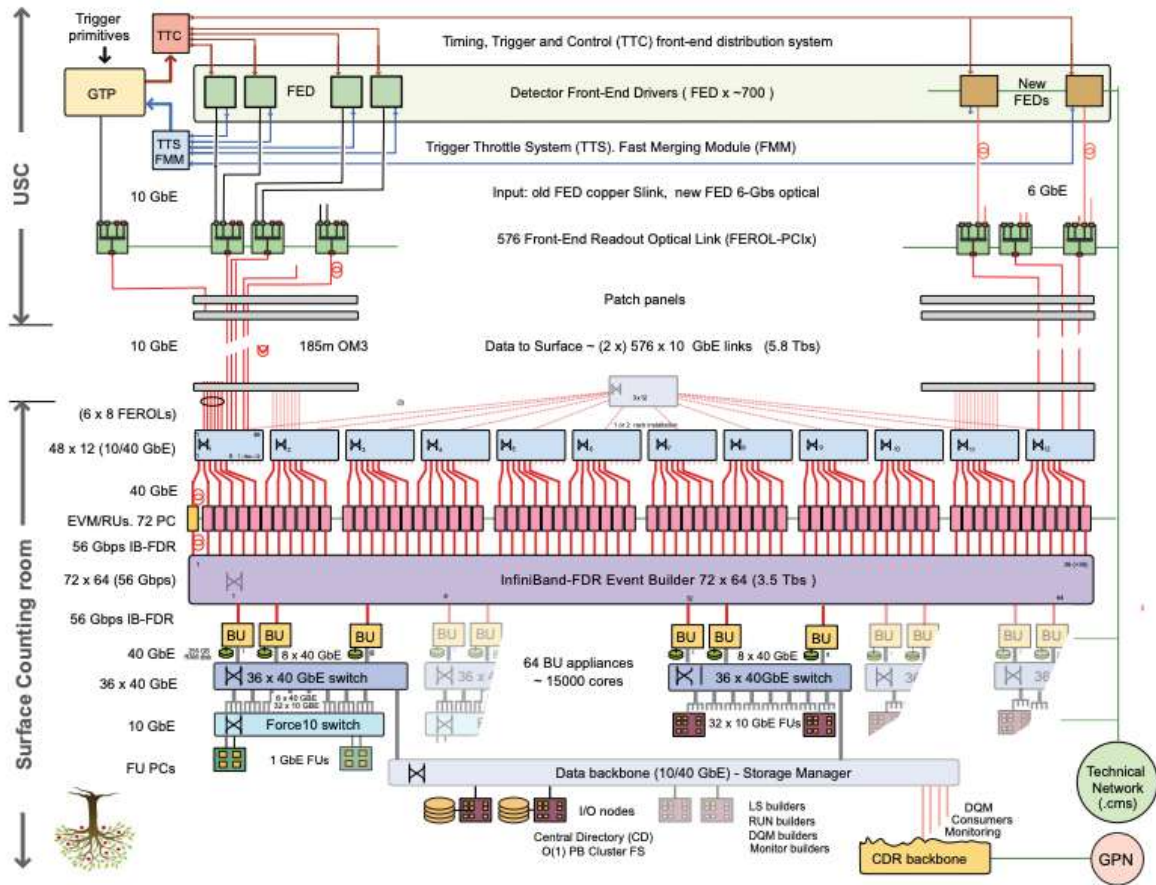


Figure 3.15: Schematic of the DAQ system architecture during LHC Run 2 [177].

ing 10/40 Gbit/s Ethernet. The Event Builder (EVB) collects the fragments, assembles them and sends the complete event to a single processing unit. Data from several FEROLs are routed to Readout Units (RUs), which read out and buffer the event fragments from the FEDs into *superfragments*. The RUs are connected via 56 Gb/s FDR Infiniband to Builder Units (BUs), which in parallel assemble the full events from the superfragments.

The HLT processing, as described in Section 3.3.6.3, is entirely file-based, which allows the DAQ and HLT systems to be independent. The HLT operates on events fully assembled by the EVB and hence can use all the detector data with full granularity and resolution. The HLT algorithms run on a large filter farm (*Event Filter*) of processing nodes called Filter Units (FUs). The FUs receive events from the BUs, where several

FUs are assigned to each BU. The FUs analyse the events made available by the corresponding BU to make the final trigger decision. Accepted events are written to files on a global cluster file system which are then transferred to the Tier-0 (T0) computing facility for full offline event reconstruction.

The T0 centre is part of the Worldwide LHC Computing Grid (WLCG) [178], which is a global grid-based computing resource for the storage, distribution and analysis of the data generated by the LHC. It combines the computing resources of about 900 000 computer cores from over 170 sites in 42 countries. The system has four tiers, starting with *T0* (CERN Data Centre), which is responsible for the safekeeping of the raw data and full offline event reconstruction. The data is then further distributed between 13 *Tier 1* centres, responsible for large-scale reprocessing and storing a share of raw, reconstructed and simulated data. Approximately 155 *Tier 2* centres provide storage and adequate computing power for specific analysis tasks, while also handling a proportional share of the production and reconstruction of simulated events. Finally, individual scientists can access the WLCG through local (or *Tier 3*) computing resources.

3.3.6.2 Level-1 Trigger

The Level-1 Trigger (L1T) is a hardware system with custom processors [174, 175, 179] with the purpose of preselecting the data arriving at the LHC at a rate of ≈ 30 MHz. The readout electronics can handle up to 100 kHz of events, making a decision whether to keep or discard an event within a fixed latency of $4 \mu\text{s}$. If an event is accepted, then the complete detector information is read out and sent downstream to the HLT.

In order to cope with the increased instantaneous luminosity and pileup conditions, the L1T trigger has undergone a major *Phase I* upgrade during *LS2*, replacing all of the Level-1 trigger hardware, cables, electronics boards, firmware and software, taking advantage of the more recent μTCA technology [145, 180].

The system combines information from the calorimeters and muon detectors [175]. The signals from both the ECAL and HCAL calorimeters and muon detectors (DTs,

CSCs, RPCs) are sent in the form of Trigger Primitives (TPs) and processed in several steps to give a basic description of the event. This includes standard particle candidates (or *trigger physics objects*) such as muons, electrons/photons, taus and jets, including basic properties such as their momentum p_T and angular direction (η , ϕ). A schematic diagram of the data flow between different subsystems comprising the LIT is shown in Figure 3.16. It is divided into two main streams: muons and calorimeters, and the combined event information is evaluated in the global trigger (μ GT) for a final decision on whether to accept an event. External binary signals coming from sub-detectors, such as the Beam Pickup Timing for the eXperiment (BPTX), HF or CTPPS-TOTEM, are also used for special triggers.

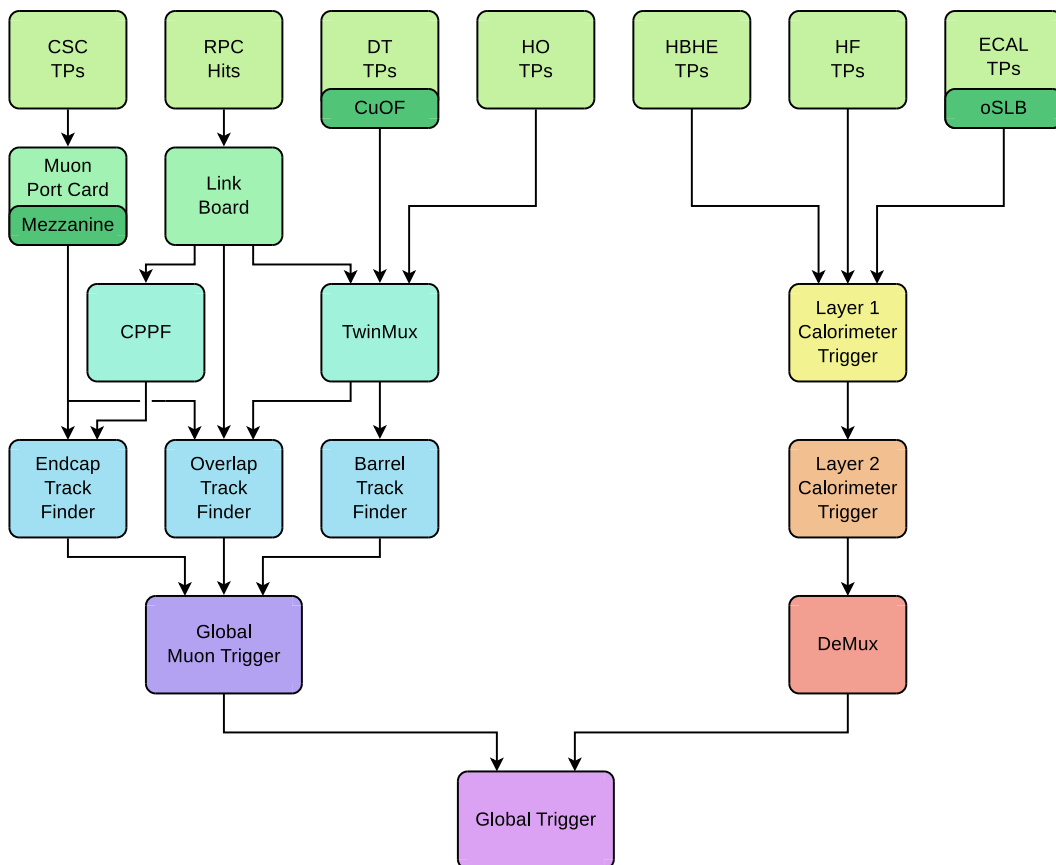


Figure 3.16: Schematic diagram of the subsystems comprising the Level-1 Trigger (LIT) [179].

LIT Muon Trigger

The track finders (TFs) use muon detector TPs to reconstruct muon track candidates with basic properties, such as coordinates, transverse momenta, timing and quality information. There are three muon TFs in the LIT:

- Barrel Muon Track Finder (BMTF) takes inputs from DTs and RPCs in the barrel.
- Overlap Muon Track Finder (OMTF) takes inputs from all three muon subsystems in the overlap region between the barrel and endcaps.
- Endcap Muon Track Finder (EMTF) takes inputs from CSCs and RPCs in the endcaps.

Each TF transmits up to 36 muons, totalling a maximum of 108, to the Global Muon Trigger (μ GMT), which resolves duplicates and sends the eight muons of the highest p_T and quality value [175, 180, 179].

LIT Calorimeter Trigger

In the calorimeters, the signals from ECAL crystals and HCAL towers are grouped into units called Trigger Towers (TTs) [175, 180, 179]. In the barrel, the 5×5 crystal arrays in the EB are grouped with the associated HB towers behind them into TTs with a granularity $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. In the endcaps (EE, HE and HF), the TTs are grouped in different ways with a granularity up to $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$. This amounts to 144 TTs in the central region and up to 54 TTs in the endcaps.

The calorimeter trigger system consists of two stages: *Layer-1* and *Layer-2*. At the first Layer-1 stage, the calorimeter trigger calculates the sum of local energy deposits from the ECAL and HCAL TPs, and calibrates the energy. At the second Layer-2 stage, the system uses the calibrated TPs to reconstruct and further calibrate the relevant physics

objects. Specific algorithms, such as dynamic clustering of energy deposits around a TT, are used to reconstruct electrons, photons, taus, hadron jets and energy sums [179]. Electrons and photons are indistinguishable to the LIT, so they are treated as single e/γ candidates. The τ algorithm aims to efficiently reconstruct τ leptons decaying to one, two, or three charged or neutral pions. The jet reconstruction algorithm uses a 9×9 TT sliding window to search for local maxima to define jet candidates. The transverse energy of the jets is defined as the scalar sum of all TT energies in the window centred on the candidate. Global event quantities such as the total scalar transverse energy of all jets, H_T , and the corresponding vector sum H_T^{miss} are computed using the reconstructed jets. Apart from jets, the total scalar sum of transverse energy over all TTs, E_T , and the corresponding vector sum, p_T^{miss} , are also computed and serve as a measure of the event's total and missing transverse momentum, respectively.

LIT Global Trigger

Ultimately, in the final stage, the Global Trigger (μ GT) combines information from the muon and calorimeter systems and generates a signal that triggers the full readout of the detector, sending it downstream to the HLT. The trigger decision is based on a programmable *menu* of algorithms, also referred to as *seeds*, which are a set of selection requirements applied to the final list of objects from calorimeter Layer-2 and μ GMT.

As part of the *Phase I* upgrade, the μ GT has been re-implemented on modern FPGAs on an Advanced Mezzanine Card (AMC) in a μ TCA crate [145, 180]. The algorithms are implemented on six MP7 (Master Processor board with Xilinx Virtex-7 FPGAs) boards, allowing up to 512 different seed algorithms to be used. A large number of possible algorithms and the re-programmable nature of FPGAs provide flexibility in modifying the menu contents.

From the technical perspective, Level-1 trigger algorithms or seeds are implemented using a Trigger Menu Editor (TME) to output an Extensible Markup Language (XML) file in a transparent format, with more details in Appendix C.5. The XML is passed

through a VHDL (VHSIC Hardware Description Language) producer to add specific metadata (e.g. UUID) and information about the module (board) distribution, for resource management. Finally, synthesis of the firmware converts the VHDL files to binary files, which are then uploaded to MP7 boards of the LIT Global Trigger (μ GT). Further details on the TME and LIT algorithms are provided in Appendix C.5.

Level-1 algorithms use information from the TPs to impose specific logical or topological requirements on the event, such as requirements on their momentum and angular acceptance, charge, quality and isolation. External binary signals coming from sub-detectors can also impose additional requirements, such as a Beam Pickup Timing for the eXperiment (BPTX) signal or a *MinimumBias* signal from the HF, which can indicate whether a bunch was filled during a bunch crossing event. More sophisticated multi-object triggers, also known as *cross-triggers*, are designed with logical combinations (AND, OR, NOT, XOR) that require coincidences of different trigger objects. Topological conditions can be applied to such triggers, such as the angular separation or requirements on the transverse or invariant masses between a pair of objects. Cross-triggers target more exclusive signatures and, by definition, have lower rates with respect to their less stringent counterparts. If the allocated rate is the same, this allows the relaxation of other conditions, such as the p_T threshold, which affects the overall acceptance of many signatures.

If an event is accepted, the complete detector information is read out and sent downstream to the HLT for a refined purity and selection.

3.3.6.3 High Level Trigger

The High Level Trigger (HLT) is implemented in software, reducing the preselected events by the LIT by another factor of 100, down to $\approx 1\text{-}2$ kHz. The final output rate reaches a sustainable level for reconstruction and storage, for subsequent offline analysis. It runs on a dedicated off-site farm of processors, including $\approx 26\,000$ CPU cores. It is only limited by the available processing power and the quality of the online calibration [164, 175]. The processing capacity of the filter farm was expanded gradually throughout *Run 2* to cope with the evolving LHC and detector conditions. A major upgrade for *Run 3* is the HLT is the inclusion of GPUs in every DAQ node, offloading a significant amount of the timing bandwidth [181], opening up a number of new possibilities.

The HLT has access to the complete detector information from the inner tracking system, calorimeters and muon detectors to perform a more precise reconstruction of physics objects with a set of complex reconstruction algorithms. The software uses the same framework used for offline reconstruction and analysis, with algorithms optimised for fast performance. They impose more stringent identification and quality criteria with respect to the LIT while ensuring a high reconstruction efficiency, purity and acceptance.

A significant reduction in event rate is achieved by utilising track and vertex information, which improves background rejection. Central to many HLT reconstruction algorithms is the particle flow (PF) algorithm (discussed in Section 4.1), which uses information from all detectors to identify candidates for muons, electrons, photons and charged and neutral hadrons, to paint a global picture of the event. In terms of jet reconstruction, the tagging of jets originating from b-quarks is also possible with algorithms that search for secondary vertices (CSV), including ones exploiting deep neural networks (DNNs) [182, 183, 184], which are described in Section 4.1.

An HLT trigger algorithm is called a *path*, which is a set of algorithmic processing steps run in a predefined order that both reconstructs physics objects and filters on these objects based on specific requirements. From the framework point of view, it is a col-

lection of specific reconstruction and filtering modules with a standard structure of increasing complexity. The first filter of an HLT path is a selection based on the decision of the corresponding L1T algorithm, which means that every HLT path is seeded by a dedicated L1T seed. This is followed by some initial reconstruction and filtering based on calorimeters and muon detectors. More complicated and processing-heavy reconstruction algorithms, such as PF reconstruction, are only run at a later stage. Intensive track reconstruction is performed in an iterative manner, with scaling degrees of stringency on the requirements of reconstructed hits in the tracker (as discussed in 4). Similarly, as with the L1T, cross-triggers with specific topological requirements are implemented.

The time and output rate of the online selection are defined by the computing power of the processor farm, which has a timing (execution time) budget of ≈ 450 ms per event. In order to comply with the limits of the computational resources, HLT algorithms are designed with fast reconstruction and filtering modules with algorithms run in optimised sequences. If at any point in the sequence the requirements of a filter are not met, the event is discarded. HLT paths selecting similar topologies are grouped into primary *datasets* for subsequent offline processing, which are associated with specific data formats into *streams* for efficient data handling by the DAQ system. Moreover, streams of data with limited event content are also saved for data quality monitoring (DQM), calibration or offline trigger studies. There are also special data streams that require less bandwidth, allowing events with objects with lower acceptance thresholds to be collected. The technique of *scouting* stores reduced event content, which can be used for analyses that do not require all the detector information; Figure 3.17 shows dimuon ($p_T > 3$ GeV) resonances reconstructed directly at HLT from the scouting stream. The process of data *parking* uses fewer computing resources by storing the data in a raw format to be fully reconstructed at a later time.

The software is based on CMS Software (CMSSW), which is the same software framework used for offline reconstruction and analysis, described in Section 6.2. The framework supports multi-threaded event processing, which optimises memory usage so

that HLT algorithms are optimised for fast performance. ConfDB (Configuration Database) is a dedicated database software system that tracks the contents of the paths, sequences, modules and their corresponding parameters in a given configuration. They are managed and modified using a special graphical user interface (GUI). An example of an HLT algorithm as viewed from ConfDB is shown in Figure A.II. A converter then transforms this information into a single *python* configuration file that is then run directly on the HLT filter farm. Further details on the ConfDB GUI and HLT algorithms are provided in Appendix C.5.

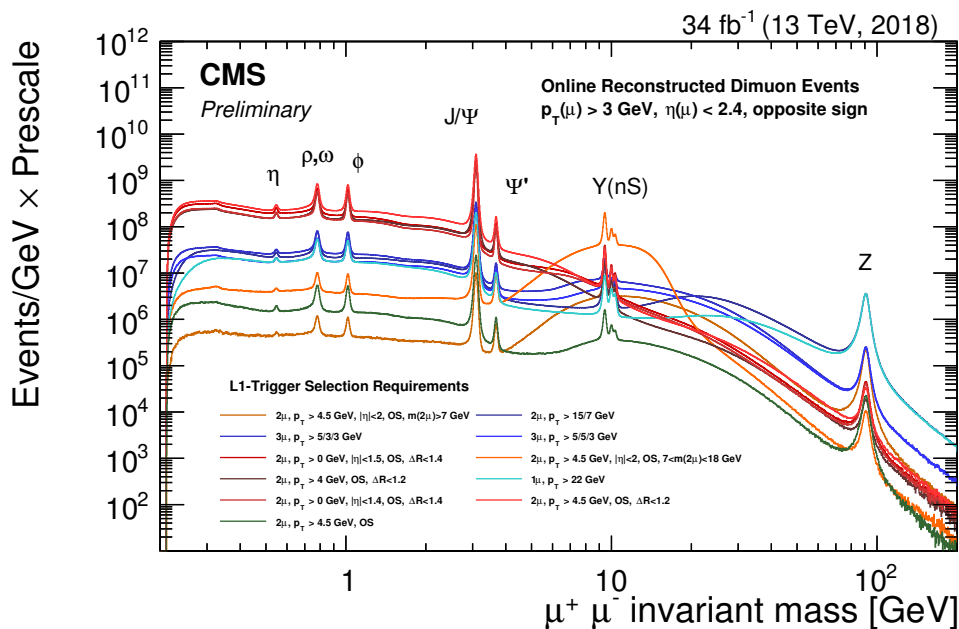


Figure 3.17: dimuon invariant mass spectra reconstructed in the High Level Trigger system of the CMS detector for various muon Level-1 trigger requirements deployed by the CMS collaboration in 2018. Well-known dimuon resonances from meson or Z boson decays (e.g. J/ψ) are indicated. [185]

3.3.6.4 Trigger System Performance

In order to receive a better overview of the trigger systems, this section presents a select number of performance measurements carried out by CMS during LHC *Run 2*, including rates and efficiencies, as well as timing.

Level-1 Trigger

The content of the trigger menus can be expressed in terms of the allocated bandwidth for different trigger group types, as shown in Figure 3.18 for the LIT menu.

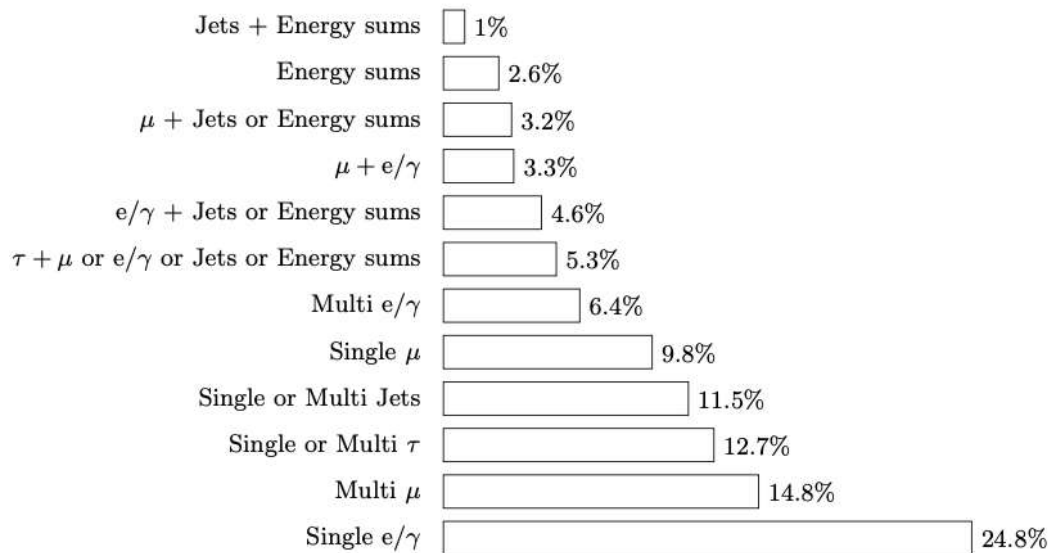


Figure 3.18: LIT bandwidth allocation for single- and multi-object triggers and cross triggers in a typical p-p physics menu [179].

The trigger rate is strongly dependent on the LHC conditions, specifically instantaneous luminosity \mathcal{L} and pileup (PU). The rate versus PU dependence of some benchmark seeds from the core of the menu is shown in Figure 3.19. One can see that for leptons, the dependence is rather linear and not as strong as in the case for single jets and p_T^{miss} (ETMHF), which has an exponential dependence. The mitigation of the high PU-dependence of certain trigger objects such as energy sums (e.g. p_T^{miss}) is one of the priorities of the trigger study groups.

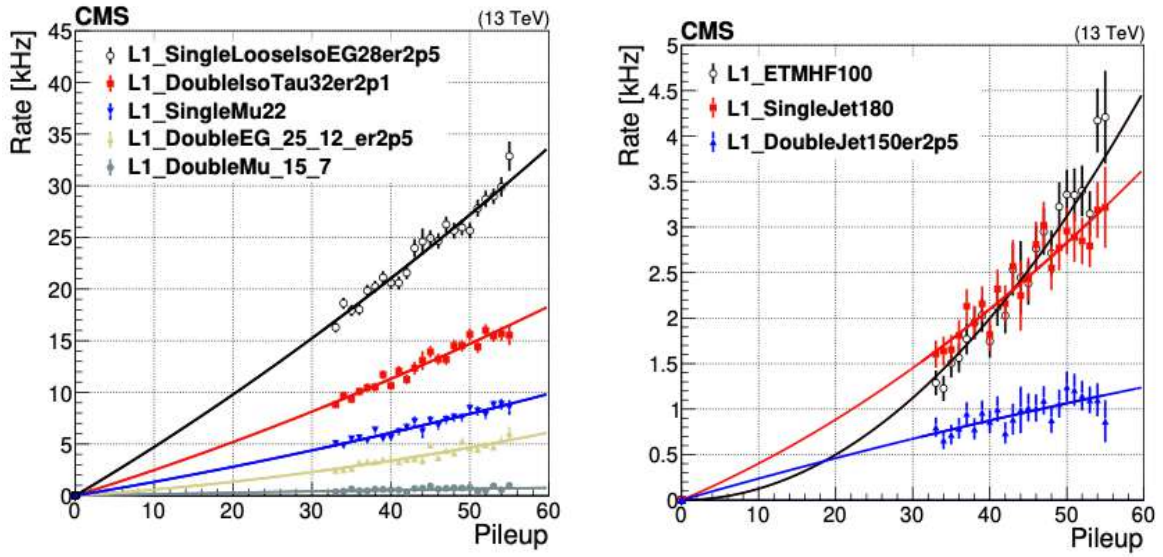


Figure 3.19: Rate vs pileup dependence for a selection of LIT single- and multi-object triggers and cross triggers [179].

A *prescale* is a number that defines what fraction of triggered events should be read out. Triggers that are not prescaled (unprescaled) are typically the ones that are used for analysis and form the core of the menus. The lowest unprescaled thresholds on benchmark single- and double-object seeds and their approximate rates at $\mathcal{L} = 1.8 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ (PU ≈ 50) are shown in Table 3.2.

Algorithm Name	Description	Thresholds	Total Rate (kHz)
L1_SingleMu22	Single muon	$p_T > 22 \text{ GeV}$	8
L1_SingleLooseIsoEG28er2p5	Single loosely isolated e/γ	$p_T > 28 \text{ GeV}, \eta < 2.5$	27
L1_SingleJet180	Single jet	$p_T > 180 \text{ GeV}$	3
L1_HTT360er	H_T	$H_T > 360 \text{ GeV}$	4
L1_ETMHF100	p_T^{miss}	$p_T^{\text{miss}} > 100 \text{ GeV}$	3
L1_DoubleIsoTau32er2p1	Double isolated τ	$p_T > 32 \text{ GeV}, \eta < 2.1$	15
L1_DoubleJet150er2p5	Double jet	$p_T > 150 \text{ GeV}, \eta < 2.5$	1

Table 3.2: Lowest thresholds and approximate rates of benchmark LIT object seeds at a $\mathcal{L} = 1.8 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ (PU ≈ 50).

The efficiency curves for some LIT physics objects with some standard thresholds and quality cuts applied are shown in Figure 3.20 and 3.21:

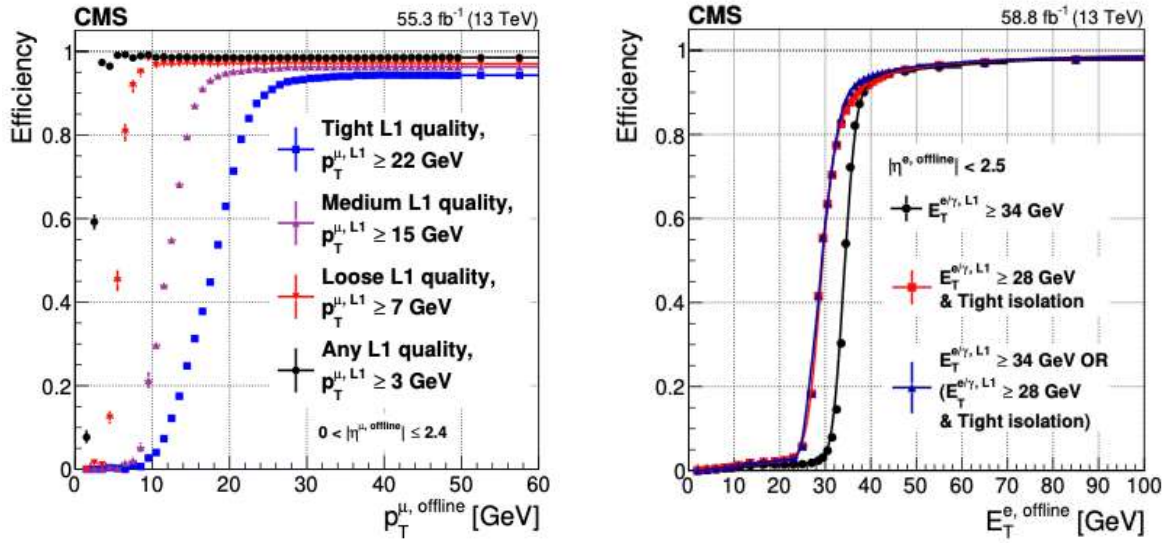


Figure 3.20: Level-1 trigger efficiencies in 2018 for muons for all possible Level-1 muon qualities as a function of muon p_T (left) and e/γ efficiency for two typical unprescaled algorithms including isolation as a function of the offline reconstructed electron E_T (right) [179].

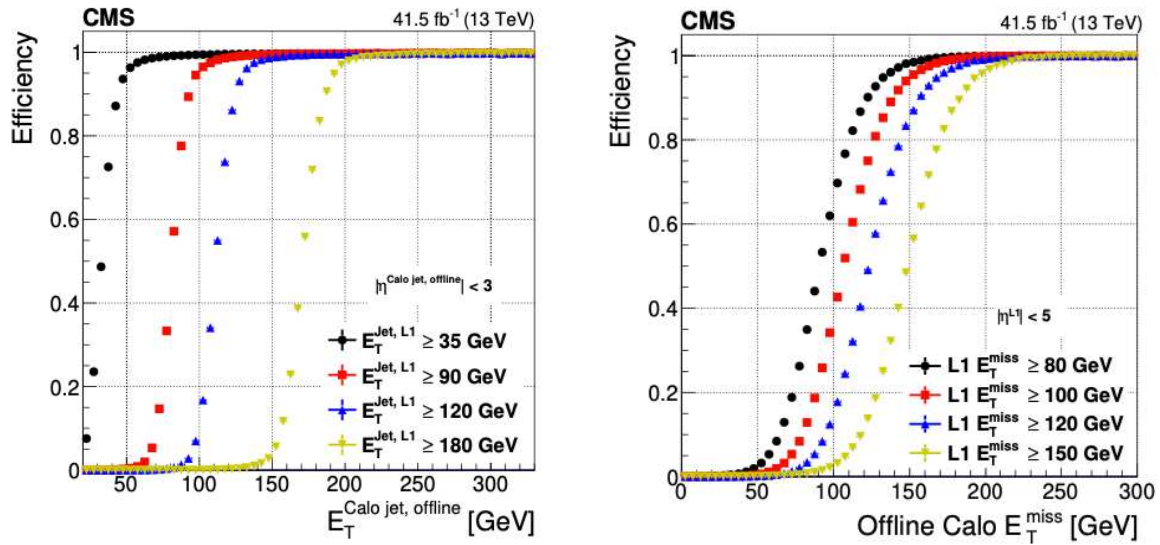


Figure 3.21: Level-1 trigger efficiencies for jet triggers for the barrel (left) and p_T^{miss} as a function of the offline calorimeter-based quantity (right) [179].

High Level Trigger

The main unprescaled HLT paths used in *Run 2* with the lowest thresholds and their corresponding rates are shown in Table 3.3:

Path	Description	Thresholds	Total Rate (Hz)
HLT_IsoMu24	Isolated single muon	$p_T(\mu) > 24$	235
HLT_Mu50	Non-isolated single muon	$p_T(\mu) > 50$	46
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8	Isolated dimuon	$p_T(\mu) > 17/8, M(\mu\mu) > 3.8$	28
HLT_Ele32_WPTight_Gsf	Isolated single electron	$p_T(e) > 32$	165
HLT_Ele115_CaloIdVT_GsfTrkIdT	Non-isolated single electron	$p_T(e) > 115$	17
HLT_Ele23_Ele12_CaloIDL_TrackIDL_IsoVL	Isolated dielectron	$p_T(e) > 23/12$	25
HLT_Photon10EB_TightID_TightIso	Isolated single photon	$p_T(\gamma) > 110, \eta(\gamma) < 1.479$	12
HLT_Photon200	Non-isolated single photon	$p_T(\gamma) > 200$	13
HLT_Diphoton30_22_R9Id*_Mass90	Isolated di-photon	$p_T(\gamma) > 30/22, M(\gamma\gamma) > 90$	40
HLT_DoubleTightChargedIsoPFTau35*	Isolated di-tau	$p_T(\tau) > 35/35$	40
HLT_PFJet500	Single jet	$p_T(j) > 500$	11
HLT_PFHHT1050	H_T	$H_T > 1050$	10
HLT_PFMET120_PFMHT120_IDTight	$p_T^{\text{miss}} + H_T^{\text{miss}}$	$p_T^{\text{miss}} > 120, H_T^{\text{miss}} > 120$	33
HLT_PFMET120_PFMHT120_IDTight_PFHHT60	$p_T^{\text{miss}} + H_T^{\text{miss}} + H_T$	$p_T^{\text{miss}} > 120, H_T^{\text{miss}} > 120, H_T > 60$	23

Table 3.3: Lowest thresholds and rates of main unprescaled HLT paths at a $\mathcal{L} = 1.8 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ (PU ≈ 50). The rate uncertainties are of $\mathcal{O}(\text{Hz})$. [186]

The allocated rate bandwidth for a typical HLT menu in 2018 is shown in Figure 3.22, indicating the rates consumed by different CMS physics groups.

CMS Preliminary (13 TeV, 2018, $2.0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$)

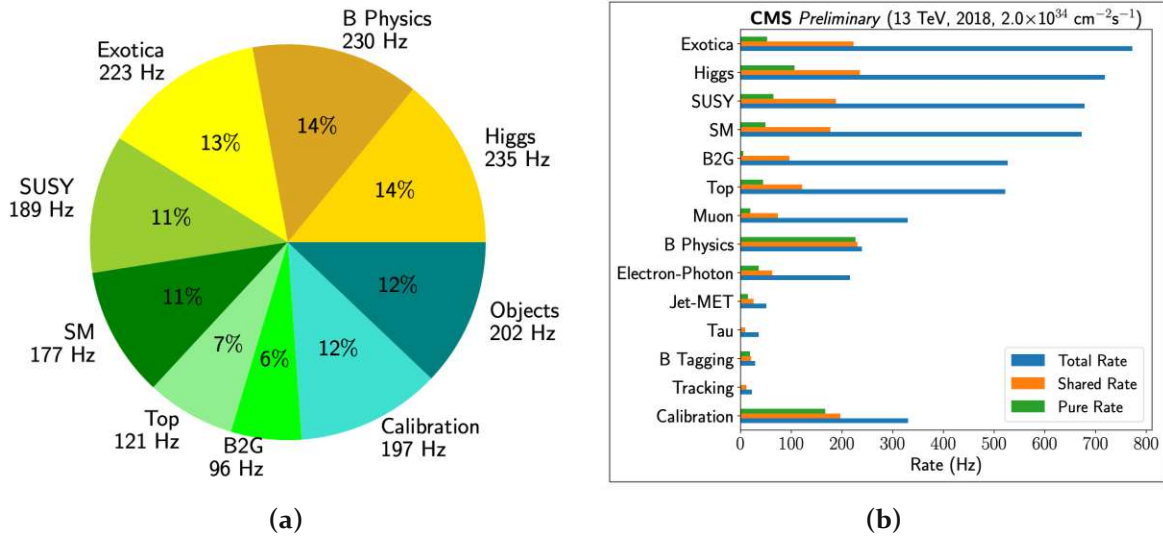


Figure 3.22: HLT rate bandwidth allocation for CMS physics groups in a typical HLT proton-proton collisions physics menu, scaled to $\mathcal{L} = 2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. The pure rate from an event is assigned to a given group if it is the only one that triggers the event. [187]

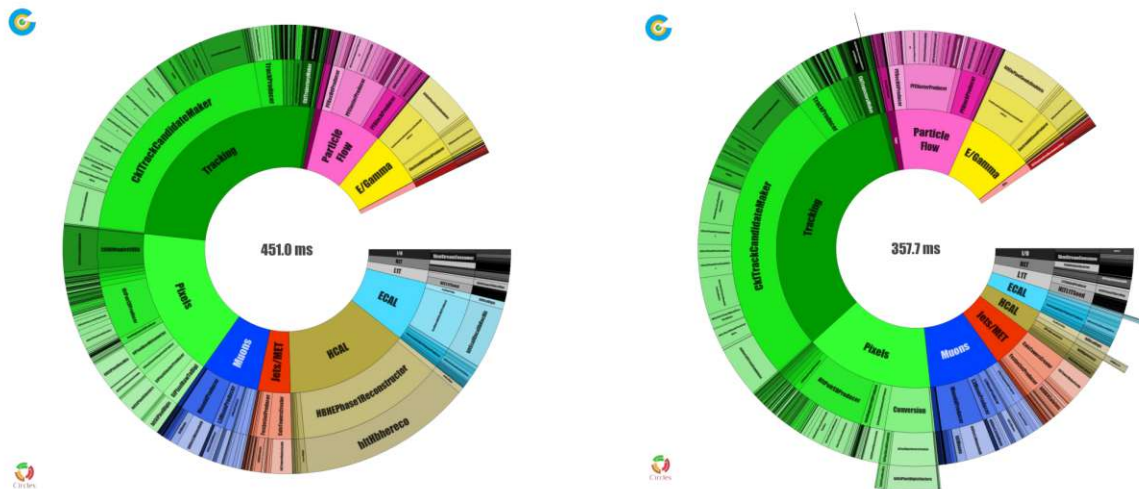


Figure 3.23: Distribution of the CPU processing time (timing or execution time) for a menu based on the 2018 definition with (right) and without (left) GPU acceleration, measured using data with $PU \approx 50$. The pie chart shows the distribution of the CPU time spent in different instances of CMSSW modules (outermost ring), the corresponding C++ class (middle ring), and by physics object or detector (innermost ring). The empty slice indicates the time spent outside of the individual algorithms. [188]

The performance of the HLT menu is also measured in terms of the total CPU and/or GPU processing time (timing or execution time), briefly discussed in Section 10.4.2. As shown in Figure 3.23 for a menu based on the 2018 definition, the total timing of the HLT menu is of the order of ≈ 450 Hz. The inclusion of offloading to GPU reconstruction in *Run 3* reduces the CPU usage by ≈ 21 %, increasing the throughput by ≈ 26 %.

Missing Momentum Triggers

The requirement of missing transverse momentum p_T^{miss} (introduced in Section 4.1) is a characteristic of many LHC searches, with non-interacting BSM particles escaping detection and manifesting themselves as a p_T^{miss} signature. Therefore, trigger paths requiring significant missing energy have been extensively used during LHC *Run 2*, including a number of searches for SUSY and exotic particles. The trigger strategy of many compressed SUSY analyses uses an online trigger selection on pure p_T^{miss} , including the *soft single-lepton analysis* (see Section 6.4.2), so a brief review of the performance of p_T^{miss} triggers is presented.

Improvements in the performance of p_T^{miss} energy sum algorithms at the trigger level are one of the priorities of the trigger study groups, especially in the context of the increasing luminosity and PU conditions at the LHC. The target is an improvement of the purity of physics objects and rate reduction without the loss of efficiency. Furthermore, in the specific case of p_T^{miss} , the mitigation of the high PU-dependence of rate is essential to prevent the increase in thresholds, as discussed in Section 4.1.

A summary of the main p_T^{miss} triggers used in *Run 2* is presented in Table 3.4. The thresholds of these p_T^{miss} paths and the corresponding seeds have changed across *Run 2*. In 2016, the paths were seeded by a variation of p_T^{miss} seeds labelled `L1_ETM[X]`, including variation with additional requirements on jets, of which the lowest unprecaled seed was `L1_ETM80`. In 2016, the paths with thresholds 90-110 GeV were precaled throughout the year, so the main threshold with the highest collected luminosity was

120 GeV. The paths with thresholds of 90 and 100 GeV were removed from the menu at the end of the year, as they contributed too much rate. In 2017, these paths were additionally seeded by [L1_ETMHF \[X\]](#) seeds that include hadronic energy from the HF in the p_T^{miss} calculation at LIT. This aligns the definition of L1- p_T^{miss} closer to that used at HLT and offline. The lowest unrescaled seed enabled for the full year was [L1_ETMHF120](#), while [L1_ETM\[110,115\]](#) were enabled for the majority of the year. With the increase of luminosity and pileup conditions, there were high LIT & HLT rates for these pure p_T^{miss} triggers. A mitigation strategy by CMS was the addition of cross triggers with a loose central H_T condition (> 60 GeV) both at L1 and HLT

([HLT_PFMET120_PFMHT120_IDTight_PFHT60](#)) to be used in a logical 'OR' with the original paths. In 2018, the seed definition was simplified to only include [L1_ETMHF \[X\]](#) seeds and [L1_ETM150](#). The primary seed remained [L1_ETMHF120](#), albeit disabled for a short period, for which [L1_ETM\(HF\)150](#) provided the correct seeding. Additional thresholds at 130 and 140 GeV were added for contingency, however, the 120 GeV threshold remained the lowest unrescaled one. Therefore, to summarise, one can consider [HLT_PFMET120_PFMHT120_IDTight](#) as the main pure $p_T^{\text{miss}} + H_T^{\text{miss}}$ path throughout *Run 2* with a rate of ≈ 33 Hz, as shown in [3.3](#).

HLT path	LIT seeds (2018)
HLT_PFMET[170-250] HLT_PFMET[NoMu]120_PFMHT[NoMu]120_IDTight	L1_ETMHF70 L1_ETMHF80 L1_ETMHF90 L1_ETMHF100 L1_ETMHF110 L1_ETMHF120 L1_ETMHF150 L1_ETM150
HLT_PFMET[NoMu]120_PFMHT[NoMu]120_IDTight_PFHT60	L1_ETMHF80_HTT60er L1_ETMHF90_HTT60er L1_ETMHF100_HTT60er L1_ETMHF110_HTT60er L1_ETMHF120_HTT60er

Table 3.4: Main p_T^{miss} -type HLT paths in *Run 2* and their corresponding seeds.

Specific details of the methodology of trigger efficiency measurements are discussed in detail in Section 10.2.3. An unbiased measurement of the p_T^{miss} trigger efficiency is typically performed in an independent data sample, selecting events with $W \rightarrow \ell + \nu$ leptonic decays. This ensures that the p_T^{miss} originates from a real source, the neutrino, rather than any fake sources of p_T^{miss} , such as detector effects or jet p_T mismeasurements, which are the main sources of fake p_T^{miss} .

The efficiency of pure PF- p_T^{miss} triggers (`HLT_PFMET[170-300]`) is shown in Figure 3.24, as a function of the offline reconstructed corrected PF- p_T^{miss} (see Section 4.1). The measurement is performed in an unbiased sample of events triggered by an isolated single-electron trigger and containing exactly one well-identified and isolated offline electron.

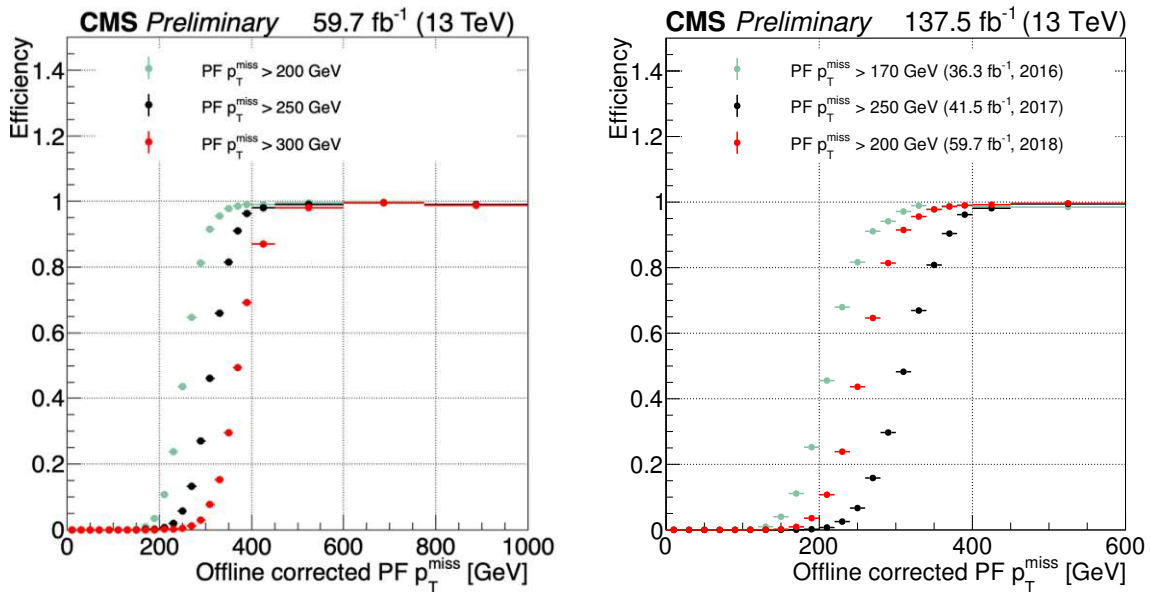


Figure 3.24: Trigger efficiency of pure PF- p_T^{miss} triggers (`HLT_PFMET[170-250]*`), measured with respect to the offline reconstructed corrected PF- p_T^{miss} , in 2018 data (left) and across Run 2, corresponding to the lowest unprescaled thresholds at a given year (right) [189].

For a threshold of 200 GeV, the efficiency reaches its plateau for offline p_T^{miss} at about 350 GeV, with a plateau efficiency close to 100 %. An independent measurement of $p_T^{\text{miss}} + H_T^{\text{miss}}$ trigger (`HLT_PFMET[NoMu]120_PFMHT[NoMu]120_IDTight`) is performed in the context of the *soft single-lepton analysis*, as described in Section 6.4.2. In summary, the independent efficiency measurements are compatible with one another,

and for the lowest unrescaled p_T^{miss} threshold of 120 GeV, the plateau is reached at ≈ 250 GeV. Taking into account the different measurements, one can roughly extrapolate the width of the turn-on to be of the order of ≈ 150 GeV, which reflects the resolution of p_T^{miss} at the HLT level.

Event Reconstruction

4.1 Global Event Description

Multi-purpose detectors have the exceptional ability to identify a range of different types of particles and combine the information to give a complete description of the underlying events. A sketch of a transverse slice of a segment of the CMS detector showing the expected responses to different types of particles is shown in Figure 4.1.

The individual *local* reconstruction of each of the subsystems is summarised above in Section 3.2. To summarise, for a general particle originating from the beam interaction region, *tracks* and *vertices* of charged particles are reconstructed from *hits* and *clusters* in the tracker, allowing a measurement of the momentum and charge. The measurements of the total energy deposits are extracted from (*super*)*clusters* in the calorimeters. Muons produce *hits* and *segments* in the muon detector layers, measuring their corresponding *tracks*. A simplified picture of different particle types is presented in the following points:

- **Electrons** and **photons** are absorbed in the ECAL to form electromagnetic showers; Electrons leave a curved track in the tracker.
- **Charged** and **neutral hadrons** can produce showers in the ECAL in the form of **jets**, which are then fully absorbed by the hermetic HCAL; Charged hadrons leave a curved track in the tracker.

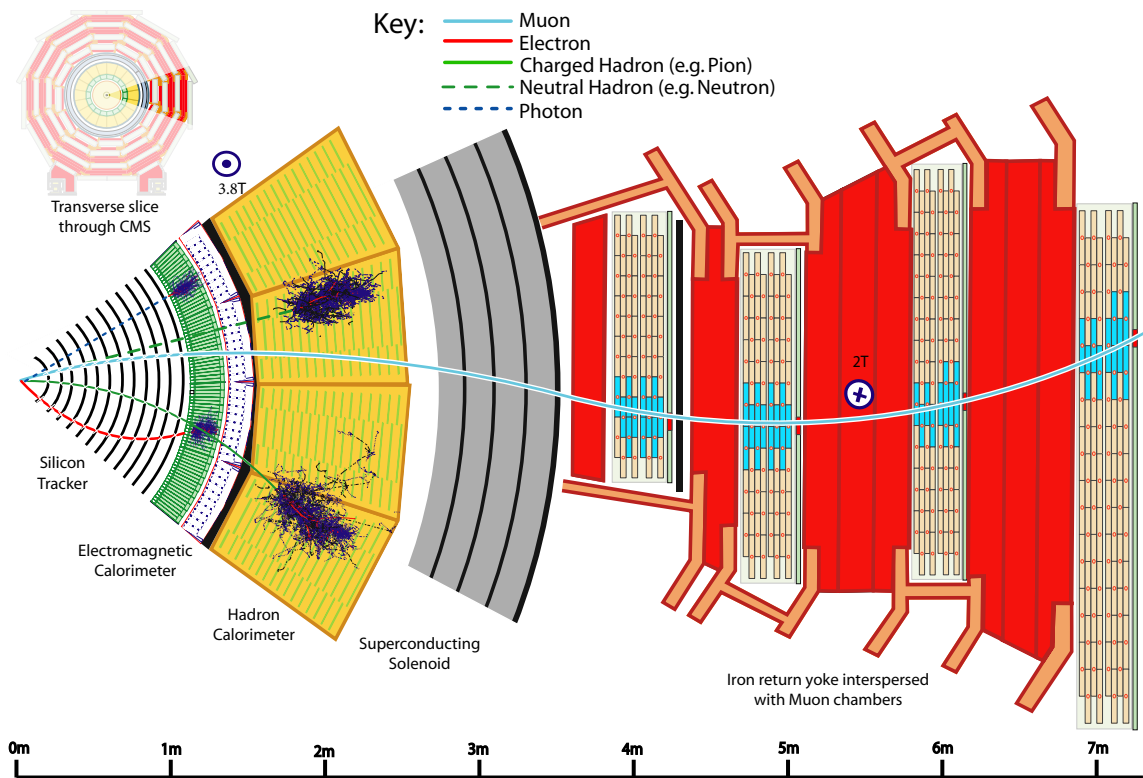


Figure 4.1: Transverse slice of one segment of the CMS detector showing the expected interactions with different types of particles. The muon and charged pion are positively charged, and the electron is negatively charged. [136]

- **Muons** are much more penetrating and typically traverse the calorimeters with little or no interactions to be detected by the muon chambers; Muons leave a slightly curved track in the tracker.
- **Taus** decay mainly into hadrons, or electrons or muons and neutrinos, which are detected by the corresponding sub-systems.
- **Neutrinos** escape the detector undetected, leaving a signature of missing momentum.

Particle-Flow Algorithm

A significantly better *global* reconstruction and identification of the *physics objects* and, correspondingly, a more complete event description can be achieved by combining the information from all the detectors. This is achieved by using the particle-flow (PF) [190] reconstruction algorithm, which was first implemented in the ALEPH experiment at LEP [191] and is used in the conceptual planning of detectors for future colliders [150]. CMS is the first experiment to take advantage of PF at a hadron collider and its design was motivated by the necessary requirements. A fine granularity and excellent resolution of the detectors are vital elements for PF, allowing for the separation of single particles. Furthermore, it is also taken advantage of at the HLT level, providing improved efficiency and purity of trigger objects. A combined fit of measurements of all identified physics objects results in improved resolutions, efficiencies and purities, which is then reflected in the final performance of all physics analyses.

The PF algorithm correlates the main elements from the detector layers, which are namely tracks from the tracker and muon systems and energy clusters from the calorimeters. A given particle is, in general, expected to give rise to several PF elements in the CMS subdetectors, as illustrated in Figure 4.2. The reconstruction of a particle, therefore, first proceeds with a *link algorithm* that connects the PF elements from different subdetectors in order to form PF blocks of associated elements. Within a given PF block, particles are reconstructed in hierarchical order, starting from muons, electrons and isolated photons, followed by charged and neutral hadrons. At the end of each iteration, elements of the reconstructed particles are then removed prior to the next iteration. Once a global event description is complete, the events are post-processed in order to increase their purity.

Calorimeter Clusters

The clustering of calorimeter deposits is necessary for a measurement of the energy and direction of stable neutral particles such as photons and neutral hadrons. Furthermore, clustering algorithms separate neutral from charged hadron energy deposits to reconstruct and identify electrons and accompanying bremsstrahlung photons. Additionally, it also assists the energy measurement of charged hadrons for inaccurate track measurements, especially in the case of high p_T and low-quality tracks.

A specific clustering algorithm was developed for the PF reconstruction in order to provide high detection efficiency and separate close energy deposits, as illustrated in Figure 4.2. It is performed separately in the ECAL and HCAL detectors, split into the barrel and endcaps. Cluster *seeds* are cells with an energy above a given threshold and their neighbouring cells, representing local maxima. *Topological* clusters are formed by aggregating neighbouring cells of a seed above a threshold that is twice the noise level. Finally, an expectation-maximisation algorithm is applied to reconstruct the final clusters within the topological clusters [192]. In the ECAL, the basic clusters are then merged together to form *superclusters* (SCs).

Charged Particles

Tracks

Particle properties such as momentum and charge can be deduced from the curvature of the tracks. Tracks are reconstructed using the measurement of the positions of the particle *hits* at key measurement points located at different layers of the pixel and strip tracker [193]. Therefore, with the additional inputs from the other sub-systems, tracks are used in the reconstruction of electrons, muons, charged hadrons, taus and jets.

The local track reconstruction comprises an iterative procedure that involves four phases:

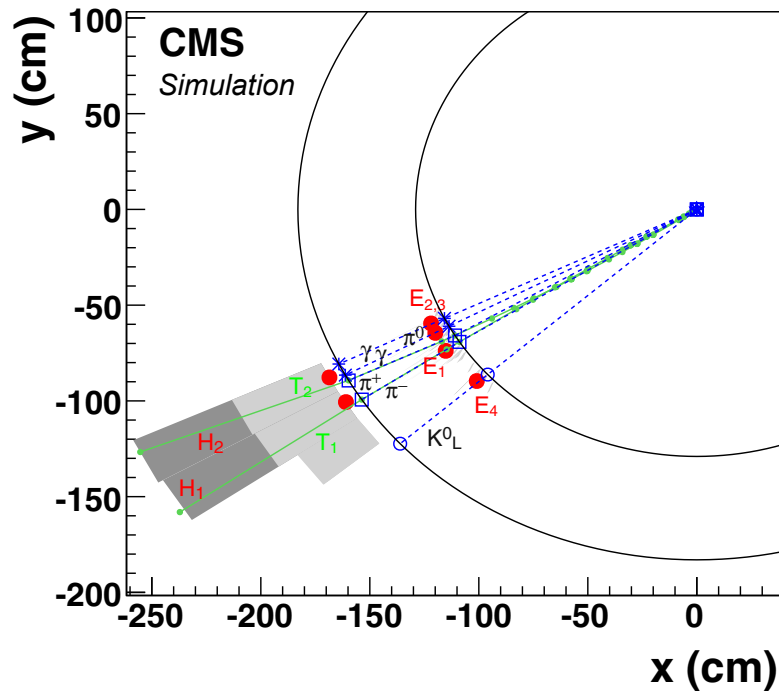


Figure 4.2: Event display of an illustrative jet made of five particles only in the (x,y) view. The ECAL and HCAL surfaces are represented as circles centred around the interaction point. The K_L^0 , the π^- and the two photons from the π^0 decay are detected as four well-separated ECAL clusters denoted $E_{1,2,3,4}$. The π^+ does not create a cluster in the ECAL. The two charged pions are reconstructed as charged-particle tracks $T_{1,2}$, appearing as circular arcs, pointing towards two HCAL clusters $H_{1,2}$. [190]

- **Seed generation:** from track candidates generated from several *hits* compatible with a charge-particle trajectory
- **Track finding:** gathering additional hits along the expected candidate track as *clusters* via pattern recognition (KF) to build the final trajectory
- **Track fitting:** final fitting to determine the properties (origin, transverse momentum, direction) of the charged particle candidate (KF or GSF)
- **Track selection:** final track selection based on additional quality criteria

The main track reconstruction algorithms are based on *seeds*, which are a coarse estimate of input tracking parameters. Prior to track reconstruction, in the first local reconstruction phase, localised hits are reconstructed with a very high *hit* detection

efficiency ($> 99\%$) in the tracker. Clusters are formed from hits on adjacent layers (*hit-multiplets*). The seeds typically are required to be compatible with particles originating from the beam interaction region. This generates track candidates, which define the initial trajectory parameters that serve as inputs to the track reconstruction algorithms.

The *seeds* can also be provided externally 'outside-in' using inputs from the other detectors, such as from the ECAL or muon detectors. For the external seeds, there exist various algorithms that, for example, focus solely on reconstructing muon tracks seeded by the muon detectors, electrons seeded from the calorimeters, and particles from secondary interactions originating from outside the pixel detector.

Track finding relies on an iterative procedure based on a combinatorial *Kalman Filter* (KF) method [194]. The tracking efficiency is increased by employing this procedure in up to ten iterations with different seeds to reconstruct tracks with lower momentum and quality. Firstly tracks that are easiest to reconstruct, corresponding to higher momenta and a larger number of hits, are found. These tracks (and corresponding hits) are then ignored in the following iterations, and the quality criteria are relaxed in order to search for more difficult tracks with higher combinatorial complexity. These include tracks generated by muons as well as displaced tracks. Once the tracks are formed, ambiguities and instances of double counting are resolved.

At this point, the full tracking information is available and any bias introduced with the imposed constraints is removed by a final fit. One of the main pattern recognition algorithms used for track fitting is a combinatorial KF method, where the algorithm extrapolates the trajectory according to the equations of motion of a charged particle in a magnetic field traversing through a material. More specialised variants of the KF algorithm are also implemented, such as the Gaussian Sum Filter (GSF) [195] for electron reconstruction.

Vertices and Pileup

The tracker, and especially the pixel detector, is crucial in the determination of the proton-proton interaction points, known as the primary interaction vertices (PVs), which are reconstructed as the intersection of associated tracks. Apart from the determination of PVs, the high granularity and good resolution allow for the reconstruction of secondary vertices (SVs) of long-lived particles, providing a measure of the lifetime of the decaying particles. Therefore, tracks are used to identify jets originating from b-quarks, through the reconstruction of displaced vertices associated with a given jet.

The determination of the proton-proton interaction point, known as the primary interaction vertex (PV), is a vital element of object identification and global event descriptions. As discussed in Section 3.1.2, additional pileup (PU) interactions arise during the same or neighbouring bunch crossings. As shown in Figure 3.4, in LHC *Run 2* the average PU was 32, whereas in *Run 3* the average PU already increased to 44, with peak PU values going up to 60-70. This results in a challenging environment and disentangling the PV from the secondary pileup PU collisions is essential in order to identify the particles originating from the primary interaction.

Reconstructed tracks are used to measure all proton-proton interaction vertices in a given event, which includes the PV from the signal collision and any secondary vertices from pileup [193]. The reconstruction of the PV comprises an initial selection of the tracks compatible with the *beamspot*, which represents the luminous region where the LHC beams collide. This is followed by clustering of the tracks that appear to originate from the same interaction vertex using a deterministic annealing (DA) algorithm [196], finding the most probable tracks and positions for each vertex. The final step is fitting for the position of each vertex using its associated tracks, which is performed using an *adaptive vertex fitter* (AVF) [197] to compute the best estimate of vertex parameters, including the 3D position and covariance matrix, as well as indicators of the success of the fit. Ultimately, the vertex with the highest quadratic sum of transverse momenta $\sum p_T^2$ of tracks associated with it is taken as the PV, while all other vertices

are considered pileup vertices.

The resolution of the PV position depends strongly on the number of tracks used to fit the vertex and the momentum of those tracks. For vertices with many tracks, the vertex position resolution is 10-12 μm in each of the three spatial dimensions [193]. The primary-vertex efficiency is estimated to be close to 100 % when more than two tracks are used to reconstruct the vertex.

Particle candidates originating from pileup vertices can cause mismeasurements of the primary interaction, and mitigation of their effects is crucial, especially in the context of high luminosity LHC runs; Figure 3.5 in 3.1.2 shows the event display of a special high PU fill recorded during LHC *Run 2*, indicating the challenging environment. Jet and energy sums (H_T, p_T^{miss}) are strongly affected by PU in terms of purity, also on the trigger level. There are a number of methods that target pileup mitigation [198], such as *charged-hadron subtraction* (CHS), used to remove charged hadron tracks associated with pileup vertices in the PF jet clustering algorithm. New approaches, such as *pileup per particle identification* (PUPPI) [199], are being developed and will be essential for the HL-LHC, with PU expected to be of order 200.

Muons

The reconstruction and identification of muons is one of the key objectives of CMS, as they provide a very clean signal in a QCD-dominated environment. Muon detectors, by themselves, already provide a high reconstruction efficiency, with high purity provided by the upstream calorimeters. The best resolution over the entire acceptance is achieved by combining information from the muon system with momentum information from the inner tracker. Ultimately, there are three different types of muon reconstruction algorithms:

- **standalone muon:** track segments are formed using hits from the DT or CSC detectors to form seeds for *standalone-muon* tracks that are fitted using pattern recognition (KF) using hits from all muon systems (including RPCs)

- **global muon:** the *standalone-muon* tracks are matched to tracks in the inner tracker (*inner tracks*), and the hits are combined to form a *global-muon track*. This is an example of *outside-in* seeding.
- **tracker muon:** *inner tracks* above a certain momentum threshold are extrapolated to the muon system. It is qualified as a *tracker muon track* if at least one muon segment matches the extrapolated track.

Approximately 99 % of muon candidates within the acceptance of the muon detectors are either reconstructed as global or tracker muons. Global and tracker muons that share the same inner track are merged into a single candidate. The momentum measurement is solely based on the tracker measurement for muon $p_T < 200$ GeV.

Identification quality criteria are applied to muon candidates in order to reduce misidentified particles and increase purity. For example, *punch-through* particles, mainly originating from charged hadron showers that manage to traverse past the calorimeters, may be misreconstructed as muons. The PF muon identification algorithm additionally uses information on the energy deposits from the calorimeters in order to improve performance. The PF algorithm considers properties of global and tracker muons only.

A measure of the hadronic energy surrounding leptons is given by the *isolation*, which is defined by the sum of p_T of inner tracks and calorimeter deposits within a cone around the lepton defined by $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.3$. Isolated muons are selected by applying a maximum requirement on this quantity. In the PF algorithm, the relative isolation is required to be less than 10 % of the muon p_T .

Further muon identification selection criteria [173] are defined based on a set of variables related directly to muon reconstruction (e.g. track fit χ^2 , the number of hits per track or the degree of matching between tracker tracks and standalone-muon tracks) as well as inputs from outside the reconstructed tracks (e.g. compatibility with PV). Their usage is dependent on the balance between the required identification efficiency and purity of an analysis, and thus the criteria are embedded within several dedicated working points (WP) or IDs:

- **loose muon ID:** muons selected by PF algorithm that are also tracker or global muons, targeting prompt muons from the PV and muons from light and heavy flavour decays
- **medium muon ID:** imposing additional requirements (e.g. track fit parameters) with respect to the *loose muon ID*, tuned towards an overall efficiency of 99.5 % from simulated W and Z decays, targeting prompt muons and muons from heavy flavour decays
- **tight muon ID:** imposing additional requirements (e.g. requiring compatibility with tracker and global muons, and an impact parameter selection compatible with the PV) targeting suppressing muons from hadronic punch-through, decay in flight and cosmic sources

There also exist *soft* and *high momentum muon IDs*, as well as MVA-based IDs, targeted towards specific cases such as low- p_T muons from decays of b-hadrons and quarkonia analyses and muons with $p_T > 200$ GeV, respectively.

Electrons and Photons

The reconstruction of electrons and photons is performed in a similar manner due to the interplay between the emission of bremsstrahlung photons from electrons and photon conversions into electron-positron pairs. As described in Section 3.3.4.1, traditional *ECAL-based* electron reconstruction is based on measurements of energy deposits in the ECAL SCs, aimed at characterising energetic and isolated electrons. However, there exist large inefficiencies are seen for electrons in jets with additional particle deposits or low momentum electrons with highly curved tracks that are not caught by the SCs.

The PF algorithm takes advantage of additional information from the inner tracker using a *tracker-based* electron seeding. An iterative tracking method is employed where electron candidates are seeded by tracks with a corresponding ECAL cluster. Selected

tracks are fitted with Gaussian-sum filter (GSF) [195] method, which is a non-linear fitting better adapted than the KF method for electrons. Photon candidates are seeded purely from ECAL SCs, without using the tracking information.

Further electron identification selection criteria are based on a set of variables related directly to electron reconstruction (ECAL and track information) as well as compatibility with PV. More specifically, the variables are related to shower shapes, track-cluster matching with ECAL, track vertices as well as a veto on electrons from converted photons. A set of *Veto*, *Loose*, *Medium* and *Tight* WPs based on a sequential selection are provided for average identification efficiencies in $Z \rightarrow ee$ events corresponding to 95 %, 90 %, 80 % and 70 %, respectively. They were optimised separately for the barrel ($|\eta_{SC}| \leq 1.479$) and endcap ($|\eta_{SC}| > 1.479$) regions, with the exact values presented in Table 4.1. Their usage is dependent on the balance between the required identification efficiency and purity of an analysis, depending on the dominant backgrounds.

Variable	Description	Veto	Loose	Medium	Tight
$\sigma_{i\eta i\eta}$	Cluster shower shape covariance in η direction	0.0115 (0.037)	0.011 (0.0314)	0.00998 (0.0298)	0.00998 (0.0292)
H/E	Cluster energy ratio between HCAL and ECAL	0.356 (0.211)	0.298 (0.101)	0.253 (0.0878)	0.0414 (0.0641)
$ \Delta\eta_{in} $	$\Delta\eta$ between SC position and track direction	0.00749 (0.00895)	0.00477 (0.00868)	0.00311 (0.00609)	0.00308 (0.00605)
$ \Delta\phi_{in} $	$\Delta\phi$ between SC position and track direction	0.228 (0.213)	0.222 (0.213)	0.103 (0.045)	0.0816 (0.0394)
$ 1/E - 1/p $	Difference between inverse of SC energy and track momentum	0.299 (0.150)	0.2421 (0.140)	0.134 (0.130)	0.0129
N_{hit}^{miss}	Expected number of missing hits for inner track	2 (3)	1	1	1

Table 4.1: Electron identification variables based on shower/cluster shape, track-cluster matching with ECAL, track vertex as well as rejection of photons conversion. A set of *Veto*, *Loose*, *Medium* and *Tight* WPs are presented. Upper bounds on the variables are listed for the barrel (endcap) regions with $|\eta_{SC}| \leq 1.479$ ($|\eta_{SC}| > 1.479$.)

There also exist special multivariate (MVA) based electron and photon IDs, trained on dedicated simulated Drell-Yan MC samples using the variables related to electron/photon reconstruction. The training is done with prompt electrons as signal and unmatched plus nonprompt electrons as background. The resulting selection is on a single MVA discriminator with several WPs corresponding to different efficiencies (e.g.

80 %, 90 %). Additional selections on the relative isolation and the impact parameters (d_{xy} , d_z) are recommended and applied on an analysis-specific basis, discussed below in Section 4.2.

Taus

The tau lepton τ with a mass $m_\tau \approx 1.78$ GeV is the only lepton that is sufficiently massive to decay into hadrons. Approximately a third of the decays are to electrons and muons, whereas almost all the remaining decay modes are hadronic. The final states are typically combinations of charged and neutral mesons (charged/neutral pions π^\pm , π^0), and a tau neutrino ν_τ . Hadronic decays of tau τ_h leptons are reconstructed using the *hadrons-plus-strips* (HPS) algorithm [200, 201, 202], which distinguishes between the individual τ_h decay modes. It combines the information from reconstructing charged hadrons and neutral π^0 candidates. The charged hadrons are reconstructed using their associated tracks in the inner tracker, whereas the π^0 candidates are reconstructed by clustering photon and electron PF candidates from photon conversions in rectangular *strips*, which are regions of pseudorapidity and azimuth ($\eta \times \phi$). The τ_h reconstruction HPS algorithm achieves an efficiency of 50 – 60 %. A new *DeepTau* algorithm has been recently developed, which uses information from all reconstructed particles in the vicinity of a hadronic tau, using a deep neural network (DNN) [203]. It has superior performance and improves the identification efficiency of hadronic tau lepton decays up to 30 %.

Jets

In hadron colliders, the dominant QCD processes result in abundant production of quarks and gluons at high energies. Due to colour confinement of the strong interaction, they undergo the process of hadronisation and fragmentation into charged and neutral hadrons (mostly pions π^\pm , π^0 , kaons K^\pm , K^0 , protons, neutrons and photons).

Ultimately, this results in parton showers that are relatively collimated in a particular direction and classified as *jets*. Electrons and muons can also form part of jets, especially in light and heavy meson decays (especially b-flavoured mesons). Typically, most of the jet energy is carried by charged hadrons ($\approx 60\%$), followed by photons ($\approx 25\%$) and neutral hadrons ($\approx 10\%$), with a minor contribution from leptons.

Once the PF algorithm identifies isolated muons, electrons and photons, the hadrons are then reconstructed from tracks and calorimeter deposits. ECAL and HCAL clusters not linked to any tracks are associated with neutral hadrons and photons. Given the jet energy fractions, precedence is given to photons in the ECAL within the tracker acceptance ($|\eta| < 2.5$), while HCAL clusters are associated with neutral hadrons. Outside of the tracker acceptance, charged and neutral hadrons cannot be distinguished, and ECAL clusters linked to a given HCAL cluster are assumed to originate from the same hadron shower, while ECAL clusters without such a link are classified as photons. The PF particles within the hadrons are finally clustered into jets using the anti- k_T algorithm [204, 205] within a distance parameter $R = 0.4$ (AK4).

Reconstructed PF jets are required to pass a set of minimal identification criteria [206] based on jet energy fractions and particle multiplicity in order to discriminate physical jets from noise. There are three dedicated WPs, of which the *loose* and *tight* IDs are designed to remove jets originating from calorimetric noise. There is also a *tight lepton veto* ID that is also designed to reject the potential background from misreconstructed electrons and muons. This effectively resolves also the ambiguity between isolated lepton candidates and jets reconstructed from single lepton candidates, also known as *jet cleaning*.

The jet energy scale (JES) and resolution (JER) is calibrated in several steps [207] in order to provide more accurate measurements of jet properties. Corrections as a function of jet p_T and η are applied, taking into account pileup interactions, jet detector response to hadrons and finally, residual differences between data and MC simulation.

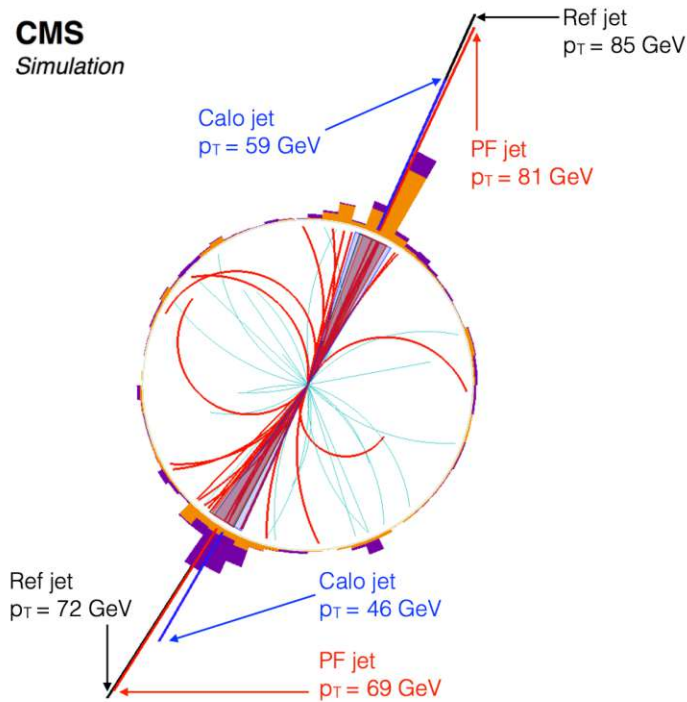


Figure 4.3: Jet reconstruction in a simulated dijet event, where the particles clustered in the two PF jets are displayed with a thicker line. The PF jet p_T , indicated as a radial line, is compared to the p_T of the corresponding generated (Ref) and calorimeter (Calo) jets. [190]

B-Jets

The efficient identification (*tagging*) of heavy-flavour jets originating from bottom quarks (*b-jets* or *b-tagged jets*) is important in precision measurements of SM processes, such as the decay of top quarks and Higgs bosons, as well as BSM searches. Therefore, being able to discriminate between heavy-flavour jets and those originating from light-flavour quarks or gluons (light-flavour jets) is important.

Given the higher mass and consequently long lifetime ($c\tau \approx 450 \mu\text{m}$) of the b-quark, its decay leads to displaced tracks forming a secondary vertex (SV), as shown in Figure 4.4.

Standard *b-tagging* algorithms primarily rely on the reconstruction of these displaced SVs [208, 209, 210]. Furthermore, they also exploit the additional properties of the

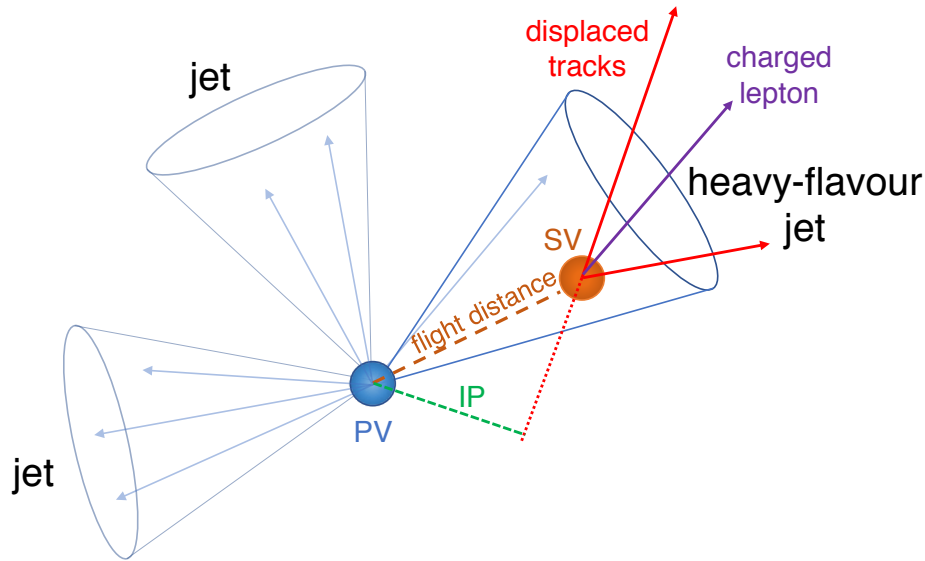


Figure 4.4: Diagram of a heavy-flavour jet with a secondary vertex (SV) from the decay of a b or c hadron resulting in charged-particle tracks that are displaced with respect to the primary interaction vertex (PV), and thus with a large impact parameter (IP) value [208].

hadrons within the jet substructure to provide further separation, such as displaced tracks, high particle multiplicity, probability of lepton in the decay chain and the momentum p_T and direction η of the underlying jets.

A commonly used b-tagging method is the *Combined Secondary Vertex* (CSVv2) algorithm combines the information of displaced tracks with the information on SV using a multivariate technique. It uses SVs reconstructed using the *Inclusive Vertex Finding* (IVF) algorithm, which considers all reconstructed tracks in the event with $p_T > 0.8$ GeV and a longitudinal impact parameter $|d_z| < 0.3$ cm. The output of the CSVv2 discriminator for jets of different flavours is shown in Figure 4.5.

There exist a number of other sophisticated algorithms, such as the Jet Probability (JP/JBP) taggers or combined multivariate analysis (cMVAv2) tagger, which combines the discriminator values of various taggers. Another improved version of the CSV algorithm was developed that uses deep machine learning (DeepCSV) [182], with the use of DNNs with more hidden layers, more nodes per layer and simultaneous training in all vertex categories and for all jet flavours.

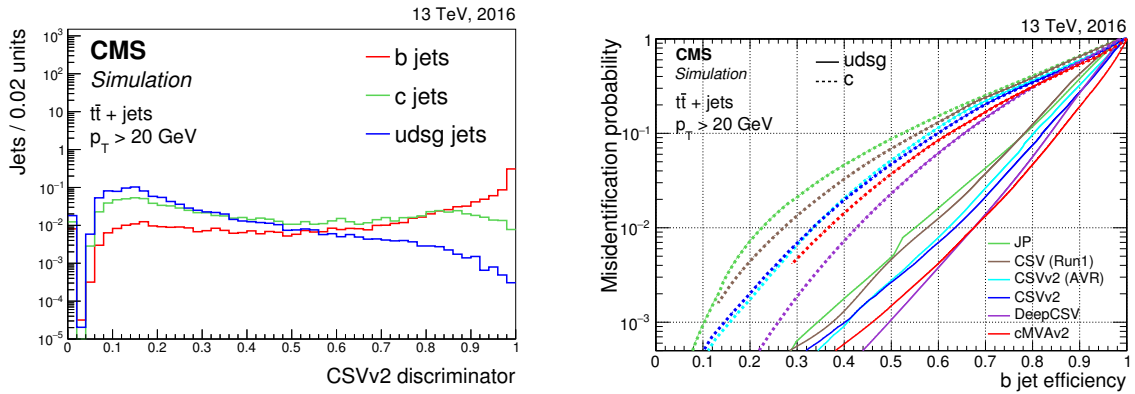


Figure 4.5: Distribution of the CSVv2 discriminator values for jets of different flavours in $t\bar{t}$ events (left) and misidentification probability for c and light-flavour jets versus b jet identification efficiency for various b tagging algorithms applied to jets in $t\bar{t}$ events (right) [208].

Three standard working points (*loose*, *medium* and *tight*) are defined for each b-tagging algorithm, which correspond to thresholds on the discriminators for misidentification probabilities of around 10 %, 1 %, and 0.1 % for light-flavour jets (*udsg*), respectively. The medium WP (defined by a CSVv2 discriminator value > 0.8484) has ≈ 63 % b-tag efficiency and a misidentification probability of ≈ 12 % and ≈ 0.9 % for c and light flavoured jets, respectively. As shown in Figure 4.5, the DeepCSV tagger generally outperforms all the other b jet identification algorithms when discriminating against c jets or light-flavour jets.

Recently, more-sophisticated taggers such as DeepJet [183] or ParticleNet [184] have been developed, which examine all jet constituents simultaneously and perform quark-gluon tagging using DNNs, with improved performance in the classification of heavy flavour jets.

Missing Transverse Momentum

Particles with weakly-interacting nature, such as neutrinos, do not interact with the detector and escape undetected. This also occurs in the case of postulated exotic particles, such as neutralinos $\tilde{\chi}_1^0$ or other dark matter candidates in various BSM models.

Due to momentum conservation, their presence can be inferred by the momentum imbalance in the events. Due to missing information of the initial momentum of the colliding partons, the measurement is quantified in the transverse plane in the form of missing transverse momentum ($p_T^{\vec{\text{miss}}}$). It is defined as the negative vector sum of the particle momenta:

$$p_T^{\vec{\text{miss}}} = - \sum_i^{\text{particles}} p_{T,i} \quad (4.1)$$

where the particles are typically the collection of PF candidates, resulting in PF- $p_T^{\vec{\text{miss}}}$. The quantity most commonly used is the magnitude of $p_T^{\vec{\text{miss}}}$, simply referred to as p_T^{miss} . The quantity of p_T^{miss} is sensitive to mismeasurements of the PF candidate momenta, induced by instrumental effects such as nonlinearity of the detector response or p_T thresholds in the calorimeters and tracker. Therefore, special corrections are applied: *type-I* corrections are calculated by propagating the jet energy corrections (described in Section 4.1) by subtracting the in the calculation of $p_T^{\vec{\text{miss}}}$.

4.2 Object Definitions

The final definitions of the reconstructed and identified particle candidates (or *physics objects*) are highly dependent on the specifics of each analysis. There is always a trade-off between the efficiency and purity of the objects, depending on the quality criteria applied in identification. The choice is finally motivated by the specific use-cases, taking into account considerations such as the dominant backgrounds or statistical power that is required, attempting to find a healthy balance. The object definitions of the *soft single-lepton analysis* in Part II are presented in this section.

Lepton Acceptance

The ability of the detectors to efficiently reconstruct soft leptons¹ is crucial in the context of compressed SUSY signatures with leptons in the decay chain. Lower p_T muons can lose sufficient energy in the calorimeters so that they do not reach the second

station of muon chambers, which is needed for standard muon reconstruction. The experiment is able to reliably identify muons down to 3 GeV and electrons down to 4.5 GeV, which borders on the limits of the initial design considerations. The analysis aims to go as low as possible in lepton p_T and considers muons with $p_T > 3.5$ GeV and $|\eta| < 2.4$ and electrons with $p_T > 5$ GeV and $|\eta| < 2.5$. Electrons that fall into the ECAL crack ($1.4442 < |\eta_{SC}| < 1.566$) between the barrel (EB) and endcap (EE) are vetoed due to low efficiencies in that region. The acceptance thresholds are determined by independent efficiency measurements, discussed in Section 6.3.2 in the context of simulation correction factors.

Electron Channel

One of the main improvements of the analysis with respect to *Run I*, which only considered muons, is the inclusion of the electron channel. In order to enable this, a dedicated study is performed to define the final identification criteria, especially in the context of low- p_T electrons that are targeted by the analysis, which are more difficult to reconstruct than muons.

The target is to find a satisfactory trade-off between a high identification efficiency of real electrons and a low misidentification rate of fakes (*fake rate* or *mismatch efficiency*). The measurement of these quantities is performed for the standard electron identification working points (or *IDs*) discussed in Section 4.1, with the *cut-based* ones explicitly defined in Table 4.1 in terms of tracker and ECAL variables.

The efficiency measurements are performed in simulated samples of privately-produced benchmark signal points with different Δm and main background samples ($W + jets$, $t\bar{t}$). The misidentification studies are additionally performed for the main nonprompt backgrounds (QCD multi-jet and Z_{inv}). Further details of the signals and dominant backgrounds are covered in Section 6. Basic preselection (Section 6.4.3.1) criteria are applied ($p_T^{\text{miss}} > 200$ GeV, $H_T > 200$ GeV, $p_T^{\text{ISR}} > 100$ GeV) in order to perform the

¹*Leptons* refers to charged light leptons (electrons, muons) and not taus.

measurements in the phase-space of interest.

The efficiency measurement is performed in a pure sample of electrons from semi-leptonic decays of the W boson. This is achieved by matching the reconstructed electrons within a $\Delta R < 0.3$ to the generated electrons and additional generator matching to the source from the hard scatter. The efficiency is defined as the ratio of the number of events that pass a given electron ID with respect to the total number of events in the sample, and it is calculated as a function of the generated electron p_T , as shown in Figure 4.6 (top).

The mismatch efficiency (*fake rate*) is defined as the fraction of fake or nonprompt electrons that pass the ID selection. Fake or nonprompt leptons are selected if the reconstructed electron is not matched to a generated electron. The misidentification efficiency is calculated as a function of the reconstructed electron p_T . The measurement is performed in a sample of reconstructed electrons that pass the ID selection, with the numerator defined by the requirement that the lepton is fake or nonprompt, as shown in Figure 4.6 (bottom). Ultimately, this definition is dependent on the composition of the different simulated samples and, more specifically, the relative proportions of prompt and nonprompt electrons. Therefore, to get a better appreciation of the absolute mis-identification efficiencies, the measurements are repeated in a sample of fake or nonprompt reconstructed electrons, with the numerator defined by the additional requirement of the ID.

Considering both the efficiency and different misidentification rate measurements across the different samples, several conclusions can be drawn. Firstly, the general purpose MVA-based IDs have too high mismatch efficiencies in the very low p_T region. The MVA training was performed and tuned for higher $p_T > 10$ GeV leptons, and thus the MVA IDs are deemed unsuitable for the analysis. Generally, in the low- p_T regime, the efficiencies of the *cut-based* working points are still relatively low. Therefore, a more detailed study of the electron ID variables, listed in Table 4.1 is performed.

The dedicated $N-1$ study involves removing each of the selections on the ID variables and re-calculating the efficiencies and misidentification rates. The results indicate

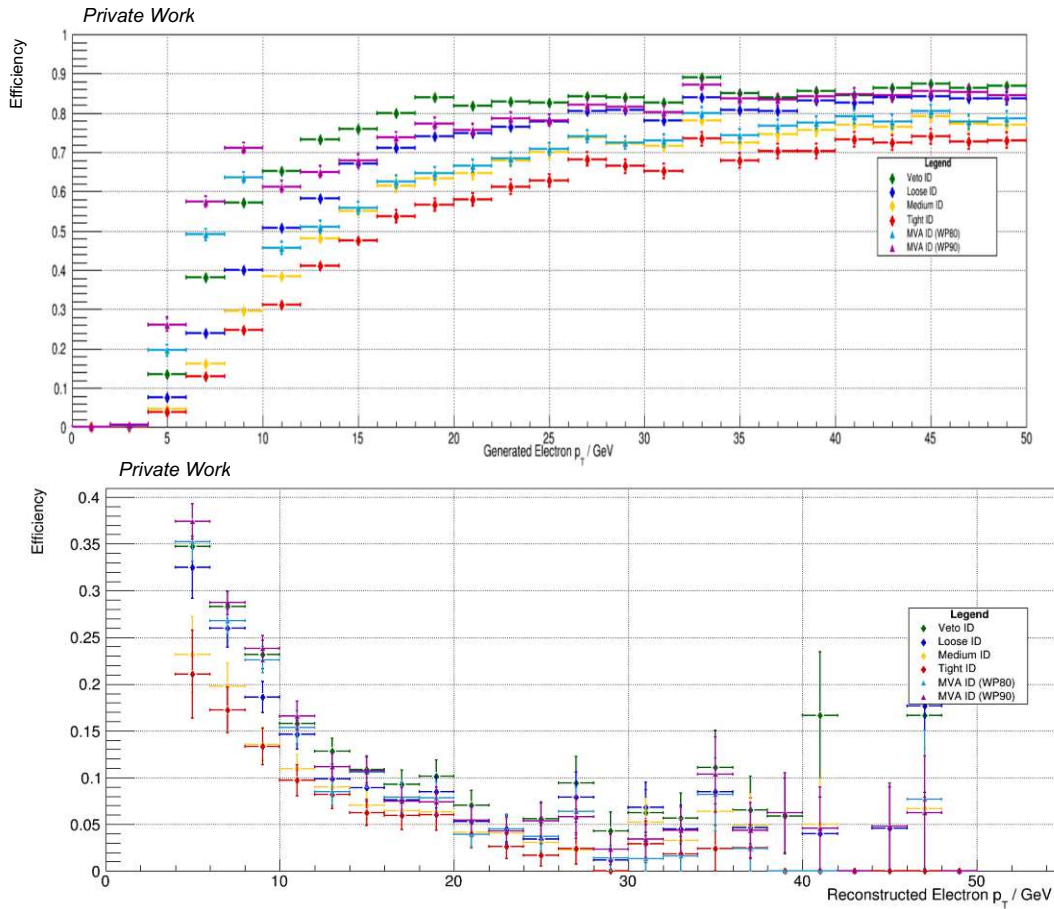


Figure 4.6: Electron ID efficiency measured in a $t\bar{t}$ sample (top) and mismatch efficiency (*fake rate*) calculated in a $T2t\bar{t}$ signal sample ($m_{\tilde{\tau}_1} = 300$ GeV, $\Delta m = 30$ GeV).

that the selection on cluster shower shape covariance $\sigma_{i\eta i\eta}$ is the dominant source of inefficiency at low- p_T . The distribution of $\sigma_{i\eta i\eta}$ and the thresholds of the *cut-based* IDs are shown in Figure 4.7, which indicates potential in fine-tuning the variable for low- p_T electrons. Conversely, the mismatch efficiency measurements indicate that relaxing this selection also translates into a significant increase in misidentification rates. Taking into account these considerations, ultimately, there are insufficient grounds for deconstructing the standard IDs, which would also require separate validation and determination of all relevant correction factors. Therefore, a choice is made to elect one of the standard *cut-based* IDs. A dedicated sensitivity study calculating the expected limits, using the procedure based on the CL_s criterion (described in Section 8), is performed with the different electron ID working points. Based on the results, together with the efficiency and misidentification rate measurements, the *Veto* ID was

selected, which has the highest efficiency for low- p_T leptons.

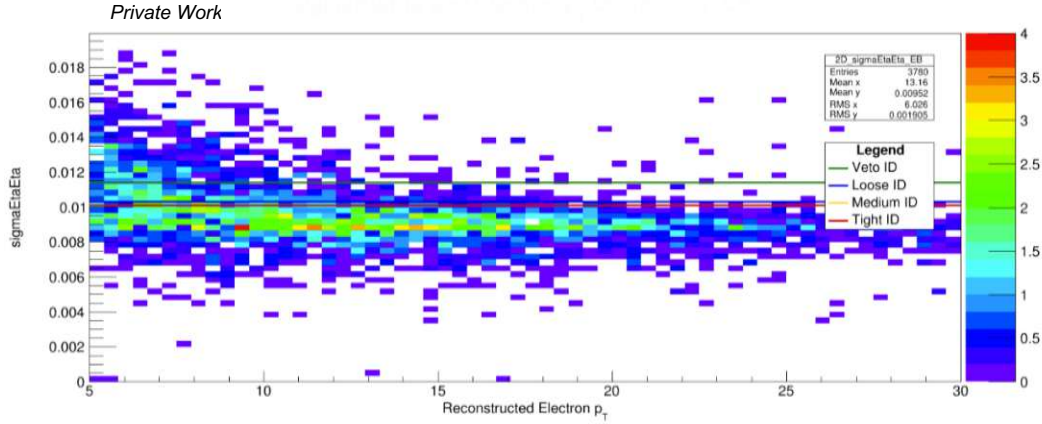


Figure 4.7: Distribution of the cluster shower shape covariance $\sigma_{i\eta i\eta}$ in a T2tt signal sample ($m_{\tilde{t}_1} = 300$ GeV, $\Delta m = 30$ GeV). The lines indicate the thresholds of the *cut-based* electron identification working points (IDs) shown in Table 4.1.

Lepton Isolation and Impact Parameters

Given the difficulties in efficiently identifying soft leptons, specific isolation and impact parameter requirements are tailored for very low- p_T leptons. Since the analysis selects leptons in a non-standard regime, dedicated measurements of the efficiency and simulation correction factors are performed for the final leptonic selection (identification, isolation and impact parameter), as discussed in Section 6.3.2.

Hybrid Isolation

Isolated leptons characteristic of prompt lepton decays from electroweak gauge W or Z bosons can be distinguished from those produced in jets using the isolation observable. This is achieved by applying a maximum requirement on the isolation variable that can be defined in several ways.

The *absolute isolation* (I_{abs}) of leptons provides a measure of the surrounding hadronic energy, defined by the sum of p_T of charged (h^\pm) and neutral (h^0) hadrons and photons (γ) within a cone around the lepton defined by a $\Delta R\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ parameter (typically chosen between 0.3 and 0.5), as shown in Equation 4.2. The summation proceeds over PF candidates, excluding leptons and charged particles not compatible with the PV. The contribution of the neutral particles from PU is subtracted [211]. In this analysis, the distance parameter is chosen as $\Delta R < 0.3$.

$$I_{\text{abs}} = \left[\sum_{h^\pm}^{\text{charged hadrons}} p_T^{h^\pm} + \sum_{h^0}^{\text{neutral hadrons}} p_T^{h^0} + \sum_{\gamma}^{\text{photons}} p_T^{h^0} \right]_{\Delta R < 0.3} \quad (4.2)$$

The *relative isolation* is defined as the I_{abs} normalised with respect to the lepton p_T : $I_{\text{rel}} = I_{\text{abs}}/p_T(\ell)$ and accounts for the fact that higher momentum leptons are also associated with higher energy deposits in its vicinity.

The analysis uses a combined *hybrid isolation* (HI) variable defined as:

$$\text{HI} = I_{\text{rel}} \cdot \min(p_T(\ell), 25 \text{ GeV}) \equiv \begin{cases} I_{\text{abs}} < 5 \text{ GeV} & \text{for } p_T(\ell) \leq 25 \text{ GeV} \\ I_{\text{rel}} < 0.2 & \text{for } p_T(\ell) > 25 \text{ GeV} \end{cases} \quad (4.3)$$

which corresponds to a switch from $I_{\text{abs}} \rightarrow I_{\text{rel}}$ at $p_T = 25 \text{ GeV}$. This results in a looser and p_T -independent isolation requirement on isolation for soft leptons that are targeted by the search, increasing the efficiency in this regime.

Impact Parameters

In order to ensure that the leptons are prompt and thus compatible with the primary vertex (PV), an additional selection on the impact parameter (IP) is applied. The IP is defined as the point of closest approach of the track associated to the lepton with the PV. The following selection is applied on the transverse (d_{xy}) and longitudinal (d_z) components of the lepton IPs, with the selection optimised for low- p_T leptons:

$$\begin{aligned} |d_{xy}| &< 0.02 \text{ cm} \\ |d_z| &< 0.1 \text{ cm} \end{aligned} \quad (4.4)$$

Summary

The final physics object selection of the *soft single-lepton analysis* is summarised in Table 4.2, including the acceptance in terms of minimum p_T thresholds and η ranges, as well as the dedicated identification working points (WPs or IDs).

Object	Acceptance		ID	Details	
	p_T	η			
Muons	$p_T > 3.5 \text{ GeV}$	$ \eta < 2.4$	Loose	HI	$ d_{xy} < 0.02 \text{ cm}$
Electrons	$p_T > 5 \text{ GeV}$	$ \eta < 2.5^2$	Veto		$ d_z < 0.1 \text{ cm}$
Taus	$p_T > 20 \text{ GeV}$	$ \eta < 2.4$			PF
Jets	$p_T > 30 \text{ GeV}$	Soft: $p_T > 30 \text{ GeV}$	$ \eta < 2.4$	Loose	AK4
B-Tagged Jets		Hard: $p_T > 60 \text{ GeV}$		Medium	CSVv2
ISR Jet	$p_T > 100 \text{ GeV}$	$ \eta < 2.4$			Leading Jet
p_T^{miss}					PF

Table 4.2: Definition of final object selection used in the *soft single-lepton analysis* presented in Part II.

²Excluding ECAL crack between EB and EE: $1.4442 < |\eta_{SC}| < 1.566$

4.3 Simulation

The simulation of events is a crucial element in the majority of studies related to the experiment, both on the experimental and theoretical sides. This spans from understanding the performance of physics objects and determination of corresponding corrections, estimations of background SM processes to evaluations of various uncertainty systematics in an analysis. Furthermore, the simulation of postulated BSM models forms the baseline for designing and optimising a dedicated search and is ultimately used in the test and interpretation of signal hypotheses against the background. Simulation corrections that are applied at the analysis level are covered in Section 6.3.2.

Event Generation and Simulation

The signal and background processes are simulated using Monte Carlo (MC) techniques in several steps, ranging from the generation of the hard scatter of the proton-proton collision to hadronisation and showering, detector response and final physics object reconstruction [212]. An illustration of the various steps in MC simulation are shown and described in Figure 4.8, demonstrating the intrinsic complexities of the procedure.

The simulation of the underlying hard-scatter process is non-trivial due to the composite structure of protons, comprising a sea of additional virtual quarks and gluons at high energies. This is made feasible by factorising and separating the treatment of the processes at different energy scales, corresponding to different momentum transfers [214]. The hard scatter is considered at the largest involved scale and is viewed in terms of scattering events of the underlying partons (quarks and gluons) that are asymptotically free. The parton distribution functions (PDFs) give the probability of finding partons in a hadron as a function of the momentum fraction of the proton's momentum carried by the parton.

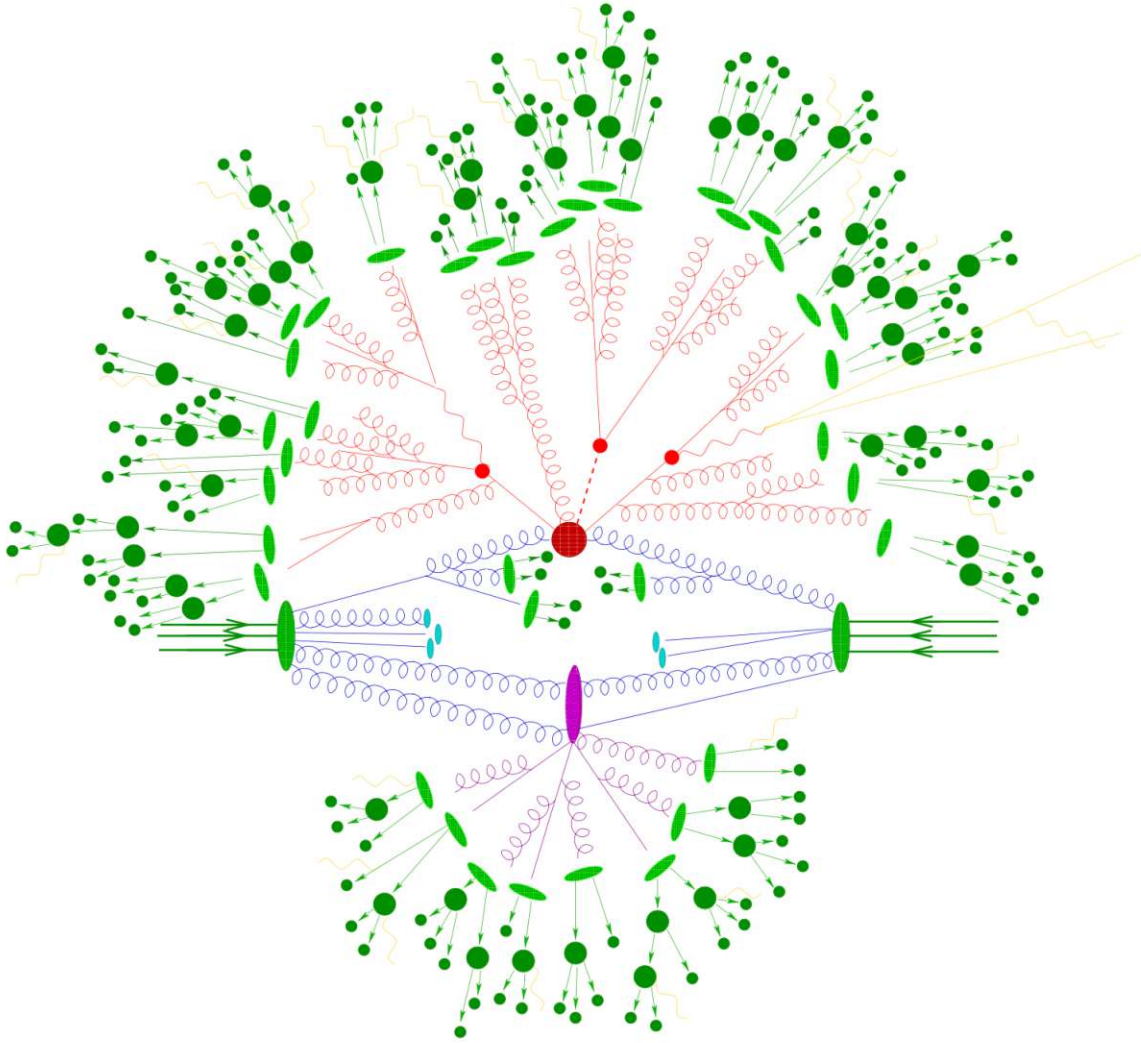


Figure 4.8: Illustration of the different components in a MC generation of a proton-proton collision, resulting in the production of a $t\bar{t}H$ event. Valence quarks of the colliding protons (big green ellipses) are represented by three arrows. Two partons from the incoming protons interact in a hard process (large red circle). Before this hard interaction, both partons undergo initial-state radiation. This follows with decays of top quarks and the Higgs boson (small red circles). Parton showers evolve from the hard scattering process. A secondary interaction between proton remnants is drawn as big purple ellipse and other beam remnants are displayed as small blue blobs. Together they form the underlying event. Hadronisation and color-neutral hadrons are indicated by small light-green ellipses, with corresponding hadron decays (dark-green circles). Photon radiation (yellow) occurs at any stage. [213]

The cross section of the hard process at a hadron collider is then given as:

$$\sigma = \int dx_1 f_{q/p}(x_1, \mu_F^2) \int dx_2 f_{\bar{q}/\bar{p}}(x_2, \mu_F^2) \hat{\sigma}(x_1 p_1, x_2 p_2, \mu_F^2, \mu_R^2) \quad (4.5)$$

where $f_{q/p}(x, \mu_F^2)$ are the PDFs depending on the momentum fraction x with respect to the parent hadron and the factorisation scale μ_F . The parton-level cross section $\hat{\sigma}(x_1 p_1, x_2 p_2, \mu_F^2, \mu_R^2)$ is computed using the matrix elements describing the underlying physical interaction and is dependent on the factorisation and renormalisation scales, μ_F and μ_R , respectively. A choice of the renormalization and factorisation scales is $\mu_R = \mu_L = Q^2$, where Q^2 is the mass scale of a process (e.g. $Q^2 = m^2$ for a particle of mass m produced in the s-channel).

The PDFs are determined using global fits of theoretical QCD predictions with data from a number of collider experiments (including DIS at HERA [74, 75] and Tevatron), such as those of the NNPDF collaboration [215, 216]. The matrix elements are generated using perturbation theory at leading order (LO) or next-to-leading order (NLO), with generators such as MADGRAPH5_AMC@NLO [217, 216] and POWHEG [218, 219]. The next simulation step of the parton showering [220], where soft gluons are emitted from the scattering partons, and their subsequent hadronisation, is performed using the PYTHIA8 [221, 222] generator, with the CUETP8MI tune defining the underlying event parameters. In simulation, the PV is chosen to be at the centre of the CMS coordinate system, where vertex smearing software is used to model the natural spread of the interaction region.

The interaction of the resulting particles with the detector is simulated with the GEANT4 [223] program, taking into account the materials and geometry of the detector. This step of the event generation, referred to as FULLSIM, is the most time and resource-consuming and is used for the generation of SM backgrounds. For large-scale signal sample production, a faster FASTSIM simulation tool [224] is used, with a simplified geometry and material interaction models, in which the detector response is parametrised. This is followed by the digitisation of detector readout electronics, in which the effects of pileup can be simulated. Ultimately, the final physics objects are reconstructed with the same algorithms as data and output into a common data format.

Event Generation Chain

The event generation and simulation chain comprises several sequential steps, ranging from the generation of the hard scatter of the proton-proton collision to hadronisation and showering, detector response and final physics object reconstruction, which is then output in standard CMS software (CMSSW) event data model (EDM) formats:

- Generation and Simulation
 - MG: generation of MADGRAPH5 gridpacks containing details of signal process
 - LHE: generation of Les Houches Event (LHE) or SLHA (SUSY Les Houches Accord) files for the definition of the matrix element
 - GEN: physics event generation, fragmentation, decay and hadronisation using PYTHIA8
 - SIM: CMS detector simulation using GEANT4 or FASTSIM
 - includes detector conditions data, such as alignment and calibrations, under a *Global Tag*
- Trigger and Pileup
 - DIGI: digitisation of detector electronics response
 - L1: simulation of the Level-1 Trigger (L1T)
 - DIGI2RAW: conversion of digitised signals into the RAW data format
 - special conditions added (e.g. realistic PU profile)
 - HLT: running the High Level Trigger (HLT)
- Reconstruction
 - RAW2DIGI: conversion of RAW data format into digitised signals
 - RECO: full event reconstruction

- Data Formatting
 - output in a AOD format with analysis objects
 - output in a miniAOD with reduced size
 - output in a *flat* nanoAOD format with further size reduction and standard physics object definitions

Part II

Compressed Stop Search

Models and Signature

5.1 SUSY Searches

A focus of the physics programme at the Large Hadron Collider (LHC) [49, 50], CERN, is physics beyond the standard model (BSM), which includes searches for supersymmetry (SUSY). As discussed in Section 2.2, there exist a number of theoretical frameworks describing supersymmetric extensions to the standard model, with the minimal extension (MSSM) a leading framework. Many MSSM models require R-parity conservation, resulting in pair-produced SUSY particles (discussed below in Section 5.1.2). Thus, they are predicted to have decay chains to SM particles that can then be detected by the ATLAS and CMS detectors [137, 136]. LHC searches focus on phenomenologically viable models motivated by theoretical considerations that are constrained by experimental measurements. As a result, the relevant phenomenological properties are described in terms of a few physical parameters, which allows a simple characterisation of the main features, instrumental in defining a strategy for discovery.

5.1.1 Simplified Signal Models

Given the large numbers of free parameters in complete SUSY models (discussed in Section 2.2) it can be challenging to design and optimise searches, as well as interpret the results in a meaningful way. In order to overcome this complexity and guide the design of analyses, a phenomenological approach is used, which solely introduces the

minimal number of additional parameters within a simplified model spectrum (SMS) [225, 226] framework. The parameters are in the form of cross sections, masses and coupling strengths and effectively result in a reduced subset of new BSM particles. The new particle content is considered decoupled from the theory, setting the masses of other sparticles at a much higher mass scale and thereby excluding their contributions to the interactions. Furthermore, such simplified models include additional assumptions on the branching fractions and lifetimes of the particles. The models typically include a single decay topology that is considered. Ultimately, such an approach enables the constraint on select model parameters, which are typically the masses of the sparticles under consideration. This facilitates direct comparison with other analyses and re-interpretation of the results in terms of other BSM models with similar final states.

A wide range of different simplified models are considered that are theoretically and phenomenologically motivated. They are categorised depending on the production mechanisms and subsequent decay modes. The production modes are divided into the underlying interaction, which could be strong or electroweak. The possible decay modes are governed by the Δm between the produced particles and the lightest decay products, which defines the available energy. In R-parity conserving models, with the decay chain ending with the lightest supersymmetric particle (LSP), the Δm is typically defined between the produced sparticles and the LSP. The simplified models used in the analysis focusing on strong production are shown in Figure 5.1 and are explained in more detail below.

5.1.2 Stop Production and Cross Sections

As discussed in 2.2.2, a specific model with a light stop \tilde{t}_1 that is mostly right-handed is considered, which is expected to be accessible by the LHC. It is motivated by naturalness and flavour considerations. Furthermore, the particular scenario where the mass spectrum is highly *compressed*, with an almost *degenerate* stop and bino-like $\tilde{\chi}_1^0$ LSP, is strongly motivated by cosmological dark matter relic density predictions.

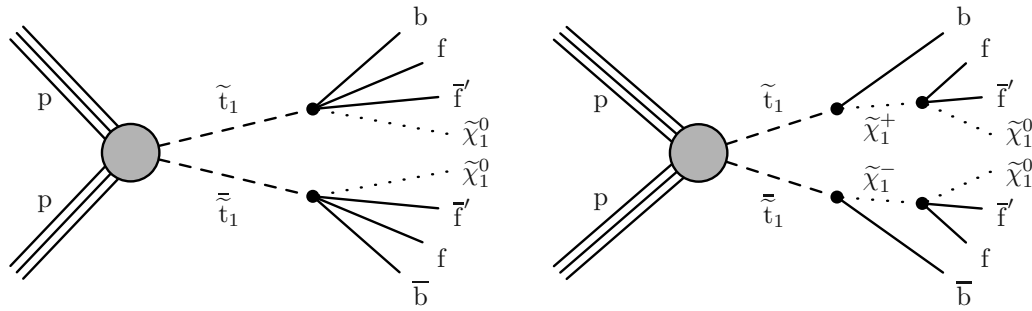


Figure 5.1: Simplified signal models for top squark pair-production with subsequent four-body (left) and chargino-mediated (right) decays into a lepton-neutrino (quark-antiquark) pair, b-quark jet and a neutralino LSP ($\tilde{\chi}_1^0$). The chargino ($\tilde{\chi}_1^\pm$) mass is assumed to be halfway between the stop and LSP. [227]

Further phenomenological motivations include the uncovered parameter space due to the difficulty of performing such an analysis in terms of experimental limitations and background separation.

R-Parity

An important dynamical conservation law in the SM is the conservation of the baryon number (B) and total lepton number (L), which have been confirmed experimentally throughout. Strictly speaking, the conservation is not assumed a priori but is a consequence of no possible renormalisable Lagrangian terms that violate the quantum numbers. Therefore, rather than imposing this requirement in the MSSM, another symmetry is added that eliminates possible B- or L-violating terms. The MSSM is defined to conserve *R-parity*, which is defined as:

$$P_R = (-1)^{3(B-L)+2s} \quad (5.1)$$

where s is the spin of the particle.

Ultimately it translates to SM particles and Higgs bosons having even R-parity ($P_R = +1$) and all SUSY sparticles have odd R-parity ($P_R = -1$). If one imposes R-parity conservation, the LSP is stable and cannot decay. Each other sparticle must then eventu-

ally decay into a state containing at least one LSP. Furthermore, another implication is that sparticles must be pair-produced in collider experiments, such as the LHC. The most copiously produced sparticles at the LHC are expected to be the strongly-interacting squarks \tilde{q} and gluinos \tilde{g} .

Stop Production

The analysis focuses on stop pair-production ($\tilde{t}_1\tilde{t}_1$) with a neutralino $\tilde{\chi}_1^0$ LSP. In hadron colliders, stops (\tilde{t}_1 or \tilde{t}_2) can be produced in pairs at the lowest order QCD in quark-antiquark annihilation and gluon-gluon fusion processes [228, 229], shown in Figure 5.2. At the LHC, the dominant (> 90 %) production mechanism is via the fusion of gluons:

$$pp \rightarrow gg \rightarrow \tilde{t}_i\tilde{t}_i + X \quad (5.2)$$

where either of the stops are pair-produced ($i = 1, 2$) in association with other particles (X). Mixed pairs between \tilde{t}_1 and \tilde{t}_2 cannot be produced at the lowest order, and at higher orders, their rates are highly suppressed.

SUSY Cross Sections

LHC analyses require accurate knowledge of the theoretical predictions for the cross sections of SUSY processes, which are essential for LHC searches, ranging from analysis design to setting exclusion limits. The production modes are divided into the underlying interaction, which could be strong (gluinos $\tilde{g}\tilde{g}$, squarks $\tilde{q}\tilde{q}$) or electroweak (charginos $\tilde{\chi}_1^\pm$, neutralinos $\tilde{\chi}^0$, sleptons $\tilde{\ell}$).

The cross sections are calculated by summing up the possible processes in perturbation theory, which is performed up to next-to-next-to-leading order (NNLO) in the strong coupling constant α_s . Calculations take into account the parton distribution

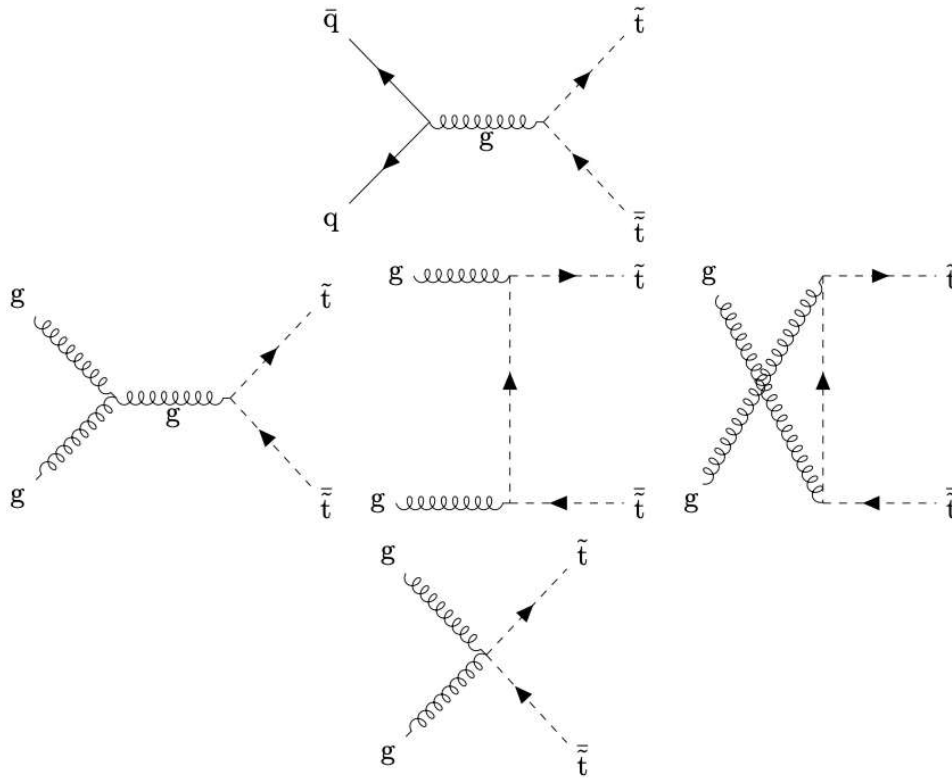


Figure 5.2: Feynman diagrams for the pair-production of top squarks at leading order, including quark-antiquark annihilation and gluon-gluon fusion [229].

functions (PDFs) (discussed later in Section 4.3) and are based on the resummation of soft-gluon emission at an accuracy up to next-to-next-to-leading-logarithmic (NNLL) [230]. They are mainly dependent on the sparticle masses and the collider centre-of-mass energy \sqrt{s} . The corresponding theoretical uncertainties are due to the variation of the renormalisation and factorisation scales, the PDFs and the value of α_s . Most calculations follow the assumption of SMS, where all other irrelevant sparticles are assumed to be heavy and decoupled from the theory.

Beyond leading-order (LO), the stop production cross section $\sigma(pp \rightarrow \tilde{t}\tilde{t}^*)$ depends not only on the stop mass but also on the masses of the gluino and the other squarks. There also exists a dependence of the cross section on the stop mixing angle $\theta_{\tilde{t}}$ (Equation 2.11). However, their effect on the total cross section has been shown to be small. The cross sections for pair-production of light stops $\tilde{t}_1\tilde{t}_1^*$ used in this analysis are calculated at NLO+NLL in a simplified SUSY scenario, with degenerate stop and gluon masses and

other squark flavours decoupled. The inclusive cross section is considered to depend only on the stop mass $m_{\tilde{t}_1}$ and varies approximately between 20 and 0.1 pb for the considered range of masses.

Electroweak production involves electroweakinos ($\tilde{\chi}_1^\pm - \tilde{\chi}_2^0$, $\tilde{\chi}_1^+ - \tilde{\chi}_1^-$, $\tilde{\chi}_1^0 - \tilde{\chi}_2^0$) and sleptons ($\tilde{\ell}^+ \tilde{\ell}^-$), reflecting the SMS that are discussed in more detail in Section 9. In most cases, the produced EWKinos are assumed to be mass-degenerate, and specific assumptions are made on their nature being wino- or higgsino-like. The dependence of the cross sections on the handedness of the slepton is also calculated. The latest cross section calculations [231, 232] for strong [233, 228, 234, 235, 236, 237] and electroweak [238, 239, 240] SUSY production at the LHC centre-of-mass energy of $\sqrt{s} = 13$ TeV are shown in Figure 5.3, as a function of the sparticle mass. Strong processes have the highest cross sections, led by gluino production followed by squark pair-production, including stop pair-production. Due to their nature, electroweak processes overall have lower cross sections at the LHC, with the lowest coming from direct slepton pair-production.

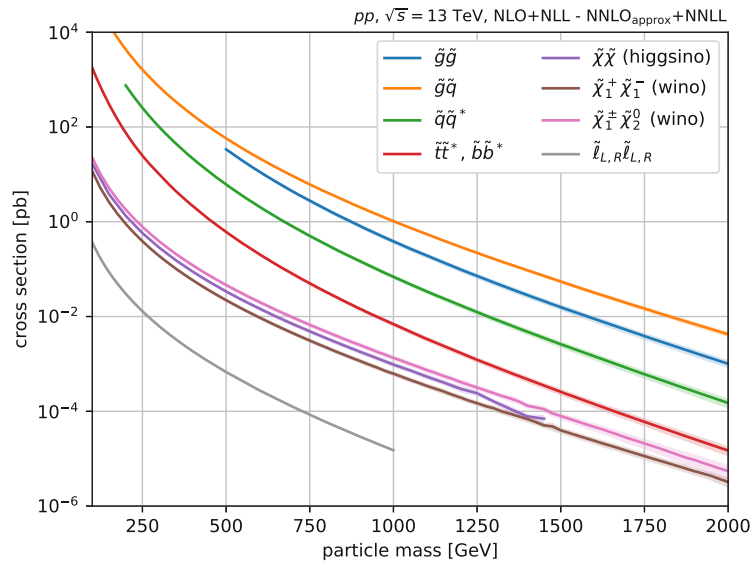


Figure 5.3: cross section predictions for various strong and electroweak SUSY processes in pp collisions at $\sqrt{s} = 13$ TeV, calculated at up to an approximate next-to-next-to-leading order ($\text{NNLO}_{\text{approx}}$), including next-to-next-to-leading logarithmic (NNLL) resummed threshold corrections. [231]

Stop Decay Modes

The possible decay modes of the stop are governed by the $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0) = m_{\tilde{t}_1} - m_{\tilde{\chi}_1^0}$ between the stop $m_{\tilde{t}_1}$ and the neutralino LSP $m_{\tilde{\chi}_1^0}$, as shown in Figure 5.4:

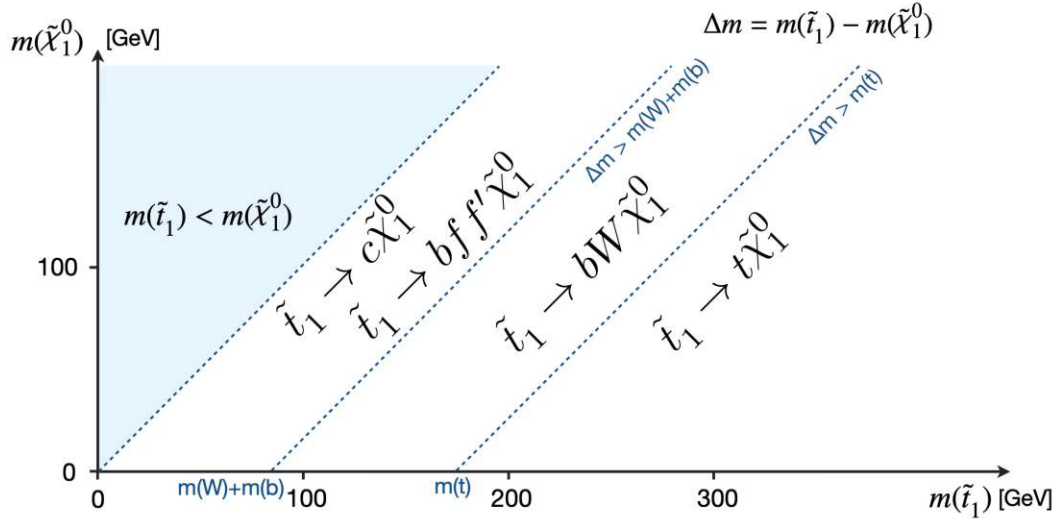


Figure 5.4: Top squark pair-production mass plane of stop (\tilde{t}_1) vs. LSP ($\tilde{\chi}_1^0$), indicating possible decay chains in the simplified model spectrum [241].

- $\Delta m > m_t$ (or $m_{\tilde{\chi}_1^\pm} > m_{\tilde{t}_1}$): two-body decays with an on-shell top quark:

$$\tilde{t}_1 \rightarrow t \tilde{\chi}_1^0 \quad (\text{or } \tilde{t}_1 \rightarrow b \tilde{\chi}_1^\pm)$$

- $m_W < \Delta m < m_t$: three-body decays with an on-shell W-boson:

$$\tilde{t}_1 \rightarrow b W^\pm \tilde{\chi}_1^0$$

- $m_b < \Delta m < m_W$ (shown in Figures 5.1, 5.5, A.1)

- four-body decays with an off-shell top t^* and W^* :

$$\tilde{t}_1 \rightarrow t^* \tilde{\chi}_1^0 \rightarrow b f f' \tilde{\chi}_1^0$$

- four-body decays with an intermediate chargino:

$$\tilde{t}_1 \rightarrow b \tilde{\chi}_1^\pm \rightarrow b f f' \tilde{\chi}_1^0$$

- two-body decay to charm (FCNC):

$$\tilde{t}_1 \rightarrow c\tilde{\chi}_1^0$$

At higher Δm there is more available energy and decay chains with on-shell top quarks or W -bosons are possible, resulting in two- or three-body final states that are targeted by a number of *classical* searches. For low Δm these modes are kinematically forbidden and the decays can only proceed via an off-shell top t^* and W^* , resulting in a four-body final state ($\tilde{t}_1 \rightarrow b f f' \tilde{\chi}_1^0$). The same final state is possible with a decay via an intermediate chargino $\tilde{\chi}_1^\pm$, if its mass is lower than the stop mass ($\tilde{t}_1 \rightarrow b \tilde{\chi}_1^\pm \rightarrow b f f' \tilde{\chi}_1^0$). In this analysis, the highly compressed case when $m_b < \Delta m(\tilde{t}_1, \tilde{\chi}_1^0) < m_W$ is considered, with a focus on the four-body final states. The general simplified models are shown in Figure 5.1. Among the various possible Feynman diagrams contributing to these decays (shown in Figure A.2), the dominant contributions are shown in Figure 5.5.

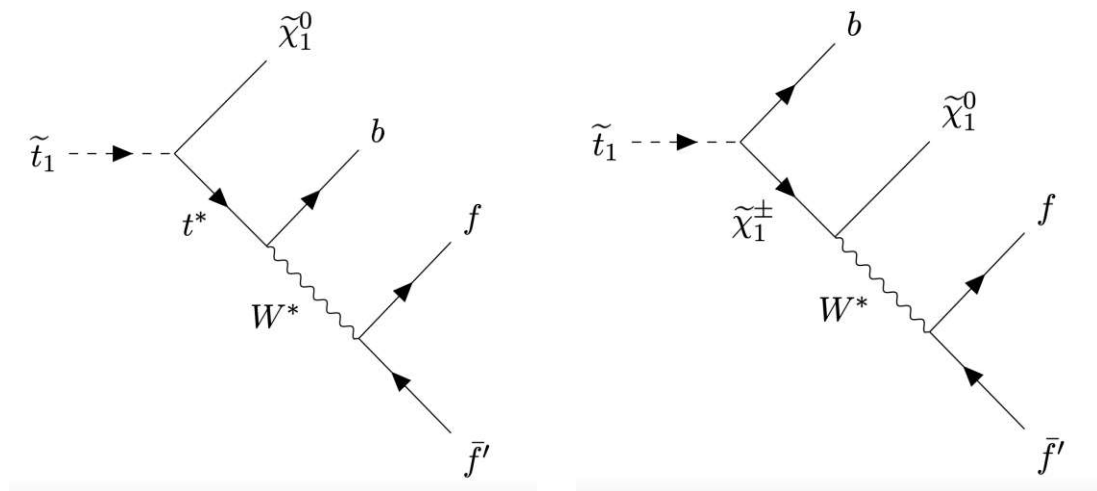


Figure 5.5: Simplified signal models for top squark pair-production with subsequent four-body (left) and chargino-mediated (right) decays into a lepton-neutrino (quark-antiquark) pair, b-quark jet and a neutralino LSP ($\tilde{\chi}_1^0$). The chargino ($\tilde{\chi}_1^\pm$) mass is assumed to be halfway between the stop and LSP. [227]

The fermions f and \bar{f}' can be either leptons or quarks, with the final states reflecting the W -boson decay modes. The branching ratios to quarks and leptons as a function of Δm are shown in Figure 5.6 (left), which reflect the branching ratios of an on-shell

W-boson for higher Δm ; the fully hadronic mode with decays to quarks is dominant ($\approx 66\%$), while each leptonic final states contribute $\approx 11\%$ each. The strategy of the experiment is to cover all possible channels, which in the case of pair-produced stops are: multileptonic, dileptonic, single-leptonic and hadronic. Given the differences in the signatures, different analysis techniques must be implemented, and each channel has a dedicated search. The hadronic channel [242, 243] has higher branching ratios, however, dedicated methods must be implemented to combat the prevalent QCD multi-jet backgrounds. The leptonic channels [5, 244, 245] have cleaner signatures with the requirement of leptons and additional observables, however, they must accept the cost of reduced acceptance due to this requirement. This analysis considers the single-lepton channel [5], so the final state comprises a lepton-neutrino (and quark-antiquark) pair, b-quark jet and a neutralino $\tilde{\chi}_1^0$ LSP.

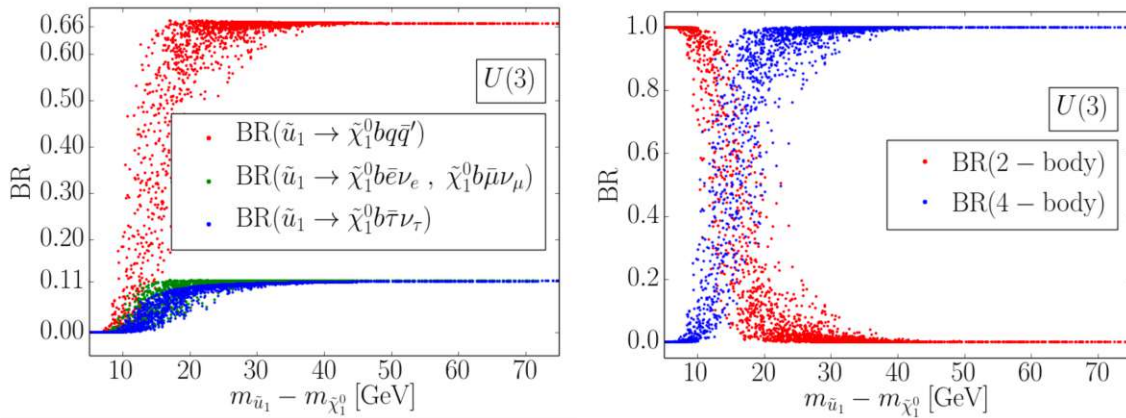


Figure 5.6: Branching ratios of the four-body decay as a function of Δm : to the dominant final states to quarks and leptons (left) and to two- (FCNC) and four-body decays; assuming a $U(3)$ symmetry in the left-handed squark sector [246].

Flavour-Violation and Decay Width

A possible competing process to the standard four-body decay involves a two-body decay to a charm quark ($\tilde{t}_1 \rightarrow c \tilde{\chi}_1^0$), shown in Figure A.1, via a flavour-violating (FV) neutral current (FCNC) interaction [246, 247], which is not forbidden in the MSSM. The branching ratios to the two- and four-body decays as a function of Δm are shown

in Figure 5.6 (right) for a specific flavour-violating scenario. In this case there is almost full dominance of the four-body decays for $\Delta m > 20$ GeV, while for very low $\Delta m < 10$ GeV, the two-body process is expected to dominate fully.

At the lowest Δm , the stops have a small width and consequently a long lifetime, which is also dependent on the possible mixing with the FCNC channel, as shown in Figure 5.6. A total width $\Gamma_{\text{tot}} \approx 10^{-12}$ GeV corresponds to a lifetime (or flight-time of the stop) of the order of pico-seconds, which is long enough to observe displaced vertices in the detector. The proper decay length of the stop could range from a few centimetres to more than a meter. The more the FV couplings are suppressed, the four-body decay dominates, resulting in a smaller total width and, consequently, a longer lifetime. The possible observation of displaced vertices would allow for conclusions on the flavour symmetry of the model.

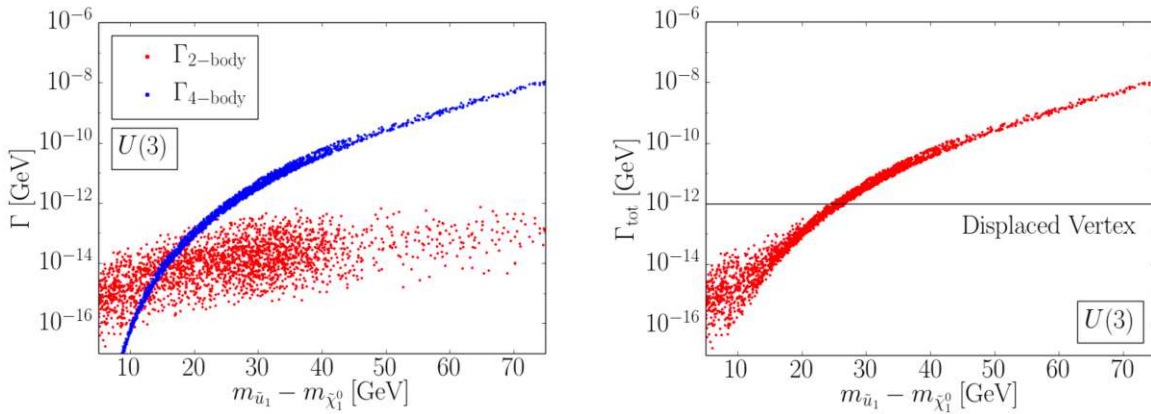


Figure 5.7: The stop two- and four-body decay widths (left) and the total widths (right) as a function of Δm . A total width $\Gamma_{\text{tot}} \approx 10^{-12}$ GeV (black line) corresponds to the value where displaced vertices can be observed. [246]

Since the analysis considers prompt leptons only, the lifetime of the stop \tilde{t}_1 is assumed to be zero. If the decay proceeds with an intermediate chargino, the property of long lifetimes is suppressed. For the chargino-mediated model, the chargino ($\tilde{\chi}_1^\pm$) mass is chosen to be halfway between the stop and LSP. Under the simplified model framework, a 100% branching fraction to the four-body or chargino-mediated decay chains is assumed.

5.2 Compressed SUSY Signature

Models with highly compressed mass spectra have a characteristic *compressed* SUSY signature distinguishing them from other SUSY signals. The kinematics are dictated by the level of compression of Δm , which defines the total available energy to the decay products. Thus, the higher the compression, the lower the available energy to the decay products and the lower (or *softer*) their momenta. In this scenario, the visible decay products would typically not pass the detector acceptance thresholds. Furthermore, for compressed models with light neutralinos $\tilde{\chi}_1^0$, the missing transverse momentum p_T^{miss} is also small in magnitude relative to more *classical* models. Therefore, these particular types of searches target a very challenging phase-space.

The difficulty of soft products can be circumvented to a first degree by the requirement of a high momentum jet in the event. The jet can emerge from initial-state radiation (ISR) originating from the hard scattering process of interacting partons. In such events, the system recoils against the ISR-jet in the opposite direction and is Lorentz boosted with respect to the laboratory frame. The massive LSPs receive the bulk of the boost, increasing the p_T^{miss} , while the visible decay products become detectable, albeit remain relatively soft. The decay products generally align together with the boost direction of the LSPs. Due to the momentum imbalance, the presence of the ISR jet is reflected in the signature in the form of high hadronic energy H_T and moderate missing transverse momentum p_T^{miss} . It is important to note that such a requirement comes at the cost of signal acceptance due to the fact that ISR is a second-order process.

Therefore, a typical compressed SUSY signature comprises at least one high momentum jet, moderate p_T^{miss} and H_T , and soft leptons and jets, as demonstrated in Figure 5.8 and is common across the different searches at the LHC.

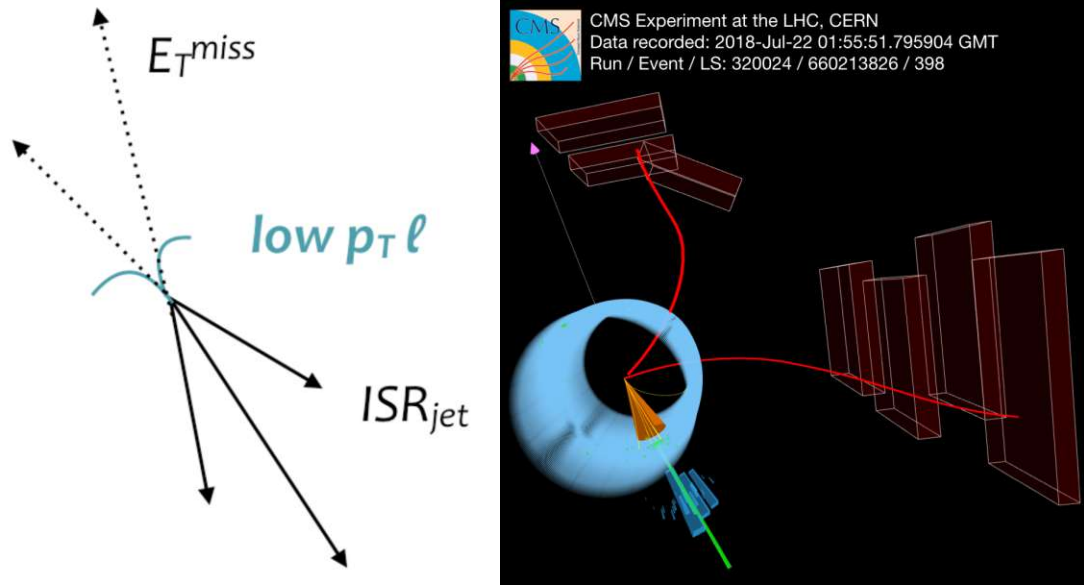


Figure 5.8: Typical initial-state Radiation (ISR) boosted topology with a high- p_T ISR jet, moderate p_T^{miss} and soft leptons and jets. The right image shows a reconstructed CMS event with an energetic jet (orange cone) containing charged particle tracks (yellow) and deposits in the electromagnetic (green) and hadronic (blue) calorimeters recoils against the inferred missing energy (purple arrow). Two low-energy muons (red lines) travel from the interaction point, bending and reversing upon themselves as they pass through alternating magnetic fields before entering the muon system. [248]

In the case of the stop pair-production with the four-body final state in the single-lepton channel, this translates to a signature containing:

- hard ISR-jet
- moderate p_T^{miss}
- moderate H_T
- soft lepton
- soft b-quark jet

Event Selection

6.1 Analysis Timeline

The *soft single-lepton analysis* focuses on stop pair-production with a nearly degenerate neutralino $\tilde{\chi}_1^0$ LSP in the single-lepton channel. The first version of the analysis was performed during LHC *Run 1* with 19.7 fb^{-1} of data and only considered the muon channel [7]. A first iteration of the corresponding LHC *Run 2* analysis with early data (12.9 fb^{-1}) [6] was completed with several improvements, such as the inclusion of the electron channel and optimisation of the final selection. The full 2016 analysis with 35.9 fb^{-1} [3, 4, 5] included a number of developments, such as improved background estimation methods, additional signal region binning, lowering of lepton thresholds as well as an additional interpretation. It was performed in parallel with an MVA [249] BDT-based [250] analysis, with a common preselection, discriminating variables and background estimation methods [4, 5]. It is recently complemented by a full LHC *Run 2* legacy result [251] with 138 fb^{-1} . Apart from the published papers, the analysis has been described in the following thesis [252], as well as several conference proceedings [8, 9, 10]. This thesis focuses on the published LHC *Run 2 cut & count* version of the analysis and potential improvements for LHC *Run 3*. A legacy LHC *Run 2* version of the analysis with 138 fb^{-1} is currently being performed in parallel with a dedicated search for long-lived stops.

6.2 Analysis Framework

The data-analysis framework that is used is based on object-oriented programming (OOP) with the ROOT system [253], which is designed to handle and analyse large amounts of data in an efficient way. Particles are treated as physics objects with given attributes such as their momentum (p_T) or direction (η, ϕ), as well as additional properties (e.g. charge, IP, isolation etc.) determined in the reconstruction process. As described in Chapter 4, the data-processing and reconstruction of objects is performed within the CMS software (CMSSW) collection framework, containing all the detector-relevant information. It is based on the C++ and Python programming languages. CMSSW is built around an event data model (EDM) framework [254, 255] needed by the simulation, calibration and alignment, and reconstruction modules that process event data for analysis. The processed data is ultimately output in relatively lightweight analysis object data (AOD) formats such as miniAOD [256] and nanoAOD [257]. The final data format that is used follows an additional post-processing step performed in Python, with analysis-specific choices (object definitions or filtering), making the subsequent analysis of large datasets more viable.

6.3 Data and Simulated Samples

Data

The sample of events under consideration comes from proton-proton data from *Run II* of the Large Hadron Collider, recorded with the CMS detector at a centre-of-mass energy $\sqrt{s} = 13$ TeV, corresponding to integrated luminosities of 12.9-35.9 fb⁻¹. The details of the specific data-taking conditions are covered in Section 3.1.2. Specific unwanted detector effects (such as HCAL or ECAL noise or beam halo particles) are suppressed by applying dedicated p_T^{miss} filters, which remove the affected events [258].

Simulation

The general specifics of Monte Carlo (MC) simulation of signal and background events are covered in Section 4.3, including the entire simulation chain. Simulated MC samples are produced using various generators and reconstructed with the same algorithms as data. The complete list of simulated samples and their cross sections is shown in Appendix B.1, including SUSY signals based on the calculations described in 5.1.2. The final object selection is summarised in Section 4.2.

6.3.1 Signal Generation

Stop Signal Generation

The information on the masses, widths, branching ratios and cross sections, including the particular stop and neutralino mixing setup, is included in an SLHA file. In the particular case of the compressed stop pair-production, it is important to properly model the additional ISR jets (see e.g. Figure 4.8). One can calculate the matrix element directly with extra jets (in MADGRAPH5) or generate only the simple event and generate extra partons in the parton shower step (in PYTHIA8). The two approaches are complementary, where matrix elements are more accurate for hard and separated partons, while the parton shower description is more accurate when the partons are collinear or soft. Matrix element calculations are computationally expensive, especially when generating extra jets. The standard approach is to implement both methods by generating the undecayed stop pair with up to two extra jets in MADGRAPH5 and then let them decay in PYTHIA8, which is adapted in this case. One caveat is that with this approach there is double counting of jets, which need to be matched with dedicated algorithms [259] that reject double-counted events. Another issue is that PYTHIA8 does not accurately describe the decay topology, where the decay is considered isotropic even off-shell, while there is an expected asymmetry between the lepton and the LSP due to the polarisation of the top quark in stop decays [260]. An example of this dis-

crepancy for a chosen stop polarisation is shown in Figure 6.1 in the distribution of $\cos \theta^*$, where θ^* is defined as the angle between the lepton and LSP in the top t^* rest frame. It is expected that the theoretical uncertainties related to polarisation effects for ISR-boosted regimes are small. Apart from this, the width of the off-shell W^* boson is taken into account as a Breit-Wigner distribution in PYTHIA8. The lifetime of the stop is assumed to be zero, and polarisation effects are neglected. Given the large number of signal points that are scanned per SUSY model, a FASTSIM detector simulation tool [224] is used where the detector response is parametrised, making large-scale signal sample production more feasible.

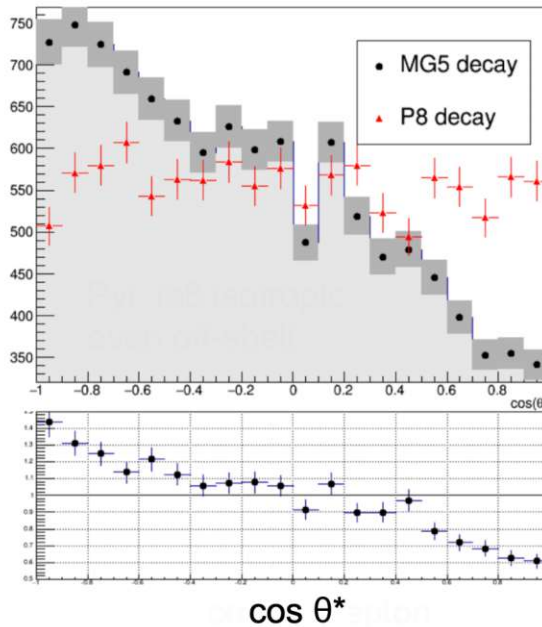


Figure 6.1: Comparison of the distribution of $\cos \theta^*$ between PYTHIA8 and MADGRAPH5. θ^* is defined as the angle between the lepton and LSP in the top t^* rest frame [261].

BSM signal models are scanned over wide ranges of the free model parameters, which, in the case of stop pair-production, are the stop and LSP ($\tilde{\chi}_1^0$) masses. The stop masses are scanned between $250 < m_{\tilde{t}_1} < 800$ GeV in steps of 25 GeV, where the lower bound is chosen based on exclusion results from *Run 1*. The mass splittings with respect to the LSP are chosen between $10 < \Delta m < 80$ GeV in 10 GeV steps, providing coverage from very high compression up to the level of the W boson mass. This amounts to 156 signal points, with different stop and LSP masses, both for the T2tt and T2bW models.

Generator Filters

In the specific case of the compressed SUSY analyses, the ISR requirement significantly reduces the signal acceptance. Together with the requirements on p_T^{miss} and lepton, this means that a significant amount of the generated events are already discarded at the level of preselection (Section 6.4.3.1). In order to increase the efficiency of how many signal events are used in the final selection with respect to the generated events, a dedicated filter has been developed and implemented at the generator level. Furthermore, it also opens up the possibility of generating more events, increasing the statistical precision, especially in depleted signal regions of the analysis (Section 6.4.4).

For standard MC production, a benchmark number of generated signal events is an average of $\approx 100 - 200$ k per signal mass point, which typically translates to a sufficient statistical precision ($< 10\%$) for a projected luminosity of $\approx 50 \text{ fb}^{-1}$. In order to recover the loss in acceptance for compressed signals and retain a similar statistical precision ≈ 500 k events are needed. Therefore, a generator filter with a reduction factor of ≈ 5 reduces the average amount of events to be stored and processed to the benchmark values. For the chosen mass point grid, the total amount of generated template events, taking into account jet-matching (reduction factor ≈ 3), would be ≈ 260 M events, which are then reduced to a final ≈ 30 M after the generation step.

A dedicated study is performed on generator-level quantities in order to achieve the required reduction factor. A focus is placed on observables that give the highest reduction while not differing greatly between generation and reconstruction levels. Furthermore, in order to be more inclusive with respect to other analysis groups targeting the other channels, a focus is made on general kinematic quantities of the compressed SUSY signature, such as the generated p_T^{miss} , p_T^{ISR} and H_T , with no additional selection on leptons.

The generator filter efficiency *turn-on* curve is evaluated as a function of the corresponding reconstructed quantities, as shown in Figure 6.2.

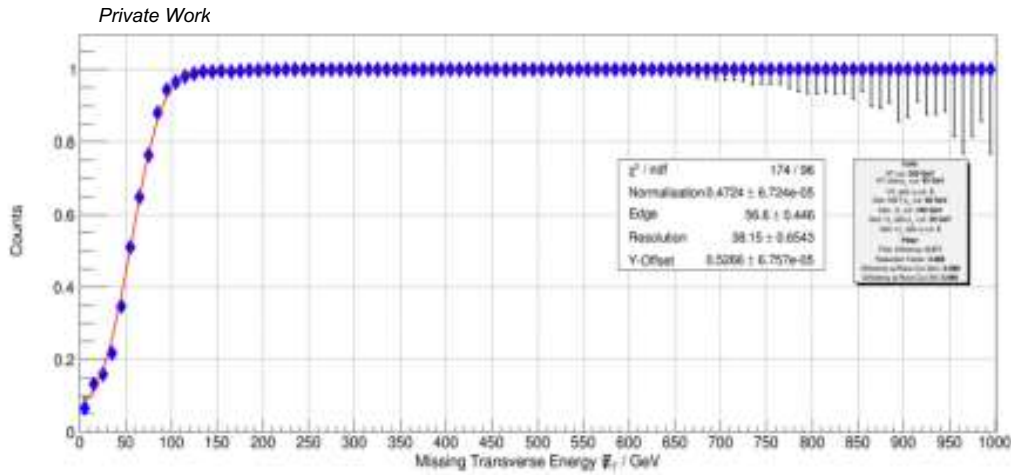


Figure 6.2: Generator filter efficiency as a function of reconstructed p_T^{miss} for a T2tt signal point with ($m_{\tilde{\tau}_1} = 300$ GeV, $\Delta m = 30$ GeV), with a selection on generated $p_T^{\text{miss}} > 60$ GeV and generated $H_T > 160$ GeV.

The shape of the *turn-on* curve reflects the differences between the generated and reconstructed observables, introduced by resolutions of the physics objects across the full simulation chain. It effectively provides a measure of the spectra distortion and the level of inefficiencies. For a given generator filter selection, the plateau of the turn-on (efficiency ≈ 100 %) needs to be achieved before any conceivable analysis selection. A scan was performed over different combinations of selections on several generated variables, with the aim of achieving the desired reduction factor while keeping the inefficiency at the level of analysis selection to a minimum (efficiency > 98 %). The optimisation of the final selection was performed on a single signal mass point ($m_{\tilde{\tau}_1} = 300$ GeV, $\Delta m = 30$ GeV) and the stability of the reduction factor and efficiency values was evaluated over the entire mass scan, taking the lowest reduction factor as reference.

Taking into account the different preselection requirements of other CMS analyses, a choice was made to select on generated p_T^{miss} and H_T . The hadronic energy sum H_T has different definitions across analyses, so a more inclusive (and effectively more relaxed) definition is used, covering most of the jet acceptance: the sum of all generated jets with $p_T > 30$ GeV and $|\eta| < 5$. Among the different sets of proposed cut values, the final filter selects on generated $p_T^{\text{miss}} > 80$ GeV and generated $H_T > 160$ GeV.

A dedicated filtering module was developed and integrated into CMSSW. The filter is implemented after the generator (GEN) step after hadronisation and pre-simulation (SIM). It is used in large-scale SUSY production of the signal grid, used by different CMS analyses. It is also used in the production of other signal models, such as TChiWZ and the higgsino pMSSM discussed in Section 9 (also indicated in Table A.1).

The total generator filter efficiency values for stop pair-production T2tt model are shown in Figure 6.3, which ranges between 0.2 and 0.5. These values are required for the luminosity and cross section reweighting procedure of the signal samples. A module has also been developed to store this information at the nanoAOD level.

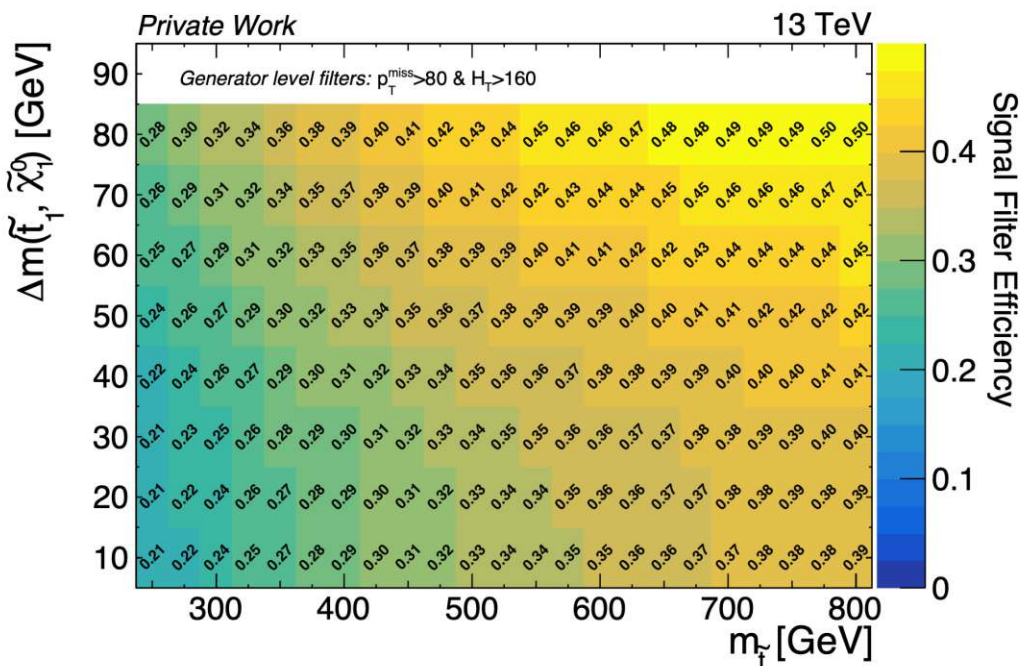


Figure 6.3: Total generator filter efficiency values for the T2tt signal model scan [252].

6.3.2 Simulation Corrections

Given that simulation plays a central role in an analysis, it is vital that the MC samples are corrected in order to reflect the collected data more accurately. This is typically achieved by calculating dedicated data/MC scale factors in well-controlled regions of parameter space (e.g. focusing on specific processes) and then applied at the analysis level.

ISR Modelling

As discussed above, the modelling of ISR is challenging to perform accurately, and it is known to have its limitations (mainly at the level of MADGRAPH5). Since the presence of ISR is an underlying characteristic of compressed signatures, it is important that it is modelled correctly. ISR-related observables in well-controlled regions (e.g. $t\bar{t}$ events or Z -peak) exhibit clear trends in the discrepancy between data and simulation. In particular, the multiplicity of ISR jets (N_{jets}^{ISR}) and the transverse momentum of the ISR jet (p_T^{ISR}) exhibit significant mismodelling. In order to correct for these effects, data/MC correction factors are determined as a function of these observables, which are then applied at the analysis level. These corrections are important for the main backgrounds of the analysis ($W + jets$ and $t\bar{t}$) as well as the signals.

The correction factors are determined as a function of N_{jets}^{ISR} in $t\bar{t}$ events and range between ≈ 0.9 for $N_{jets}^{ISR} = 1$ to ≈ 0.5 for $N_{jets}^{ISR} > 6$. They are applied to the $t\bar{t}$ MC samples, as well as the considered signals, as the production of top squarks is expected to be affected by the same mismodelling, due to the similarities in the topology. For electroweak production, the correction factors are determined in Drell-Yan (Z/γ^*) events on the Z mass peak as a function of p_T^{ISR} . The correction factors range between ≈ 1.2 for $100 < p_T^{ISR} < 150$ GeV to ≈ 0.8 for $p_T^{ISR} > 600$ GeV. In the soft single-lepton analysis, they are applied to the $W + jets$ MC samples as a function of the W -boson p_T , which serves as a proxy for p_T^{ISR} in boosted topologies. They are also used to correct electroweakly-produced SUSY signals, discussed in Section 9, where the correction is

applied as a function of the p_T of the system of pair-produced EWKinos. Additional global normalisation factors are calculated between corrected and uncorrected samples, in order to ensure that the total number of events for a given sample stays constant.

Lepton Scale Factors

The efficiency of the lepton reconstruction and final object selection (identification, isolation and impact parameter, described in Section 4.2) is measured both in data and simulation. The corresponding data/MC ratio (or *scale factors*) are used to correct the simulated samples for any remaining differences between simulation and observation. The scale factors are determined for background (FULLSIM) and signal (FASTSIM) samples. Additional FULLSIM/FASTSIM scale factors need to be taken into account for signal to correct for any residual differences introduced by the simplified FASTSIM detector simulation.

Backgrounds (FULLSIM)

Measurements of efficiencies of the reconstruction and final object selection of leptons are typically performed by exploiting well-measured physical processes ($Z, J/\psi$ decays) with the *tag-and-probe* method [262]. This is typically performed using events with leptonic decays of the Z boson $Z \rightarrow \ell\ell$, which are selected by requiring the invariant mass of the two leptons to be within a window close to the Z boson mass (*Z-peak*). The method requires one of the leptons (*tag*) to satisfy stringent selection criteria, while the other (*probe*) is subject to the selection criteria under investigation (identification, isolation, impact parameter). The invariant mass distributions are fitted with a double Gaussian distribution for the Z-peak and a falling exponential distribution for the background. The ratio of the fits of probe leptons is measured as the lepton efficiency. This procedure is performed for different ranges of p_T and $|\eta|$ for both data and simu-

lation, separately for muons and electrons. The lepton scale factors are calculated as the ratio of the data and MC efficiencies data/MC .

The collaboration provides *central* measurements for the standard reconstruction and identification working points (or *IDs*). However, since this analysis considers very low- p_T leptons in a non-standard regime (Section 4.2), with specific isolation and impact parameter requirements, the scale factors need to be measured independently. They are shown in Figure 6.4 for the hybrid isolation and impact parameter (right), as well as the *loose* and *veto* IDs (left), for muons and electrons, respectively. The muon scale factors are close to unity, and the electron scale factors range down to ≈ 0.9 for the *veto* ID in the lowest electron p_T bins. For $p_T(\ell) > 10$ GeV, they are compatible with the centrally-derived scale factors, while for the very low- $p_T < 10$ GeV regime, the privately measured ones are used.

Signal (FASTSIM)

Apart from the standard (FULLSIM) lepton scale factors, additional $\text{FULLSIM}/\text{FASTSIM}$ scale factors need to be taken into account to correct for any residual differences introduced by the simplified FASTSIM detector simulation used in SUSY signal production. Similarly, due to the non-standard low- p_T regime and specific isolation and impact parameter requirements, the $\text{FULLSIM}/\text{FASTSIM}$ scale factors need to be measured independently.

The efficiency of the combined isolation and impact parameter selection was determined in the standard stop pair-production $T2t\bar{t}$ (FASTSIM) and dedicated signal points with full detector simulation (FULLSIM), and the scale factors are calculated from the ratio of the two. The $\text{FULLSIM}/\text{FASTSIM}$ scale factors are shown in Figure 6.5 as a function of lepton p_T and $|\eta|$ and are measured to be close to unity. These correction factors are used together with the standard lepton scale factors in order to correct the signal samples.

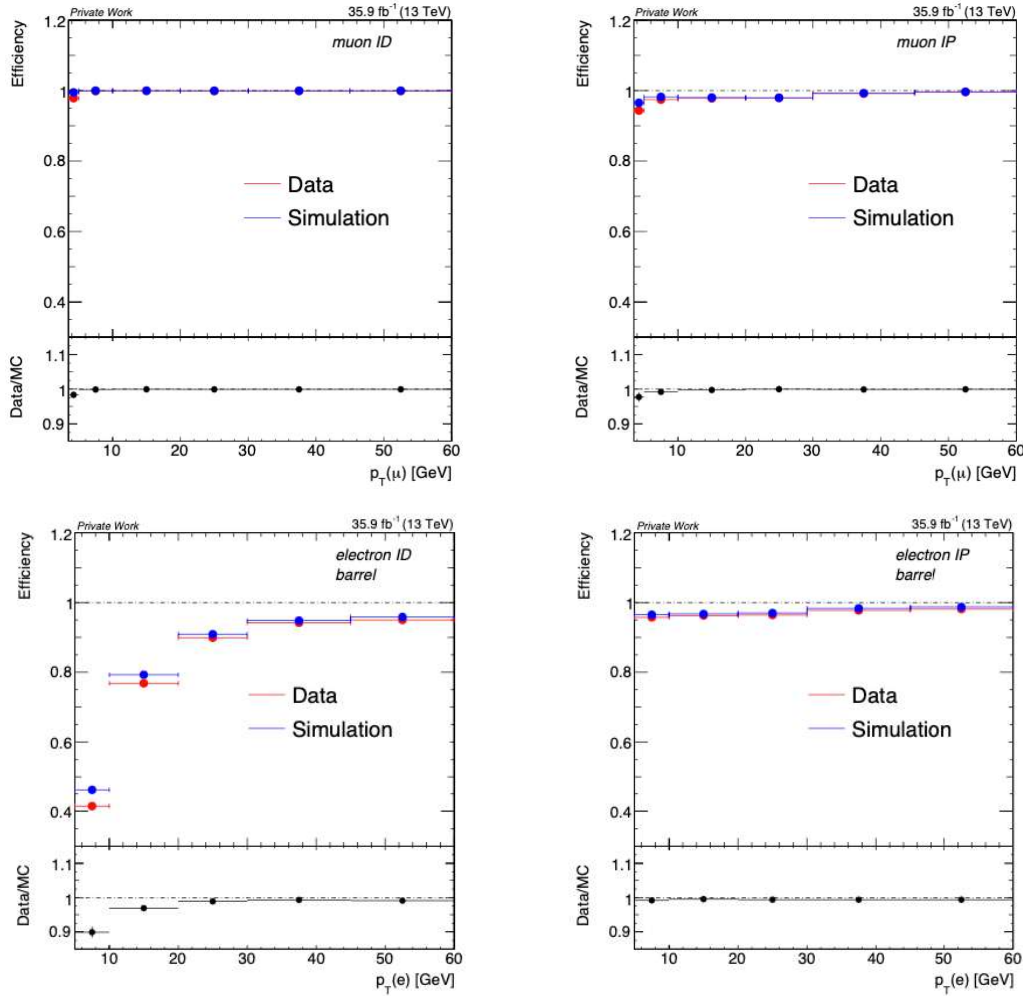


Figure 6.4: Lepton scale factors for ID (left) and hybrid isolation & impact parameter (right) both for muons (top) and barrel ($|\eta_{SC}| \leq 1.479$) electrons (bottom), binned in lepton p_T . They are determined using the *tag-and-probe* method on the Z-peak ($60 < m_{\ell\ell} < 120$ GeV). [263]

B-Tagging Scale Factors

Similarly to leptons, the efficiencies of reconstruction and selection criteria of the b-tagging algorithm working points (described in Section 4.1) are used to determine dedicated data/MC scale factors in order to correct the simulation. The efficiencies are determined for the heavy- and light-flavoured jets as a function of b-jet p_T and $|\eta|$. The analysis relies solely on the multiplicity of b-tagged jets (described in Section 6.4.4 below), so the b-tagging scale factors are determined in the form of weights correspond-

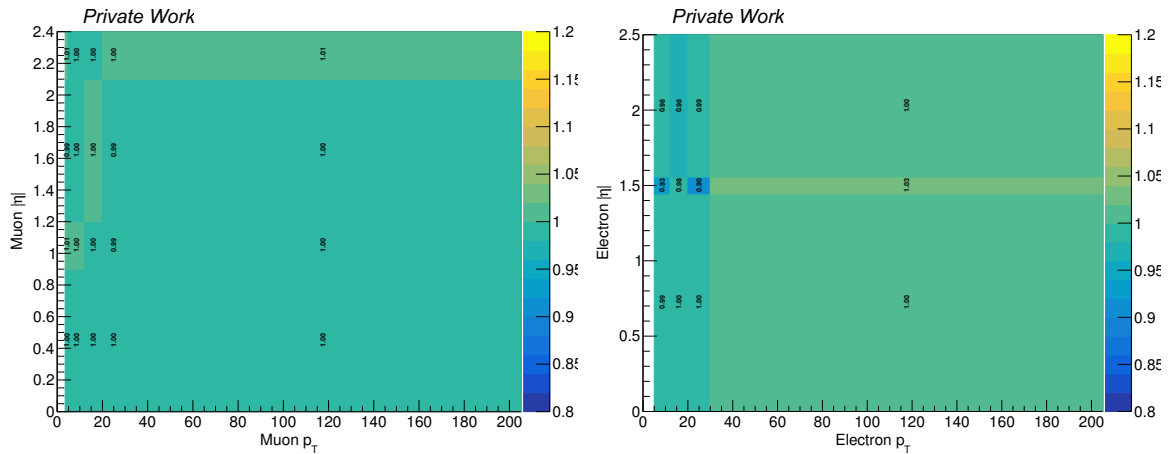


Figure 6.5: 2D $\text{FULLSIM}/\text{FASTSIM}$ scale factors for hybrid isolation & impact parameter for muons (left) and electrons (right) binned in lepton p_T and $|\eta|$.

ing to the probability that the particular event falls into a particular N_{b-jets} category. Corresponding $\text{FULLSIM}/\text{FASTSIM}$ scale factors are also calculated.

Pileup

Furthermore, with the changing data-taking conditions of LHC, pileup effects need to be corrected with the final observed profile, as shown in Figure 3.4. Since this information is not available a priori, the corrections are calculated once the full dataset is collected. The differences between the assumed and observed pileup distributions are calculated as a function of the average pileup per bunch crossing and applied to simulation. Their values are highly dependent on the final profile shapes, especially in the tails of the distributions.

6.4 Event Selection

In searches for new physics at the LHC, events are chosen based on observables of the particular signature that is being sought for, while suppressing the SM backgrounds. The selection is guided by the objective of maintaining sensitivity to a wide range of signal parameters, which in the case of simplified models are the BSM particle masses.

6.4.1 Cut & Count Analysis

A conventional approach that is implemented in this search is a 'cut-&count' analysis. It involves the sequential application of constraints (or *cuts*) on discriminating variables, with the purpose of maximising the signal over the backgrounds.

The yield for a given signal process is given as:

$$N_{sig} = N_{data} - N_{bkg} = \mathcal{A} \times \epsilon_{tot} \cdot \sigma_{sig} \cdot \int \mathcal{L} dt \quad (6.1)$$

where N_{data} is the total number of candidate events in data, N_{bkg} is the number of estimated background events, σ_{sig} is the signal cross section, $\int \mathcal{L} dt$ is the integrated luminosity, and $\mathcal{A} \times \epsilon_{tot}$ is the *acceptance* \times *efficiency*. The total efficiency is defined as:

$$\epsilon_{tot} = \epsilon_{det} \cdot \epsilon_{trg} \cdot \epsilon_{reco} \cdot \epsilon_{analysis} \quad (6.2)$$

and it accounts for reconstruction and identification losses both from the detector and during offline processing, while the acceptance \mathcal{A} is defined as the fraction of events that pass all the cut requirements of the analysis. Therefore, the $\mathcal{A} \times \epsilon$ is a measure of the final number of events that pass the analysis selection.

A set of *preselection* cuts are applied in order to focus on the phase-space of interest, with specific *signal regions* (SRs) chosen to maximise the signal-to-background ratio. *Control regions* (CRs) are designed to allow for precise estimation of the leading backgrounds, while *validation regions* (VRs) are used to verify these estimation methods (Figure 7.1). The CRs and VRs are chosen in such a way that they are dominated by the relevant backgrounds, where the sensitivity to the signal is minimal. The estimation of backgrounds is covered in Section 7. All the regions in the analysis are chosen such that they are orthogonal to one another and do not overlap.

With a 'cut-&count' approach, one essentially performs a counting experiment where the signal and background yields follow a Poissonian distribution. The number of data events in the signal regions is *counted* and compared against the estimated background yields. The final result follows from a statistical analysis based on the likelihood prin-

ciple to calculate the probability of a BSM signal, given the estimated background and observed data, taking all sources of systematic uncertainties into account. The exact details of the method are covered in Section 8.

6.4.2 Trigger

A number of compressed SUSY searches, including the soft single-lepton analysis, rely on an online trigger selection on pure p_T^{miss} , which is compatible with an ISR-boosted signature, discussed in Section 5.2. The usage of such triggers does not impose specific restrictions on leptons, making it suitable for compressed signatures with very soft leptons. In comparison, the lowest unrescaled threshold for single lepton paths is $p_T(\mu) > 24$ GeV (shown in Table 3.3), which is incompatible with leptonic compressed SUSY signatures.

The trigger selection is defined by the following HLT paths $p_T^{\text{miss}} + H_T^{\text{miss}}$ with thresholds ranging between 90 and 120 GeV:

- `HLT_PFMET90_PFMHT90_IDTight`
- `HLT_PFMET100_PFMHT100_IDTight`
- `HLT_PFMET110_PFMHT110_IDTight`
- `HLT_PFMET120_PFMHT120_IDTight`

Pure p_T^{miss} are discussed in Section 3.3.6.4, including the seed collection, history and performances. An independent measurement of the performance of the triggers is additionally performed to assess the efficiency in the phase-space of interest, as well as evaluate the systematic uncertainties associated with the trigger selection. Specific details of the methodology of trigger efficiency measurements are discussed in more detail in Section 10.2.3. The measurement is performed in an orthogonal control sample from the *Single Electron* primary dataset, using data collected with a single-electron trigger (`HLT_Ele27_eta2p1_WPTight_Gsf`) in order to not introduce a bias in the measurement. A requirement on the leading electron $p_T(e) > 30$ GeV is made in order to

be on the plateau of the electron trigger turn-on. An additional selection corresponding to the preselection cuts on the leading jet $p_T^{\text{ISR}} > 100$ GeV and $H_T > 200$ GeV is applied in order to perform the measurement in a phase-space kinematically similar to that of the analysis. The measurement is made as a function of the offline reconstructed PF p_T^{miss} , both for all the thresholds separately as well as a combination of all using a logical 'OR', as shown in Figure 6.6. The efficiency of the separate turn-ons of the two lowest thresholds (90, 100 GeV) do not reach full efficiency, due to the effect of the prescales that were applied during data-taking in order to control the rates. Therefore, the total efficiency of the logical 'OR' is essentially driven by the two higher thresholds (110, 120 GeV).

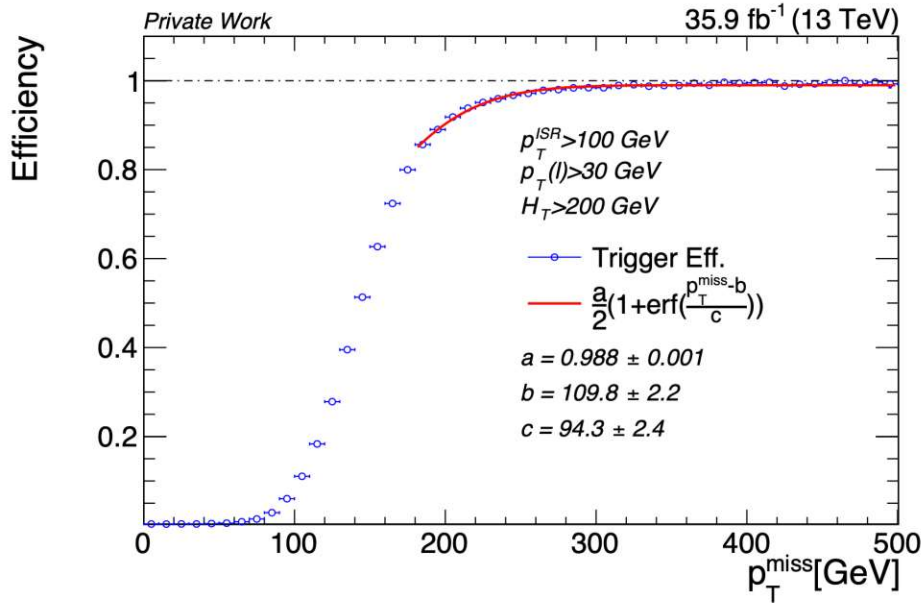


Figure 6.6: Pure- p_T^{miss} and H_T^{miss} trigger efficiency as a function of p_T^{miss} as measured privately by the analysis in the single electron dataset (`HLT_Ele27_eta2p1_WPTight_Gsf`) after a selection on $p_T > 30$ GeV, jet $p_T^{\text{ISR}} > 100$ GeV and $H_T > 200$ GeV. The measurement is made as a combination of all thresholds using a logical 'OR'. [264, 252]

In summary, the independent efficiency measurements are compatible with one another, and for the lowest unprescaled p_T^{miss} threshold of 120 GeV, the plateau of the turn-on is reached at ≈ 250 GeV with an efficiency of $\approx 99\%$. This indicates a reasonable minimum requirement on p_T^{miss} in the signal regions of an analysis. As discussed in the following sections, the data on turn-on is only used in preselection and valida-

tion regions which have a maximum precision of $\mathcal{O}(10\%)$. At the level of preselection of $p_T^{\text{miss}} > 200$ GeV and $H_T > 300$ GeV, the efficiency is $\approx 90\%$. The signal and control regions of the analysis were chosen well beyond the plateau, with a requirement on $C_T > 300$ GeV that translates to $p_T^{\text{miss}} > 300$ GeV, ensuring low systematic uncertainties from the trigger selection. The combined efficiency is parametrised with an error function (Equation 10.9), as shown in Figure 6.6, that is then used to correct the simulated MC samples in all signal and control regions. Such a parametrisation is preferred over using the emulated decision in MC, as it is less sensitive to any data/MC differences on the turn-on. In order to account for possible biases in the measurement method, a systematic uncertainty of 1% is applied.

Further details of trigger-related concepts are covered extensively in Section 10, in the context of improvements of the trigger strategy.

6.4.3 Preselection and Backgrounds

6.4.3.1 Preselection

In order to focus on the relevant parameter space, a set of *preselection* criteria are applied. In the first stage, requirements are made on kinematic features of the targeted signature, such as on energy sums and multiplicity of jets. This is followed by requirements on leptons and background-suppressing cuts.

The requirement of missing transverse momentum p_T^{miss} (introduced in Section 4.1) is a characteristic of many BSM searches. Many SUSY models include stable and weakly-interacting LSPs, which escape detection and manifest themselves as a p_T^{miss} signature. Such a requirement significantly suppresses dominant QCD multi-jet backgrounds present in hadron colliders. In ISR-like topologies, the p_T^{miss} is boosted to moderate levels. A jet compatible with ISR is selected by requiring the leading jet in an event to have momentum $p_T > 100$ GeV and pseudorapidity $|\eta| < 2.4$. This simple definition is appropriate as non-ISR signal jets are not expected to have high momenta due to the low mass-splittings Δm . The presence of an ISR jet is reflected in the form of high hadronic energy H_T , defined as the scalar sum of the p_T of all jets with $p_T > 30$ GeV

and pseudorapidity $|\eta| < 2.4$. Due to the momentum imbalance in ISR-like topologies, the two observables p_T^{miss} and H_T are thus strongly correlated.

The leading jets of the dominant backgrounds in the analysis ($W + jets$ and $t\bar{t}$, covered below in Section 6.4.3.2) are typically softer than for signal, which is reflected in slightly lower hadronic activity. In order to provide some separation, a cut of $H_T > 300$ GeV is applied. This separation is discussed in more detail in the context of a simultaneous selection on p_T^{miss} and H_T in the below Section 6.4.4.3. Furthermore, the backgrounds are expected to have a higher jet multiplicity with several energetic jets. Therefore, further background suppression is achieved by requiring events to have at most two jets with $p_T > 60$ GeV.

Compressed SUSY signatures include very soft leptons and their p_T s are dependent on the Δm . As discussed in Section 4.2, muons with $p_T > 3.5$ GeV and electrons with $p_T > 5$ GeV are considered. The single-lepton topology is selected by requiring at least one soft light lepton in the event, while vetoing hadronically-decaying taus and additional light leptons with $p_T > 20$ GeV.

The dominant QCD multi-jet backgrounds in hadron colliders are already significantly reduced with the requirement of a lepton and relatively high p_T^{miss} . QCD-induced processes are predominantly produced in a dijet configuration with a back-to-back topology. Therefore, to further reduce residual contributions, a maximum requirement on the angle between the leading and sub-leading jets with $p_T > 60$ GeV, $\Delta\phi(j_1, j_2) < 2.5$ rad is imposed.

Other key observables used in the analysis, such as transverse mass m_T or requirements on b-jets, are exploited in the final SR selections, which are discussed in Section 6.4.4. A summary of the preselection requirements is shown in Table 6.1. The distribution of the lepton p_T at the level of preselection is shown in Figure 6.7.

The plots show the shape distributions of the main backgrounds in the analysis, with a comparison against benchmark signal models, following a set of reweighting procedures discussed in Section 6.3.2. A good agreement between data and simulation is

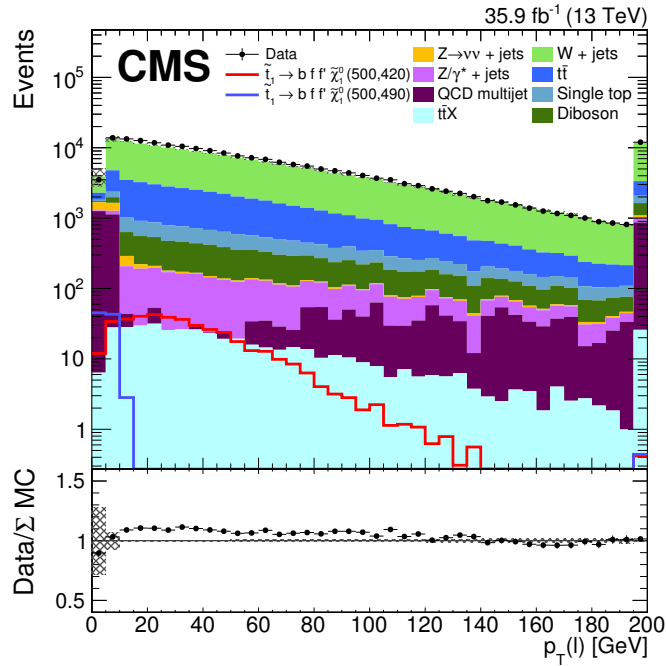


Figure 6.7: Distributions of the lepton p_T at the level of preselection. Two T2tt ($\tilde{t}_1 \rightarrow b f f' \tilde{\chi}_1^0$) benchmark signal points are chosen with $m_{\tilde{t}_1} = 500$ GeV and $\Delta m = \{10, 80\}$ GeV. The first bin only considers muons from 3.5 GeV, causing the visible drop. The last bin in the plot includes events beyond 200 GeV. [227]

seen, taking into account that the ratio only considers statistical uncertainties (as the systematic uncertainties are not propagated to the preselection level). The normalisation of the simulation with respect to data is additionally corrected by the results of the data-driven prompt background estimation technique for the main backgrounds. The method itself is not sensitive to the overall normalisation, as the $^{\text{data}}/\text{MC}$ scaling proceeds from the high to low p_T region, as described in Section 7.1.

6.4.3.2 Backgrounds

In hadron colliders, the environment is dominated by QCD multi-jet processes, with many jets and, consequently, high hadronic energy. They are strongly suppressed already at the level of preselection, with the requirement of a lepton and relatively high p_T^{miss} , as well as the dedicated anti-QCD cut $\Delta\phi(j_1, j_2) < 2.5$ rad. Following preselection, the main residual backgrounds relevant to the analysis are shown in Figure 6.7 and comprise mainly electroweak processes.

The dominant backgrounds in the analysis that closely mimic the targeted signals are the production of a W -boson in association with extra jets ($W + jets$) and the pair-production of top quarks ($t\bar{t}$). Although in most of the cases, W -bosons and top-pairs decay hadronically, the W -boson has a 20% chance to decay into light leptons (electrons and muons) via $W \rightarrow \ell\nu$ and, consequently, in approximately 30% of the cases $t\bar{t}$ decays semi-leptonically via a real W ($t \rightarrow Wb$) [265]. Therefore, both backgrounds contain leptonically-decaying W -bosons via the electroweak interaction, resulting in final states containing real and promptly-produced leptons and genuine p_T^{miss} generated from the accompanying neutrinos. Moreover, the decay chains of the strongly-produced signals $T2tt$ ($\tilde{t}_1 \rightarrow bff'\tilde{\chi}_1^0$) and $T2bW$ ($\tilde{t}_1 \rightarrow b\tilde{\chi}_1^\pm \rightarrow bff'\tilde{\chi}_1^0$) themselves contain leptonic decays of t and correspondingly W , as discussed in 5.1.2. The presence of extra jets for $W + jets$ and the hadronically-decaying leg of the stop pairs $t\bar{t}$, lead to moderate H_T values, while the p_T^{miss} can also be artificially enhanced due to the mismeasurement of the jet momenta. The Feynman diagrams of some dominant production modes and leptonic channels are shown in Figure 6.8. The processes have relatively large cross sections at the LHC, which are up to 5 orders of magnitude larger than stop pair-production and at the level of preselection, they make up approximately 70% and 20% of the total expected SM backgrounds, respectively. Both the $W + jets$ and $t\bar{t}$ backgrounds are estimated simultaneously with the method described in Section 7.1.

Similar processes that include top quark decays are also relevant, albeit to a much lesser extent, as a result of lower cross sections as well as acceptance, due to the different topologies. These processes include the production of a single top quark (Single Top (ST)), as well top pair-production in association with an extra boson $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}\gamma$ (collectively labelled as $t\bar{t}X$).

Other rarer processes include decays of EWK bosons into real leptons, such as Drell-Yan (DY) processes (Z/γ^*) or the production of two bosons (WW , WZ , ZZ), collectively labelled VV . It can happen that one of the leptons is not correctly reconstructed (and thus *lost*), generating a single-leptonic signature.

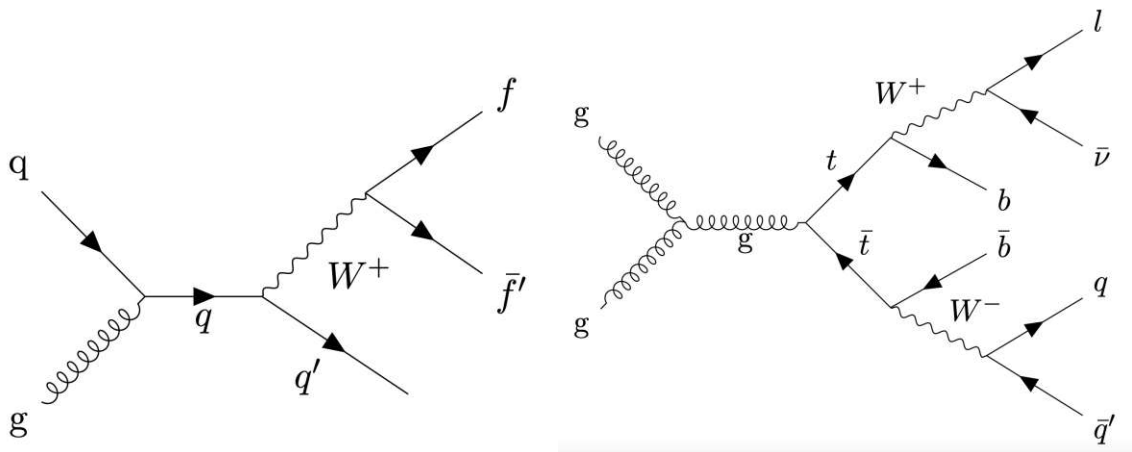


Figure 6.8: Feynman diagrams of leading backgrounds: possible of W-boson production mode in association with jets ($W + jets$) and dominant gluon-gluon fusion production of top quark pairs ($t\bar{t}$), decaying via real Ws into lepton-neutrino pairs.

In the case of VV, such a signature can also result via a combination of leptonically- and hadronically-decaying vector bosons.

Even though some processes do not contain real prompt leptons, they can still emerge as significant backgrounds in the analysis, passing all the reconstruction, identification and isolation criteria. These backgrounds arise when a reconstructed lepton is misidentified (also referred to as *fake*) or not promptly produced (referred to as *non-prompt*). Fake leptons mostly arise when quark- or gluon-initiated jets or hadrons are misidentified as leptons. Hadrons from energetic particle showers can *punch-through* the calorimeters and reach the muon chambers and be misreconstructed as muons. Electrons can be misidentified in the ECAL from neutral pions or unidentified photon conversions. Real nonprompt leptons originate from the semi-leptonic decays of heavy-flavour (c, b) quarks, as well as hadron decays, such as light-meson decays in flight. Despite the different possible sources and varying origins, they are estimated simultaneously with the methods described in Section 7.2. Thus, in this thesis, they are interchangeably referred to as nonprompt (or *fake*) leptons, and refer to both cases simultaneously.

Specific to this analysis, QCD multi-jet processes are a source of nonprompt leptons, where one of the jets is misidentified as a lepton. Simultaneously, the mismeasurement of the momenta of other jets in the event can lead to the generation of fake p_T^{miss} that passes the analysis selection. Another significant nonprompt contribution is the decay of the Z-boson to neutrinos (Z_{inv}), where the neutrinos generate significant genuine p_T^{miss} and a nonprompt lepton can be misreconstructed in association with one of the jets. Furthermore, for prompt backgrounds (e.g. $W + jets$ and $t\bar{t}$), there exists the possibility that a real lepton is lost while a fake one is found, which augments this background type.

Therefore, the main backgrounds in this analysis are categorised between prompt and *nonprompt* (or *fake*)¹. The leading prompt and sub-leading nonprompt backgrounds are explicitly estimated using separate methods, discussed in Section 7.²

6.4.4 Signal Regions

Following the preselection requirements, the total SM backgrounds are larger by approximately a factor 10^3 with respect to the benchmark signals, as seen in Figure 6.7. In order to further separate the signal from the backgrounds, a set of additional criteria exploiting the main observables is applied to define the signal regions (SRs). The regions are sub-divided (or *binned*) in order to maintain sensitivity to a wide range of SUSY particle masses and Δm .

6.4.4.1 Lepton p_T and B-Jet Multiplicity

The analysis targets very compressed mass spectra with low Δm and, consequently, very soft leptons, so the SRs are primarily defined by requiring the lepton $p_T < 30$ GeV. The control regions (CRs) used in the main background estimation, discussed in Section 7.1, are defined by inverting the lepton $p_T > 30$ GeV.

Given that the kinematics are strongly dependent on the mass splittings, in order to provide sensitivity to different Δm , two general sets of SRs are created, labelled SR1 and

²Light leptons from purely leptonic tau decays are considered prompt in this analysis.

SR2. SR1 targets very low Δm where the b-jets are expected to be too soft to be detected and are therefore vetoed. In this region, the $W + jets$ background dominates further, while the b-jet veto significantly reduces the $t\bar{t}$ background. SR2 requires at least one soft b-jet with $30 < p_T < 60$ GeV, targeting higher mass splittings. Conversely, the $W + jets$ background is suppressed with the b-jet requirement, while the $t\bar{t}$ background becomes dominant in SR2.

As demonstrated in Figure 6.7, there is a very strong dependence of the lepton p_T signal spectrum on the Δm . Therefore, all SRs are further split in terms of the lepton p_T : low (L): 5-12 GeV, medium (M): 12-20 GeV and high (H): 20-30 GeV. There is an additional very low (VL) bin with $3.5 < p_T < 5$ GeV for muons, added in low- m_T regions only. This splitting allows to retain sensitivity to a wide range of Δm , while enhancing the signal-to-background ratio.

6.4.4.2 Transverse Mass m_T

The transverse mass m_T of semi-invisibly decaying particles is calculated from the transverse momenta p_T s of their decay products, allowing their identification and measuring their masses. It can be computed from an object's p_T and p_T^{miss} as

$$m_T \equiv \sqrt{2p_T^{\text{miss}}p_T(1 - \cos \Delta\phi)} \quad (6.3)$$

where $\Delta\phi$ is the angular difference between the object \vec{p}_T and \vec{p}_T^{miss} . In semi-invisible decays of dominant SM processes that include W-bosons, such as $W + jets$ and $t\bar{t}$, the transverse mass m_T spectrum exhibits a kinematic edge in the form of a Jacobian peak around the mass of the W-boson m_W . The distributions of the transverse mass m_T at the level of preselection are shown in Figure 6.9, clearly showing the kinematic edge beyond $m_T \approx m_W \simeq 80$ GeV.

In order to isolate this peak, an additional splitting in m_T (a: $m_T < 60$ GeV, b: $60 < m_T < 95$ GeV and c: $m_T > 95$ GeV) is done for all SRs, which further improves the sensitivity of the search. The low- m_T regions with a, b : $m_T < 95$ GeV are dominated by the prompt $W + jets$ and $t\bar{t}$ backgrounds, whereas beyond the kinematic edge c :

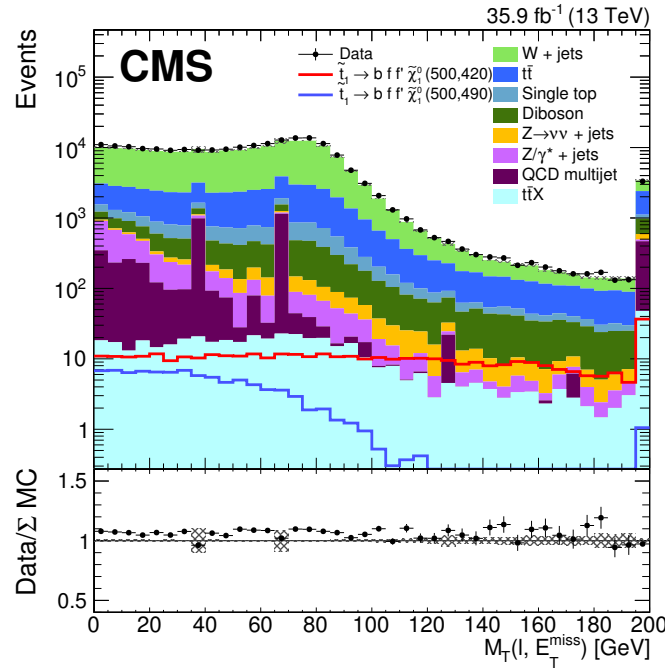


Figure 6.9: Distributions of the transverse mass m_T at the level of preselection [227]. Two T2tt ($\tilde{t}_1 \rightarrow b f f' \tilde{\chi}_1^0$) benchmark signal points are chosen with $m_{\tilde{t}_1} = 500$ GeV and $\Delta m = \{10, 80\}$ GeV.

$m_T > 95$ GeV in the tails of the m_T distribution, the nonprompt backgrounds become significant. Since $m_T \propto \sqrt{p_T^{\text{miss}} p_T}$, low m_T regions are more sensitive to signals with low Δm and correspondingly lower p_T and lepton p_T , as seen in Figure 6.9. This further justifies the chosen splitting and is the reason why the VL lepton p_T bin is added for low m_T regions only.

6.4.4.3 C_T Variable

Regarding the kinematic selection on energy sums, a simultaneous selection on p_T^{miss} and H_T (or p_T^{ISR}) is performed, which takes advantage of the strong correlation between the variables in compressed SUSY signatures.

The selection is made by introducing a new variable C_T , which has slightly different definitions in terms of the hadronic observable used depending on the signal region:

- $C_{T1} \equiv \min(p_T^{\text{miss}}, H_T - 100 \text{ GeV})$
- $C_{T2} \equiv \min(p_T^{\text{miss}}, p_T^{\text{ISR}} - 25 \text{ GeV})$

where the subscript number corresponds to the SR the requirement is applied in.

At a first level, the C_T variable selects ISR-like events where $p_T^{\text{miss}} \sim H_T$ and avoids strange topologies with high p_T^{miss} and low H_T , or vice-versa. In SR2, a selection on the ISR jet momentum p_T^{ISR} is more effective than H_T in reducing the $t\bar{t}$ background, which is dominant in that region. Moreover, as discussed in Section 6.4.3.1, the signals have slightly lower hadronic activity than the dominant backgrounds, motivating the offset in the diagonal for H_T (or p_T^{ISR}). The exact values of the offsets were chosen following a numerical optimisation maximising a Figure-of-Merit (FoM) based on the signal-to-background yields.

The signal regions all have a minimum requirement on $C_T > 300$ GeV that includes on $p_T^{\text{miss}} > 300$ GeV, which is well above the plateau of the pure p_T^{miss} triggers that are used. Both SRs are split into two bins of C_T : $300 < C_T < 400$ GeV and $C_T > 400$ GeV, labelled X and Y, respectively.

There is a further discussion on the C_T selection in Section 10.5 in the context of a new trigger strategy targeting a *low-MET* region; this includes Figure 10.8 indicating the characteristic “L” shapes in the 2D plane of the kinematic variables, as well as 10.9 presenting 2D distributions of benchmark signals against the dominant backgrounds.

6.4.4.4 W-Boson Rapidity and Charge Asymmetry

The production of W-bosons at the LHC shows specific features in terms of its rapidity, helicity, differential cross sections and charge [266], primarily related to the PDFs of the colliding protons (Section 4.3). These features can be exploited in BSM searches to suppress the backgrounds further. W-boson decays exhibit a charge asymmetry with a preference towards positive values of W_s , while in stop pair-production, the charged leptons are produced symmetrically. Therefore, in SR1 regions with low $m_T < 95$ GeV, dominated by $W + jets$, the leptons are required to be negatively charged $Q(\ell) = -1$. Furthermore, leptons from W-boson decays are more often produced in the forward direction, so the pseudorapidity is restricted to $|\eta| < 1.5$ in SR1. These restrictions are not applied in SR2, which has lower acceptance and is dominated by $t\bar{t}$.

Variable	Preselection					
p_T^{miss}	$> 200 \text{ GeV}$					
H_T	$> 300 \text{ GeV}$					
p_T^{ISR}	$> 100 \text{ GeV}$					
N_{jets}^{hard}	$\leq 2 (p_T > 60 \text{ GeV})$					
$\Delta\phi(j_1, j_2)$	$< 2.5 \text{ rad}$					
Lepton requirement	$\geq 1 \mu$ with $p_T > 3.5 \text{ GeV}$ or e with $p_T > 5 \text{ GeV}$					
Lepton veto	veto on τ and additional ℓ with $p_T > 20 \text{ GeV}$					
Common SR Selection						
C_T	$> 300 \text{ GeV}$					
p_T^{miss}	$> 300 \text{ GeV}$					
	SR1			SR2		
H_T	$> 400 \text{ GeV}$			$> 300 \text{ GeV}$		
p_T^{ISR}	$> 100 \text{ GeV}$			$> 325 \text{ GeV}$		
N_{b-jets}	0			$N_{b-jets}^{\text{soft}} \geq 1 (30 \leq p_T \leq 60 \text{ GeV})$		
$ \eta(\ell) $	< 1.5			< 2.4		
	SR1a	SR1b	SR1c	SR2a	SR2b	SR2c
m_T	$< 60 \text{ GeV}$	$60\text{--}95 \text{ GeV}$	$> 95 \text{ GeV}$	$< 60 \text{ GeV}$	$60\text{--}95 \text{ GeV}$	$> 95 \text{ GeV}$
$Q(\ell)$	-1		± 1	± 1		

Table 6.1: Definition of preselection and general signal region requirements, indicating splitting in m_T with specific requirements on lepton η and charge $Q(\ell)$.

6.4.5 Summary

To summarise the final analysis selection, there are two main signal regions: SR1 and SR2, targeting lower and higher Δm , respectively. They are mainly separated with the requirement of a soft b-jet, as well as some differences in the kinematic cut definitions in $C_T \propto p_T^{\text{miss}}, H_T, p_T^{\text{ISR}}$. Each signal region is further divided into bins of m_T , lepton p_T and C_T , with additional restrictions on lepton η , charge $Q(\ell)$ in $W + jets$ dominated regions. The general signal region requirements are presented in Table 6.1, whereas the final signal region binning for both SR1 and SR2, together with the corresponding control region definition, is shown in Table 6.2.

Signal Region	C_T (GeV)	m_T (GeV)	$p_T(\ell)$ (GeV)	Control Region
SRVLaX	300–400	< 60	3.5–5	CRaX
SRLaX			5–12	
SRMaX			12–20	
SRHaX			20–30	
SRVLbX		60–95	3.5–5	CRbX
SRLbX			5–12	
SRMbX			12–20	
SRHbX			20–30	
SRLcX		> 95	5–12	CRcX
SRMcX			12–20	
SRHcX	20–30			
SRVLaY	> 400		< 60	
SRLaY		5–12		
SRMaY		12–20		
SRHaY		20–30		
SRVLbY		60–95	3.5–5	CRbY
SRLbY			5–12	
SRMbY			12–20	
SRHbY			20–30	
SRLcY		> 95	5–12	CRcY
SRMcY			12–20	
SRHcY	20–30			

Table 6.2: Final signal region binning for both SR1 and SR2, indicating splitting in C_T , m_T and lepton p_T , together with the corresponding control region definition. The very low (VL) lepton p_T bin considers muons only.

Ultimately, there are 44 SRs, which are labelled:

$$\text{SR}[1, 2] [\text{VL}, \text{L}, \text{M}, \text{H}] [\text{a}, \text{b}, \text{c}] [\text{X}, \text{Y}]$$

indicating the splitting in the different observables and 12 CRs in total, mirrored in the lepton p_T cut at the 30 GeV threshold.

Background Estimation Methods

The estimation of the dominant backgrounds in the analysis that mimic the signal is performed using various background modelling methods, implementing combined approaches using both data and simulation (*data-driven*). Dedicated control regions (CRs) are designed such that they are dominated by the targeted backgrounds. The determination of correction factors between data and simulation in these CRs allows making a precise extrapolation back to the signal regions (SRs). Background components with real prompt and nonprompt (or *fake*) leptons in the final state are estimated separately with specifically tailored methods. The remaining sub-leading backgrounds are estimated directly using simulation, with a 50 % uncertainty on the cross section.

7.1 Prompt

The dominant backgrounds in the analysis originate from the $W+jets$ and $t\bar{t}$ processes, with real leptons that are promptly-produced in the final state. Since the backgrounds contain real leptons, they are considered *irreducible*, as they cannot be reduced using improved or more stringent lepton identification criteria. Therefore, a precise and accurate method of estimation of these backgrounds is vital.

As discussed in Section 6, the control regions (CRs) are defined by inverting the lepton p_T selection ($p_T > 30$ GeV). Dedicated $data/MC$ transfer factors (TFs) are determined in these CRs and then applied to the simulation in the SRs in order to give the estimate of the final background contribution in the SRs. In other words, the $data/MC$ scaling is

extrapolated from a high p_T to a low p_T region.

The number $N_{\text{prompt}}^{\text{SR}}$ of estimated prompt background events in the signal region determined via:

$$\begin{aligned} N_{\text{prompt}}^{\text{SR}}(X) &= N_{\text{MC}}^{\text{SR}}(X) \times TF^{\text{CR}} \\ &= N_{\text{MC}}^{\text{SR}}(X) \times \left[\frac{N_{\text{data}}^{\text{CR}} - N_{\text{nonprompt}}^{\text{CR}} - N_{\text{rare}}^{\text{CR}}}{N_{\text{MC}}^{\text{CR}}(X)} \right] \end{aligned} \quad (7.1)$$

where X corresponds to the background processes: $W + jets$ or $t\bar{t}$. The data/MC transfer factor TF^{CR} is determined in the CRs, by subtracting the nonprompt and rare contributions from data and taking the ratio to the total number of prompt simulated $W + jets$ or $t\bar{t}$ events. The nonprompt contributions are estimated via the *tight-to-loose* method described below in Section 7.2.3, while the rare processes are taken directly from simulation. The TF is then used to correct the simulated yields in the SRs to yield the final estimates.

In order to validate this method, the estimation procedure is repeated in dedicated validation regions (VRs). The validation regions are designed to be enriched in the particular background that is being estimated. Furthermore, they are chosen to be kinematically similar while not overlapping with the SRs and CRs, as well as having little sensitivity to the signals. The validation regions for the prompt background estimation method are shown in Figure 7.1. One of the VRs (VW) is defined by lowering the C_T selection to $200 < C_T < 300$ GeV, which is kinematically similar to the main SRs with similar background compositions. The other VR (VB) requires at least one hard b-jet with $p_T > 60$ GeV ($N_{b\text{-jets}}^{\text{hard}} \geq 1$), enriching the sample in $t\bar{t}$ events.

The predictions in all validation regions are compared to the observed events to perform a closure test; As shown in Figure 7.2, the predictions in all regions are compatible with the data within the uncertainties, thereby validating the method.

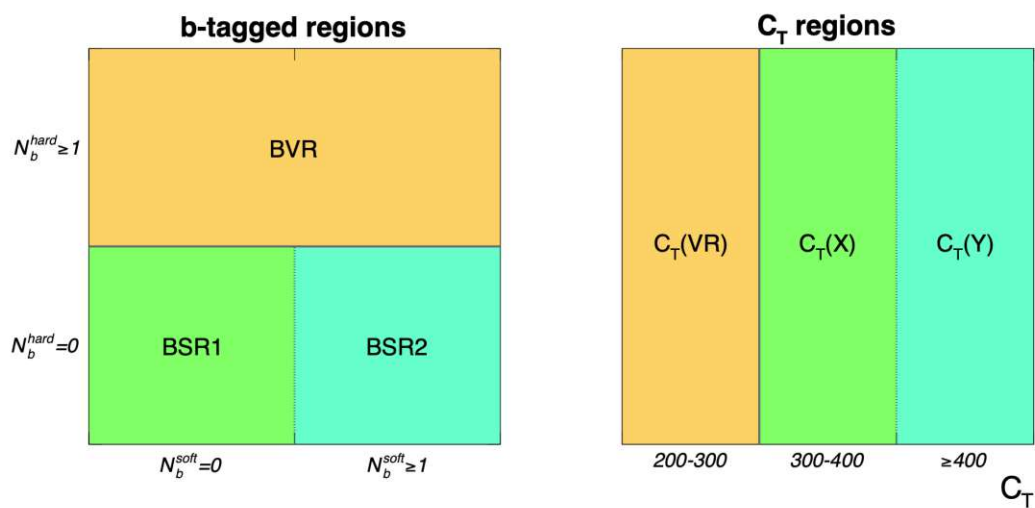


Figure 7.1: Schema of the validation regions (VRs) in terms of $N_{b\text{-jets}}^{\text{soft}}$, $N_{b\text{-jets}}^{\text{hard}}$ and C_T , relative to the SRs and CRs [252].

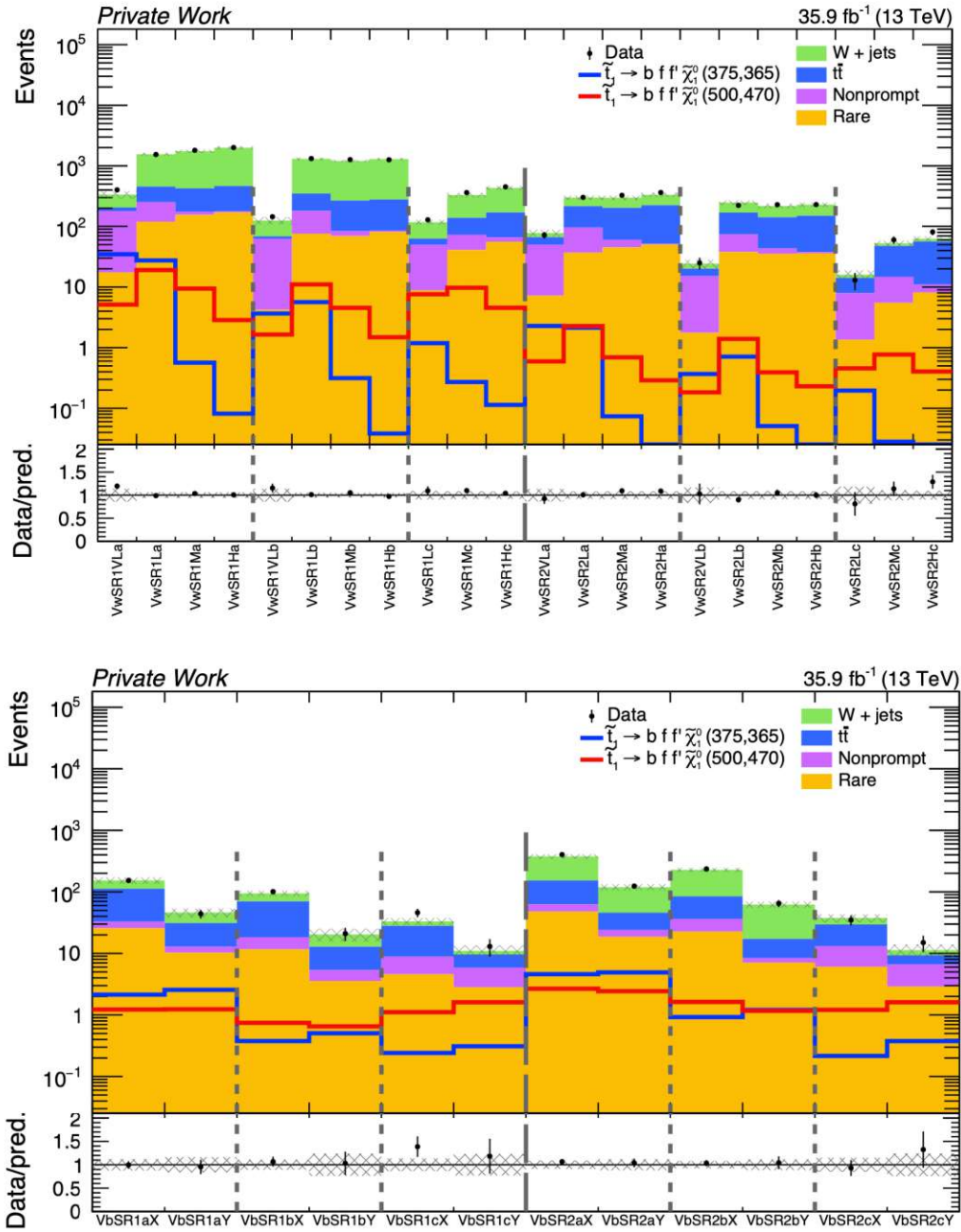


Figure 7.2: Closure test of prompt background estimation method in validation regions VW (top) and VB (bottom), indicating compatible predictions within uncertainties and validation of the method. [252]

7.2 Nonprompt

Events with nonprompt leptons together with moderate p_T^{miss} form a non-negligible background in the analysis. These backgrounds arise when a reconstructed lepton is misidentified (also referred to as *fake*) or not promptly produced (referred to as *non-prompt*). Fake leptons mostly arise when quark- or gluon-initiated jets or hadrons are misidentified as leptons. Hadrons from energetic particle showers can *punch-through* the calorimeters and reach the muon chambers and be misreconstructed as muons. Electrons can be misidentified in the ECAL from neutral pions or unidentified photon conversions. Real nonprompt leptons originate from the semi-leptonic decays of heavy-flavour (c, b) quarks, as well as hadron decays, such as light-meson decays in flight. Despite the different possible sources and varying origins, they are estimated simultaneously; therefore they are interchangeably referred to as nonprompt (or *fake*) leptons, which refer to both cases simultaneously.

Such backgrounds are also considered as *reducible*, as they can be suppressed using improved or more stringent lepton quality criteria, discussed in Section 4. Due to the jet fragmentation being extremely difficult to model (discussed in Section 4.3), the probability of a jet being misidentified as a lepton is not modelled well in MC. Therefore, in regions where there is a significant nonprompt contribution, precise modelling is vital.

The production of $Z + jets$ with Z-decays to neutrinos $Z(\rightarrow \nu\nu) + jets$ and the QCD multi-jet backgrounds, with no prompt leptons by definition, are the dominant non-prompt backgrounds. There exists the possibility that a real lepton (mainly from $W + jets$ or $t\bar{t}$) is lost while a misidentified one is found, which augments this background type. In SRs with suppressed prompt background contributions, fake backgrounds have a higher relative contribution and require precise modelling. The most significant SRs for nonprompt leptons are in the tails of the m_T spectrum ($SRc : m_T > 95 \text{ GeV}$) beyond the kinematic edge of the W mass (Figure 6.9), as well as very low lepton p_T regions (Figures 6.7).

The nonprompt backgrounds have been estimated using several different data-driven background modelling methods, using simulation together with normalisation to data in dedicated control regions. In the first iteration of the Run 2 analysis [6], Z_{inv} and QCD multi-jet were estimated separately. The contribution of $Z + jets$ was estimated by emulation of decays to neutrinos ($Z \rightarrow \nu\nu$) using events from leptonic decays of the Z-boson $Z \rightarrow \ell\ell$. The contribution from QCD multi-jet was estimated using an *ABCD*-method, scaling from QCD-enriched orthogonal sideband regions. In the subsequent full 2016 analysis [5], the nonprompt contribution from all different sources was simultaneously estimated using the *tight-to-loose* method.

7.2.1 Emulation of Z-Boson Decays to Neutrinos

The production of $Z + jets$, where the Z-boson decays to neutrinos ($Z \rightarrow \nu\nu$), henceforth referred to as Z_{inv} , constitutes a non-negligible background of the analysis. The neutrinos lead to moderate p_T^{miss} , while a nonprompt lepton can be misreconstructed in association with one of the jets, leading to a signature similar to the signal.

Due to the similarity in the topology and kinematics, the invisible decay of the Z-boson can be emulated using events from its leptonic decay channel $Z \rightarrow \ell\ell$. The procedure of emulation of Z_{inv} from $Z \rightarrow \ell\ell$ is performed by removing the pair of charged leptons from the event and recalculating the \vec{p}_T^{miss} vector by adding the lepton \vec{p}_T s to it. The magnitude of the modified \vec{p}_T^{miss} is referred to as the emulated $p_{T_{em.}}^{miss}$.

A data-driven estimate of the background is obtained by the determination of $^{data/MC}$ correction factors from the $Z \rightarrow \ell\ell$ control sample, which is then applied to the simulation of Z_{inv} in the SR. The number of Z_{inv} events with an extra nonprompt lepton entering the SR is estimated as:

$$N_{\nu\nu\ell} = N_{\nu\nu\ell}^{MC} \cdot R_{\ell\ell}^{data/MC} \equiv N_{\nu\nu\ell}^{MC} \cdot R_{\ell\ell}^{data/MC} \cdot R_{(\ell\ell)\ell/\ell\ell}^{data/MC} \cdot SF_{\nu\nu/\ell\ell}^{data/MC} \quad (7.2)$$

The subscripts indicate the Z decay mode, including the possible requirement of an additional soft lepton. $N_{\nu\nu\ell}^{MC}$ is the Z_{inv} simulation yield in the SR, and $R_{\ell\ell}^{data/MC}$ is the

inclusive correction factor. The correction factor is broken down into the following sub-factors, which are determined separately:

- $R_{\ell\ell}^{data/MC}$: inclusive data/simulation ratio of cross sections for Z production, including efficiency \times acceptance ($\mathcal{A} \times \epsilon$)
- $R_{\ell\ell/\ell\ell}^{data/MC}$: data/simulation ratio of the probability to observe an additional third soft lepton in an event with a $Z \rightarrow \ell\ell$ decay, emulating a SR-like lepton selection
- $SF_{\nu\nu/\ell\ell}^{data/MC}$: double ratio of $\mathcal{A} \times \epsilon$ for the Z boson decays to neutrinos and leptons in data and simulation

Data from early *Run 2* is used, corresponding to an integrated luminosity of 12.9 fb^{-1} . In order to maximise statistics, both the $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ channels are considered. The $Z \rightarrow \mu\mu$ events are extracted from the *SingleMuon* dataset and are required to have passed the [HLT_\(Iso\)Mu24](#) trigger, while the $Z \rightarrow ee$ events are taken from the *SingleElectron* dataset and are required to have passed the [HLT_Ele27_WPTight_Gsf](#) trigger. The MC samples comprise primarily simulated Drell-Yan (Z/γ^*) events, as well as minor contributions from leptonic $t\bar{t}$ and di-boson (VV) processes.

The following selection on the leptonic pair is applied:

- Oppositely charged leptons (e^+e^- or $\mu^+\mu^-$)
- Lepton p_T cuts defined by the trigger plateaus
 - p_T of the leading lepton has to be higher than 26 GeV and 29 GeV for the $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ channels, respectively
 - p_T of the sub-leading lepton has to be higher than 20 GeV
- Baseline selection corresponding to the analysis' muon and electron object definitions
 - Tighter relative isolation $I_{\text{rel}} = I_{\text{abs}}/p_T < 0.12$

- Mass of the dilepton system has to be within 15 GeV of the Z mass: $|m_{\ell\ell} - m_Z| < 15 \text{ GeV}$

A high purity of the sample in $Z \rightarrow \ell\ell$ Drell-Yan events is ensured by restricting the mass of the dilepton system to be within a small Z-mass window and using tighter lepton identification criteria.

Furthermore, the following preselection requirements are applied in order to enter a phase-space closer to that of the analysis.

- $p_T^{\text{ISR}} > 100 \text{ GeV}$
- Veto on 3rd jets with $p_T > 60 \text{ GeV}$
- Veto on b-tagged jets
- Veto on taus with $p_T > 20 \text{ GeV}$

Additional selections are applied on the emulated $p_{T_{em.}}^{\text{miss}}$ to further converge on the parameter space of the SRs. A lower limit on $p_{T_{em.}}^{\text{miss}} > 75 \text{ GeV}$ and correspondingly $C_{T_{em.}} = \min(p_{T_{em.}}^{\text{miss}}, H_T - 100 \text{ GeV}) > 75 \text{ GeV}$ is applied. The final selection choices are motivated by a reasonable balance between the amount of statistics and the targeted phase-space.

The invariant mass $m_{\ell\ell}$ of the dilepton system is shown in Figure 7.3, displaying the Z-mass peak at $\approx 91 \text{ GeV}$. The transverse momentum $p_T(\ell\ell)$ of the dilepton system on the Z-mass peak is shown in Figure 7.4. The original p_T^{miss} distribution and its emulated counterpart is shown in Figures 7.5.

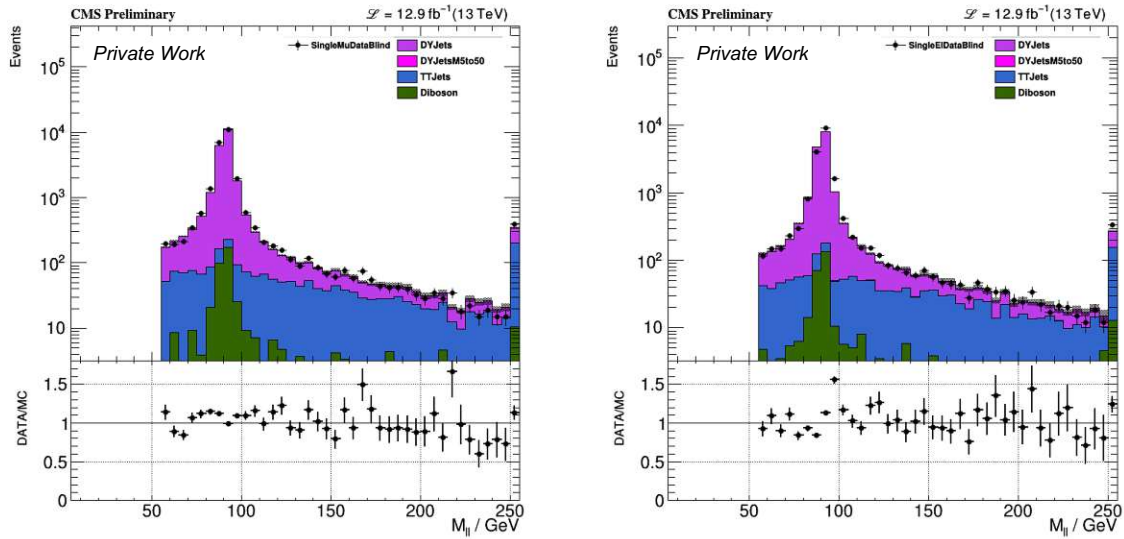


Figure 7.3: Distributions of dimuon (left) and dielectron (right) system's mass $m_{\mu\mu}$ with the requirement $C_{\text{TI},em.} > 75 \text{ GeV}$.

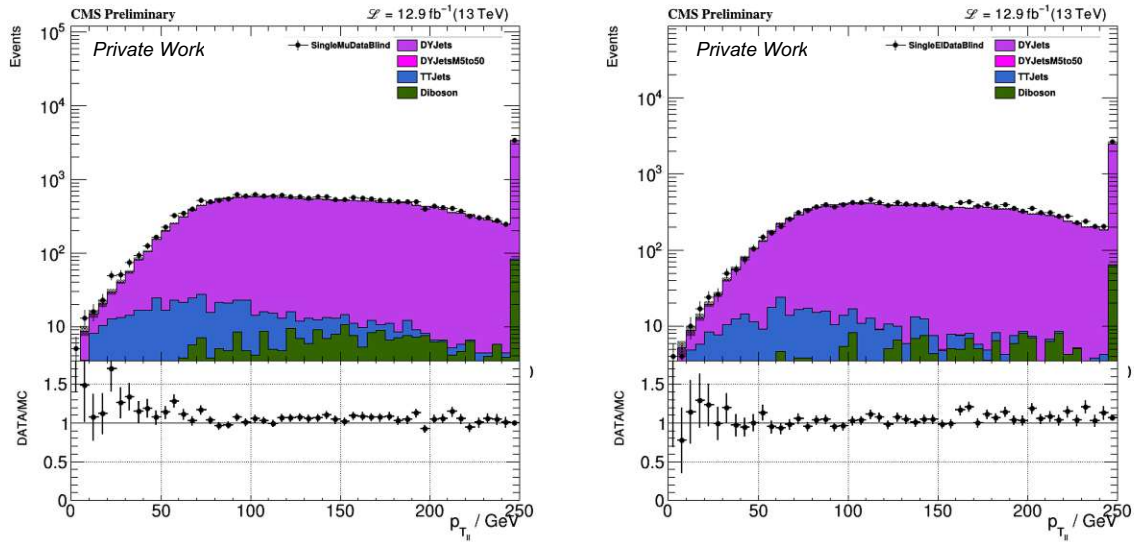


Figure 7.4: Distributions of (left) dimuon and (right) dielectron system's transverse momentum $p_T(\mu\mu)$, with the requirement $C_{\text{TI},em.} > 75 \text{ GeV}$ and $|m_{\ell\ell} - m_Z| < 15 \text{ GeV}$.

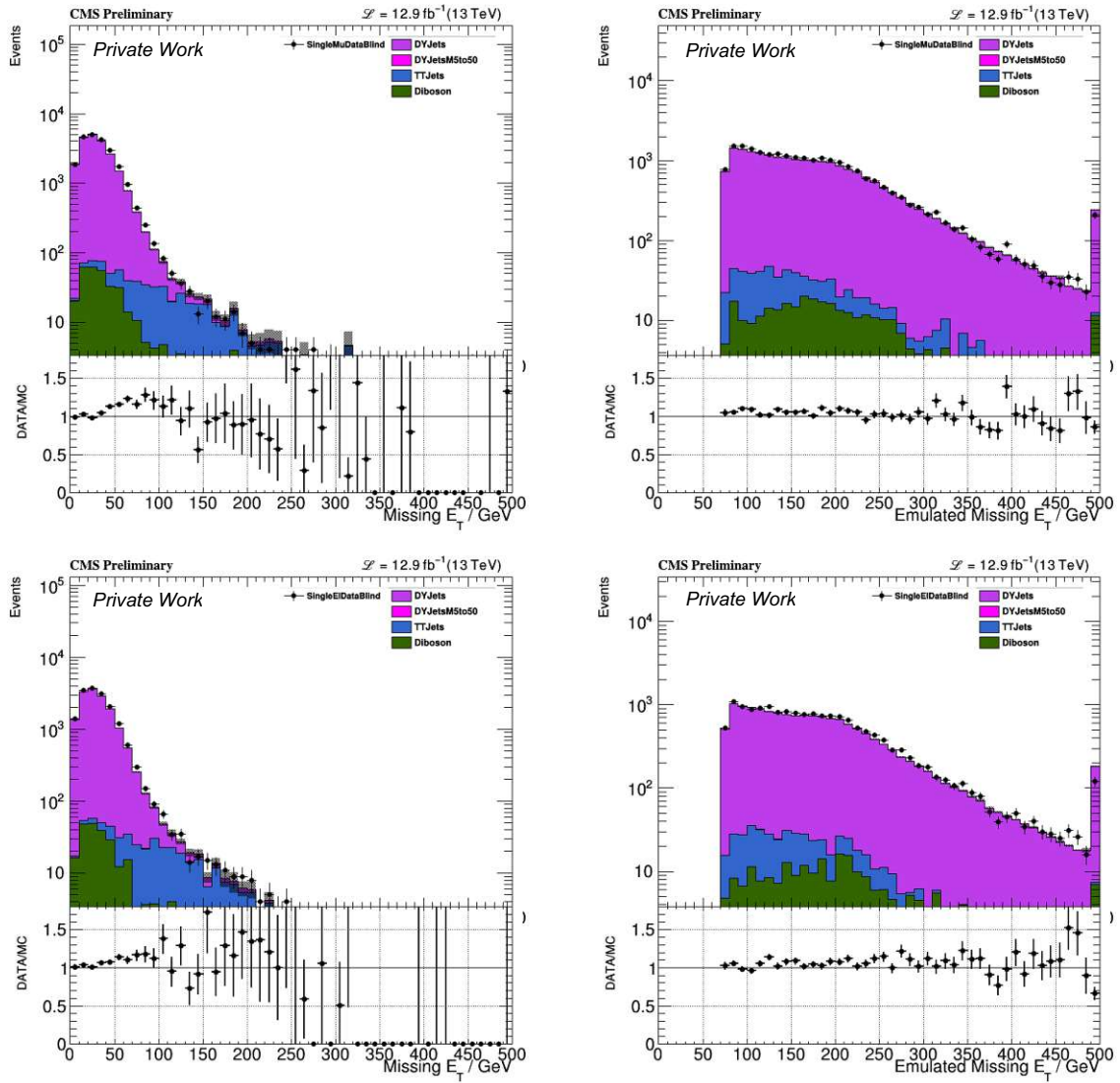


Figure 7.5: Distribution of p_T^{miss} before (left) and after emulation (right), for muons (top) and electrons (bottom) with the requirement of $C_{Tl,em.} > 75$ GeV and $|m_{\ell\ell} - m_Z| < 15$ GeV. Emulation refers to adding the dimuon system's \vec{p}_T to the p_T^{miss} .

The cross section ratio $R_{\ell\ell}^{\text{data}/\text{MC}}$ is determined on the Z-mass peak. For a selection equivalent to SRI, with $C_{\text{Tl},em.} > 300$ GeV, $R_{\ell\ell}^{\text{data}/\text{MC}}$ is measured as 0.99 ± 0.03 (stat.).

The probability of observing a third soft lepton is determined by applying an equivalent selection on the additional leptons as in the SRs, including vetos on extra hard leptons. Due to limited statistics when imposing the additional selection on a third lepton, the $|\eta|$ acceptance on the leptons is relaxed from $|\eta| < 1.5$ to 2.4 for muons and 2.5 for electrons, respectively.

The quantity $R_{\ell\ell/\ell\ell}^{\text{data}/\text{MC}}$ is a double ratio and therefore is expected to be stable for small to moderate variations of the selection.

Finally, the double ratio $SF_{\nu\nu/\mu\mu}$ is assumed to be one, with an uncertainty that is negligible compared to the other sources of systematics.

The total correction factor is determined by combining the results from the $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ channels. The correction factors $R_{\ell\ell\ell} \equiv R_{\ell\ell} * R_{\ell\ell/\ell\ell}$ for both the electron and muon channels are shown in Figure 7.6 as a function of $C_{\text{Tl},em.}$.

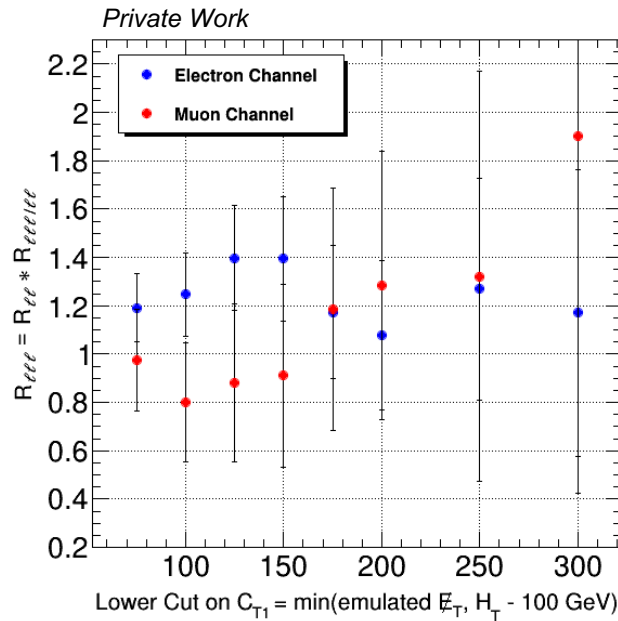


Figure 7.6: Correction factors $R_{\ell\ell\ell} \equiv R_{\ell\ell} * R_{\ell\ell/\ell\ell}$ for different minimum requirements on $C_{\text{Tl},em.}$ for both the muon and electron channels. The quoted uncertainties are statistical.

Given that there is no significant trend in terms of tightening the $C_{Tl,em.}$ requirement, the correction factor determined with the lowest statistical uncertainty is used; This corresponds to the SRl inclusive ratio determined with the selection $C_{Tl} > 75$ GeV. A conservative flat systematic uncertainty of 50 % is applied on the final correction factor to account for the extrapolation in $C_{Tl,em.}$. To verify that the simulation models the $p_T(\ell)$ distribution of this background sufficiently well, the estimation process is repeated upon splitting into three $p_T(\ell)$ bins: SRL1, SRH1 and SRV1, corresponding to 5-12 GeV, 12-20 GeV and 20-30 GeV, respectively, as shown in Table 7.1. Within uncertainties, the correction factor is stable across the bins.

SR	$R_{\ell\ell\mu} \equiv R_{ll} * R_{\ell\ell\mu/\ell\ell}$	$R_{\ell\ell e} \equiv R_{\ell\ell} * R_{\ell\ell e/\ell\ell}$
$SRL1$	0.81 ± 0.19	1.24 ± 0.17
$SRH1$	1.63 ± 1.05	1.01 ± 0.30
$SRV1$	2.96 ± 2.00	1.24 ± 0.52

Table 7.1: Correction factors $R_{ll} * R_{ll/\ell\ell}$ for $N_{\nu\nu}^{MC}$ for $C_{Tl} > 75$ GeV, split into p_T bins.

The correction factors used in the analysis are 0.97 ± 0.21 (stat.) ± 0.49 (sys.) for the muon channel and 1.19 ± 0.14 (stat.) ± 0.60 (sys.) for the electron channel.

7.2.2 QCD Estimation with ABCD Method

Due to the nature of hadron colliders, strong QCD processes dominate and manifest themselves as *jets* from the hadronization process of quarks and gluons. The QCD multi-jet background is characterised by multiple jets and hadronic activity. The background can be significant in leptonic searches when a jet is misidentified as a lepton, due to the high QCD cross sections, despite relatively low misidentification rates.

Considering the complexity of modelling the QCD background processes (discussed in Section 4.3), along with the combinatorics of misidentification of other physics objects, such as leptons, it is vital for the background to be accurately estimated. In the context of this analysis, the requirement of moderate p_T^{miss} together with a specific anti-QCD selection (discussed in Section 6.4.3.1) suppresses this background to the level that almost none of the simulated QCD events survive the signal region (SR) selection. In order to verify that the simulation is able to correctly characterise data around the SR, a dedicated data-driven background estimation method is implemented.

The estimation proceeds by scaling simulation to data from a QCD-enriched orthogonal sideband region via an *ABCD* method. An *ABCD* method involves the inversion of selections on SR observables in order to construct a control region (CR) enriched in the background that is being estimated. The estimated yield in the SR is then calculated via the product of transfer factors in the sideband regions of the inverted variables. The signal contribution in the sideband regions is assumed negligible, and the variables are assumed to be independent and largely uncorrelated.

Inverting the anti-QCD selection around the SR, by construction, creates a QCD-enriched control region. The main anti-QCD observables are hybrid isolation ($\text{HI} = I_{\text{rel}} \cdot \min(p_T(\ell), 25 \text{ GeV})$) and the angle between the leading two jets $\Delta\phi(j_1, j_2)$. High hadronic activity translates to large lepton isolation, while the QCD background typically consists of dijet events, with the jets oriented in a back-to-back topology.

The baseline regions relevant to the ABCD method are the following:

- A: SR with applied isolation ($HI < 5 \text{ GeV}$) and applied $\Delta\phi_{j_1,j_2} < 2.5 \text{ rad}$
- B: inverted $\Delta\phi_{j_1,j_2}$ and applied isolation
- C: inverted isolation and applied $\Delta\phi_{j_1,j_2}$
- D: inverted $\Delta\phi_{j_1,j_2}$ and inverted isolation

Figure 7.7 demonstrates the regions in the isolation- $\Delta\phi_{j_1,j_2}$ parameter space:

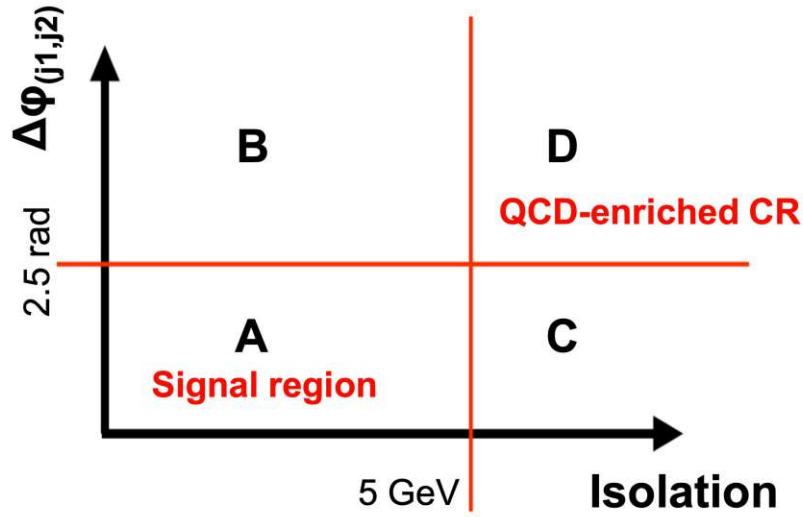


Figure 7.7: ABCD method sideband regions with inverted anti-QCD cuts on $\Delta\phi_{j_1,j_2}$ and lepton isolation.

Under the assumption that isolation, $\sigma_{I\eta I\eta}$ and $\Delta\phi_{j_1,j_2}$ shapes are fairly uncorrelated, a quantitative estimate of the QCD background in the SR can be expressed as:

$$N_{QCD} \simeq SF_{\not{I}\not{A}}^{data/MC} \cdot \frac{N_{\not{I}A}^{MC} \cdot N_{IA}^{MC}}{N_{\not{I}\not{A}}^{MC}} \quad (7.3)$$

where indices I and A correspond to isolation and $\Delta\phi_{j_1,j_2}$, respectively and the slash symbolises an inverted cut. $SF_{\not{I}\not{A}}^{data/MC}$ is the ratio between data, with all non-QCD (electroweak) MC contributions subtracted, and QCD simulation yield in the fully inverted region.

The same baseline method is used for both the electron and muon channels, with a selection corresponding to the SR, apart from the cuts which are inverted or relaxed. The method was optimised in terms of increasing the statistics and purity of the QCD sample and sideband regions. This is achieved by taking advantage of the relevant observables, such as the lepton impact parameters and the ECAL cluster shape covariance $\sigma_{I\eta I\eta}$, which forms part of the standard electron identification criterion (see Section 4.1).

- p_T^{miss} is relaxed to 200 GeV in region B for both channels to enhance statistics in the sideband regions
- In the muon channel, the impact parameter d_{xy} is relaxed from 0.02 cm to 0.05 cm in region C
- In the electron channel, $\sigma_{I\eta I\eta}$ is additionally inverted in region C

All these modifications are also applied to the fully-inverted region D for a consistent extrapolation with the transfer factors.

Data from early Run 2 is used, corresponding to an integrated luminosity of $L = 12.9 \text{ fb}^{-1}$, with the set of simulated background samples being the same as in the standard analysis. Figure 7.8 demonstrates that inverting the hybrid isolation cut at 5 GeV forms a QCD-enriched region. The fully inverted region D is formed by additionally inverting the cut on $\sigma_{I\eta I\eta}$ or loosening the cut d_{xy} , for the electron and muon channels, respectively. Applying the isolation cut forms the sideband region B.

The distribution of $\Delta\phi_{j_1, j_2}$ in Fig. 7.9 indicates that one moves to a QCD-enriched region by inverting the cut at 2.5 rad . Applying the $\Delta\phi_{j_1, j_2}$ cut and additionally inverting the cut on $\sigma_{I\eta I\eta}$ or relaxing the cut d_{xy} , for the electron and muon channels, respectively, forms the sideband region C.

The standard electron identification criterion requires $\sigma_{I\eta I\eta} < 0.0114$ in the ECAL Barrel (EB: $|\eta| \leq 1.479$) and $\sigma_{I\eta I\eta} < 0.0352$ in the ECAL Endcaps (EE: $1.479 < |\eta| <$

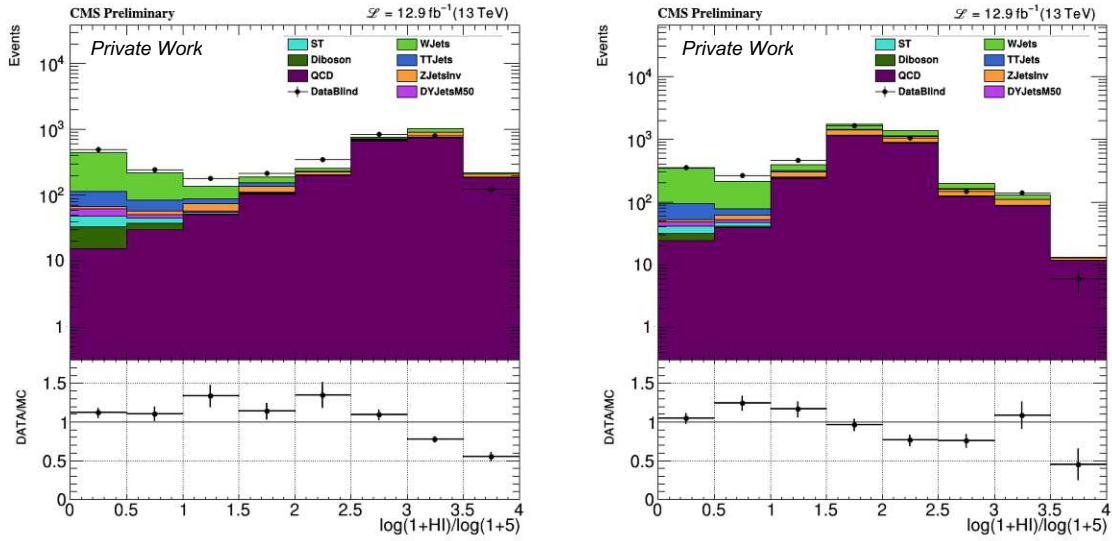


Figure 7.8: Hybrid isolation (HI) distribution with a baseline selection corresponding to SR1 applied, with the cut on $\Delta\phi_{j_1, j_2}$ inverted and p_T^{miss} relaxed to 200 GeV for the muon channel (left) and the electron channel (right). For (left) the cut on $\sigma_{I\eta I\eta}$ is not applied and for (right) the cut on d_{xy} is not applied. In the x-axis, 1 corresponds to the division between the SR: $\text{HI} < 5$ GeV.

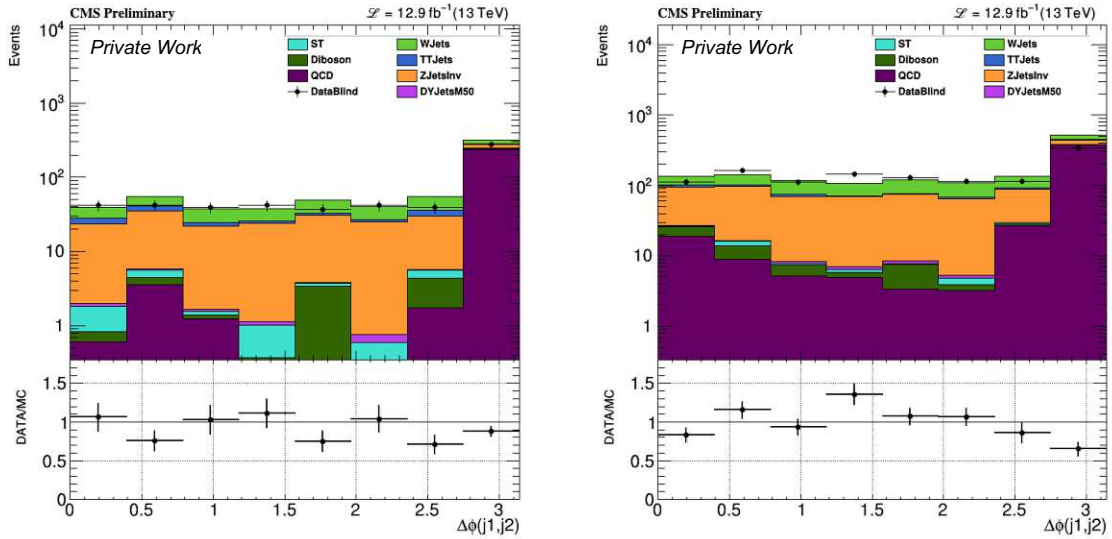


Figure 7.9: $\Delta\phi_{j_1, j_2}$ distribution with a baseline selection corresponding to SR1 applied and with the cut on hybrid isolation (HI) inverted for the muon channel (left) and the electron channel (right). For (left) the cut on $\sigma_{I\eta I\eta}$ is not applied and for (right) the cut on d_{xy} is not applied.

2.5). Figure 7.10 shows the distribution of the $\sigma_{I\eta I\eta}$ variable, indicating that inverting the cut leads to a further enrichment in QCD events.

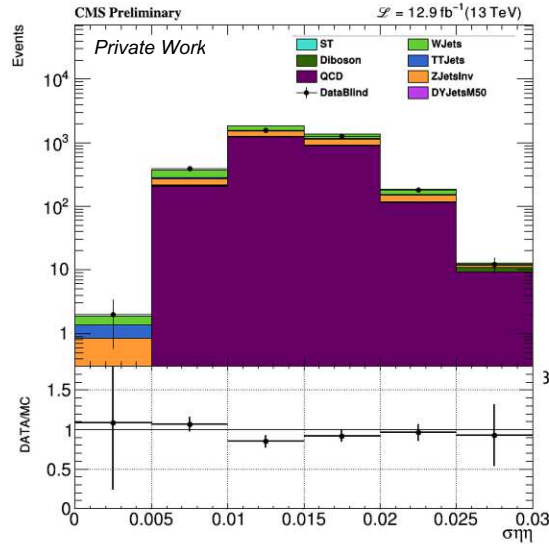


Figure 7.10: $\sigma_{I\eta I\eta}$ distribution with a baseline selection corresponding to SRI applied and cuts on $\Delta\phi_{j_1, j_2}$ and hybrid isolation (HI) inverted for the electron channel. p_T^{miss} is relaxed to 200 GeV.

In order to verify the estimation method, closure tests comparing the estimate to the simulated yield were performed in lepton p_T - and m_T -inclusive SRI:

- Muon channel:
 - the data/MC scale factor as $SF_{\cancel{A}}^{\text{data}/\text{MC}} = 0.94 \pm 0.04$
 - estimate $N_{QCD} \simeq 0.32 \pm 0.12$ (stat.) ± 0.16 (sys.)
 - QCD MC yield is 4.31 ± 3.35 (stat.)
- Electron channel
 - the data/MC scale factor as $SF_{\cancel{A}}^{\text{data}/\text{MC}} = 0.79 \pm 0.07$
 - estimate $N_{QCD} \simeq 0.74 \pm 0.22$ (stat.) ± 0.39 (sys.)
 - QCD MC yield is 2.57 ± 2.43 (stat.)

The estimates agree with the simulated yields within the limits of uncertainty. For SR2, which requires a soft b-tagged jet, the contribution of the QCD background is expected to be even smaller. This is supported by the fact that the corresponding data and MC yield in the QCD-enhanced control region $N_{\cancel{e}eA}^{\text{data}}$ is only of the order of a few events.

Sources of systematic uncertainties related to the estimation are primarily covered with a conservative 50 % systematic uncertainty on the final estimate, to account for factorisation of the $SF^{\text{data}/\text{MC}}$ from simulation across the regions and any correlations between the variables. Additionally, a 30 % systematic uncertainty is applied to the electroweak subtraction due to the impurity of the QCD-enriched region.

7.2.3 Tight-to-Loose Method

The different sources of nonprompt backgrounds can be simultaneously estimated using a data-driven *tight-to-loose* method [267]. The advantage of this method over the separate methods described in Section 7.2.1 and 7.2.2 is that the measurement of the transfer factors is performed directly in data, and the method need not distinguish between different processes. Therefore, it is less sensitive to potential MC mismodelling of nonprompt-enriched regions. This method was used in the *soft single-lepton analysis* with the full 2016 dataset [3, 4, 5].

Similarly to the *ABCD* method, one defines a set of *loose* lepton identification and isolation criteria to select nonprompt candidates. The *tight* criteria correspond to the final leptonic selection of the analysis, presented in Table 4.2. Events where at least one lepton fails the tight criteria but passes the loose selection are reweighted by a transfer factor based on the probability that the lepton passes both criteria. This probability, the tight-to-loose ratio ϵ_{TL} , is determined in a control region (CR) that is constructed to be enriched in nonprompt leptons, known as the *measurement region* (MR). A loose-not-tight (L!T) sideband control region (also known as the *application region*; AR), orthogonal to the SR, is formed by applying the same kinematic cuts as the SR, while requiring the lepton to fail the tight and pass the loose selection.

The predicted number of fake leptons entering the SR is estimated via:

$$N_{\text{fake}}^{SR} = \frac{\epsilon_{TL}}{1 - \epsilon_{TL}} \cdot (N_{\text{data}}^{L!T} - N_{\text{prompt}}^{L!T}) \quad (7.4)$$

where $N_{\text{data}}^{L!T}$ is the data yield in the L!T CR and $N_{\text{prompt}}^{L!T}$ is the prompt simulated yield in the L!T CR.

The *loose* definition is defined by relaxing the isolation and impact parameter selections, which are the main leptonic observables that counter misidentification against jets. The hybrid isolation threshold is increased by a factor of four to $HI < 20$ GeV, which translates to $I_{\text{abs}} < 20$ GeV for $p_T(\ell) < 25$ GeV and $I_{\text{rel}} < 0.8$ for $p_T(\ell) > 25$ GeV, while the impact parameter requirements are relaxed to $|d_{xy}| < 0.1$ and $|d_z| < 0.5$. A graphical representation of both the tight and loose regions in the parameter space of IP and hybrid isolation variables is shown in Figure 7.11.

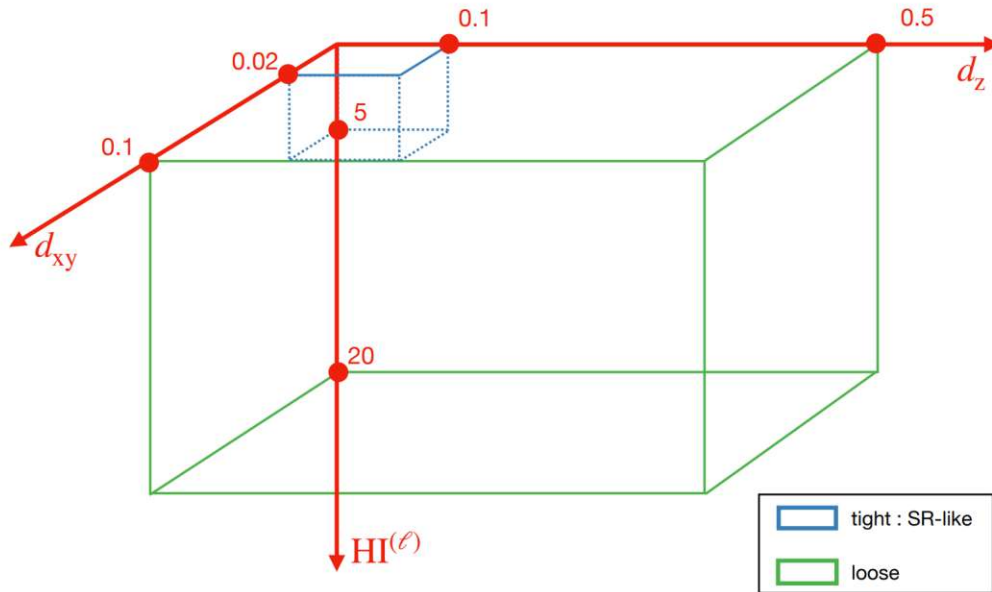


Figure 7.11: A graphical representation of the loose region (green) and the SR-like tight region (blue) [268].

The estimate is data-driven, where the simulated events with prompt leptons are subtracted from data in the AR and MR. This requires the separation of events with leptons into prompt and nonprompt categories, which is performed using MC generator-

matching of generated and reconstructed leptons, with tight requirements on ΔR and the p_T ratio.

Detailed studies have shown that within the measurement and application regions, there is a non-negligible contribution of nonprompt leptons originating from the misidentification of hadronically-decaying taus. More specifically, they are associated with *punch-through* hadrons traversing past the calorimeters or neutral pions misidentified in the ECAL. Investigations concluded that they are not modelled well with this approach due to their differing isolation shapes that the method extrapolates in. Therefore, these events have been tagged using ΔR matching between the reconstructed nonprompt leptons and generated taus coming from W s, Z/γ s and excluded from the estimation procedure. As such, they are included as part of the main background estimation method for prompt $W + jets$ and $t\bar{t}$. As opposed to fake leptons, their fraction increases with $p_T(\ell)$ and therefore is more relevant in the control regions, whereas in the SRs, their contribution relative to dominant prompt backgrounds is maximum of the order of a few percent. Also, in high p_T bins of SRs a and b with $m_T < 95$ GeV, dominated by prompt backgrounds, their size is comparable to other nonprompt contributions.

The MR is selected as an orthogonal dataset selected by an H_T -based trigger, which is by definition dominated by QCD multi-jet events and thus enriched in nonprompt leptons. The events are extracted from the *JetHT* dataset and are required to have passed the lowest unrescaled `HLT_PFHT800` trigger¹. Apart from the trigger plateau selection requiring very high $H_T > 900$ GeV, additional cuts on $p_T^{\text{miss}} < 40$ GeV and $m_T < 30$ GeV are applied to reduce the prompt contamination and increase the purity of the control sample. The lepton with the leading p_T is considered for the measurement, and the tight-to-loose ratio ϵ_{TL} is measured as a function of lepton p_T and η .

¹Additionally, jet triggers `HLT_PFJet450` and `HLT_AK8PFJet450` are added in order to recover an inefficiency of the H_T -based triggers in part of the 2016 dataset (Run H).

In the high p_T region in the muon channel, there are relatively large $^{\text{data/MC}}$ discrepancies, and the prompt contamination is relatively high. In this very high H_T region, the description of the prompt background by the simulation is unreliable, despite the ISR reweighting that is applied (as discussed in Section 6.3.2). Therefore, for muons with $p_T > 50$ GeV, the measurement of ϵ_{TL} is performed in an alternative low- H_T MR. Events in this region are extracted from the *SingleMu* dataset and are required to have passed the `HLT_Mu50` muon trigger (without an isolation requirement), and the measurement is performed with the triggering muon. The same purity requirements on p_T^{miss} and m_T are applied. Furthermore, a requirement on $p_T^{\text{ISR}} > 100$ GeV is applied to enter a phase-space closer to that of the analysis.

The lepton p_T distributions for both tight and loose leptons in the baseline and alternative low- H_T MR are presented in Figure 7.12, indicating an improved $^{\text{data/MC}}$ agreement. The binning for $p_T < 30$ GeV corresponds to the p_T -binning of the SRs, whereas, for $p_T > 30$ GeV, the binning is optimised based on the dependence of ϵ_{TL} on the lepton p_T .

The tight-to-loose ratio ϵ_{TL} measurements in both data and simulation are shown in Figure 7.13 as a function of lepton p_T . The measurement is split into two $|\eta|$ bins at $|\eta| = 1.5$, corresponding to the SRI selection.

The final nonprompt contributions in the signal and control regions are then estimated with Equation 7.4 using the measured ϵ_{TL} ratios to reweight the AR, after subtracting the prompt contribution estimated from simulation. The same ϵ_{TL} measurements and systematics have been used in the nonprompt background estimation of the MVA-based analysis.

The systematic uncertainties related to the estimation method are covered in Section 8.1: for the higher lepton p_T region, the systematic uncertainties are very large. In this region, the simulation description is not accurate due to the limitations of ISR modelling discussed in Section 6.3.2, also indicated in Figure 7.12. Due to a high contamination of prompt leptons, the related error is propagated into the prompt subtraction systematics. Therefore, the method is limited in this region, as also indicated by

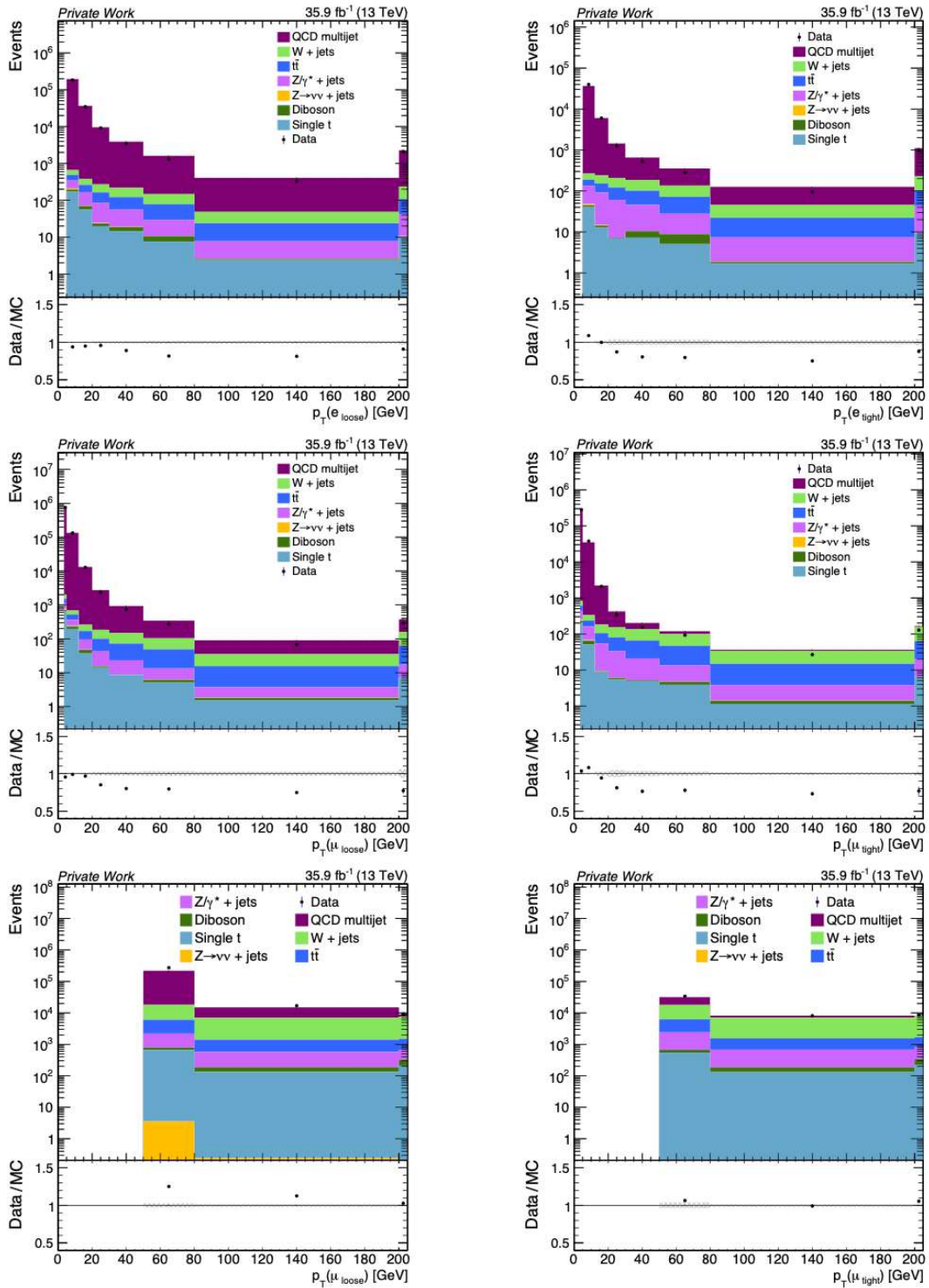


Figure 7.12: Lepton p_T distributions for loose (left) and tight (right) selection criteria for electrons (top) and muons (middle) in the baseline MR, as well as in the low- H_T MR for muons (bottom).

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar. The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

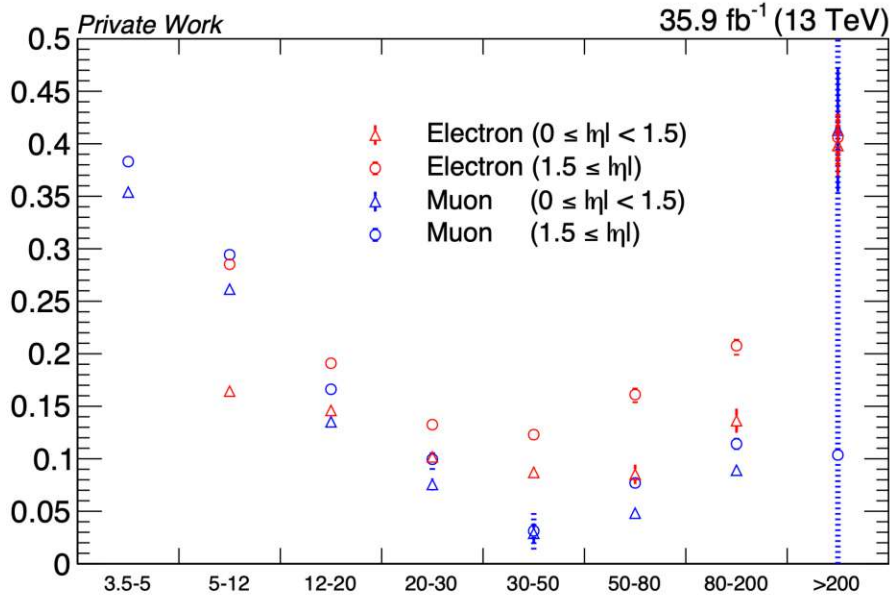


Figure 7.13: Tight-to-loose ratios ϵ_{TL} in data simulation for electrons and muons, as a function of lepton p_T .

the non-closure systematics. Nevertheless, this region corresponds to the main analysis control regions, which are dominated by prompt leptons and thus the nonprompt component has a negligible effect on the final estimates.

In ARs where there is a higher prompt contribution, it can happen that there is a statistical overlap of the prompt subtraction with zero, leading to negative nonprompt estimates. In these cases, the estimation is constrained using a Maximum Likelihood Fit (MLF) to ensure positive yields. The differences between the estimates with simple prompt subtraction and the MLF are shown in Figure 7.14.

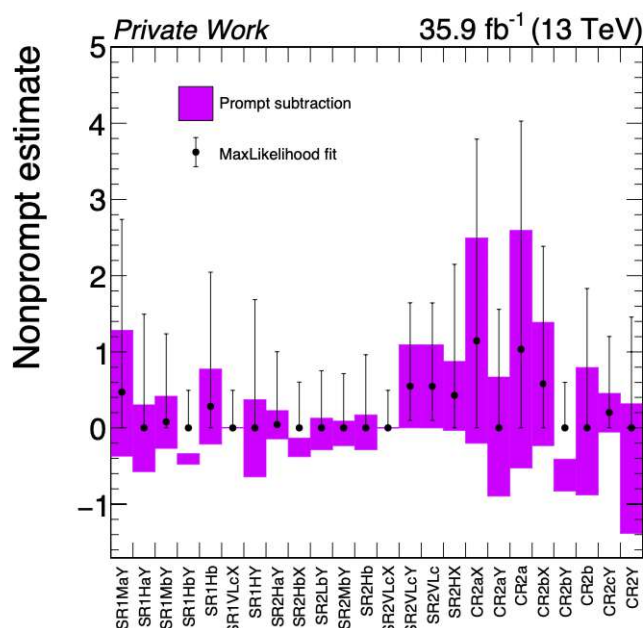


Figure 7.14: Comparison of nonprompt estimates based on a simple prompt subtraction and MLF in SR bins where there is a statistical overlap with zero [252].

Due to the size of the systematic uncertainties in some regions, albeit not affecting significantly the final results, shows that this method has its limitations, requiring further refinements. Potential improvements include a MR with a lower H_T requirement, where the MC description of the prompt contamination is more reliable, making the estimation procedure less sensitive to differences between data and simulation. A consideration is to use events triggered by single-lepton triggers and measure sub-leading leptons (à la *tag-&-probe*), which would also be a phase-space closer to that of the final analysis selection. This approach has been adopted in the context of the full Run 2 analysis, performing the ϵ_{TL} measurement in a region selected by prescaled low- p_T single-jet triggers, resulting in a higher purity and lesser dependence on the ISR reweighting of contributions from prompt leptons [268].

Results

8.1 Systematics

Several sources of background and signal systematic uncertainties are taken into account, which have been discussed throughout include: background estimations, ISR modelling, pileup dependence, jet energy scale and resolution (JES/JER), object efficiencies or the luminosity uncertainty, as well as additional ones spanning from theoretical of experimental effects.

Initial-State Radiation

Since the analysis relies on significant ISR, one of the leading systematic effects is due to the simulation of the ISR p_T spectra. As discussed in Section 6.3.2, the mismodelling of ISR needs to be corrected with dedicated correction factors for the main background ($t\bar{t}$, $W + jets$) and signal samples. Systematic uncertainties are added related to the methods that are used to determine the factors and correct the samples, taking into account any possible differences in shapes. The uncertainty is determined by calculating its effect on the signal acceptance: in $t\bar{t}$ and signal samples, it is evaluated using half the correction size, whereas for $W + jets$ the full size of the correction is used. The relative size of the uncertainties with respect to the total background estimation is approximately between 4-10 % for the $W + jets$ sample and less than 1 % for the $t\bar{t}$

sample in the signal regions. Since the backgrounds are normalised from data, only the effect on the lepton p_T spectrum shape impacts the background estimate. For the signals, the uncertainty ranges between 5-7 %.

Lepton Scale Factors

As discussed in Section 6.3.2, lepton scale factors between data and simulation are calculated for the final object selection (Figure 6.4). On top of the statistical uncertainty of the scale factor measurement, a flat 1 % systematic uncertainty is added in the SR for the backgrounds, to take into account any residual induced shape uncertainty between the high- p_T CRs and the low- p_T SRs, as the yields are normalised in the CRs. Additional scale factors are also determined to account for the simplified FASTSIM detector simulation (Figure 6.5) for the signals. A 2 % uncertainty is applied on the $FULLSIM/FASTSIM$ scale factors, together with an additional flat 2 % uncertainty to account for any residual effects.

B-Jet Tagging Scale Factors

The analysis relies on the multiplicity of b-tagged jets, so variations in the efficiency and purity of the b-jet identification could move events between the SRs and change the fractions of $W + jets$ and $t\bar{t}$ events in each bin. The size of this effect is estimated by varying the b-tagging scale factors (Section 6.3.2) within the uncertainties. The variations are performed separately for light and heavy-flavoured jets. For the signal samples, corresponding $FULLSIM/FASTSIM$ scale factors are also varied. For backgrounds, the b-tagging uncertainties are $< 1\%$ in most SRs, while for signal, the uncertainty is between 1 and 3 %.

Jet Energy Scale and Resolution

Systematic uncertainties related to the miscalibrations of the jet energy scales (JES) are calculated by varying the corresponding jet energy corrections (JEC) within one standard deviation for jets with $p_T > 10$ GeV and by 10 % for jets with lower momenta. The uncertainties due to the differences in jet energy resolution (JER) between data and simulation are compensated by smearing the jet transverse momenta according to a Gaussian distribution with a width that is given as a function of the jet p_T and η . The effect of each JES and JER variation is propagated to all related hadronic (and p_T^{miss}) observables. The corresponding JES systematic uncertainties reach up to 2.1 % for backgrounds and 4 % for signal, whereas the JER uncertainties go up to ≈ 1 % for signal and backgrounds.

Renormalisation and Factorisation Scales

The effect of renormalisation and factorisation scales (as described in Section 2.1.2) in the calculation of the signal cross sections also needs to be taken into account. This is performed by modifying the renormalisation and factorisation scales by eight different combination of factors 0.5, 1 and 2 (avoiding unphysical combinations). The uncertainty is estimated from the envelope of the yields of the different combinations and is between 2-3 %.

Pileup

As discussed in Section 6.3.2, all simulated background samples were reweighted to match the pileup profile of the data. Residual uncertainty on the background estimate related to this procedure is estimated by varying the minimum bias cross section used in weights by 5 %.

Trigger Efficiency

Taking into account the trigger efficiency measurement in Figure 6.6 described in Section 6.4.2, the efficiency at the plateau is $\approx 100\%$. Therefore, only an uncertainty of 1% is applied to the simulated background and signal samples in the SRs to account for any biases in the measurement.

Luminosity and Cross Section

All processes that are not predicted from data are subject to an uncertainty of 2.5%, reflecting the relative precision of the integrated luminosity, which is applied to the signal yields in the SRs. An uncertainty of 50% on the cross section of all background processes whose yields are predicted directly from simulation, which should additionally cover for the luminosity uncertainty.

Prompt Estimation

The prompt background estimations (Section 7.1) treat both the $t\bar{t}$ and $W + jets$ simultaneously. Superimposed (and area-normalised) lepton p_T distributions for $W + jets$ and $t\bar{t}$ in the signal regions Figure 8.1, indicate that the p_T distributions in these processes are very similar and their separate estimates would be highly anti-correlated. However, residual differences between the shapes, which can be significant in high- m_T regions, need to be accounted for in the estimation. A conservative 20% limit for the variation in the cross section ratio between $W + jets$ and $t\bar{t}$ is assumed, which is then propagated to the transfer factors (TF^{CR}) used to normalise and estimate the prompt backgrounds. The systematic uncertainties related to the $W + jets$ and $t\bar{t}$ lepton p_T shape differences range between 0.1-2.2% in all SRs.

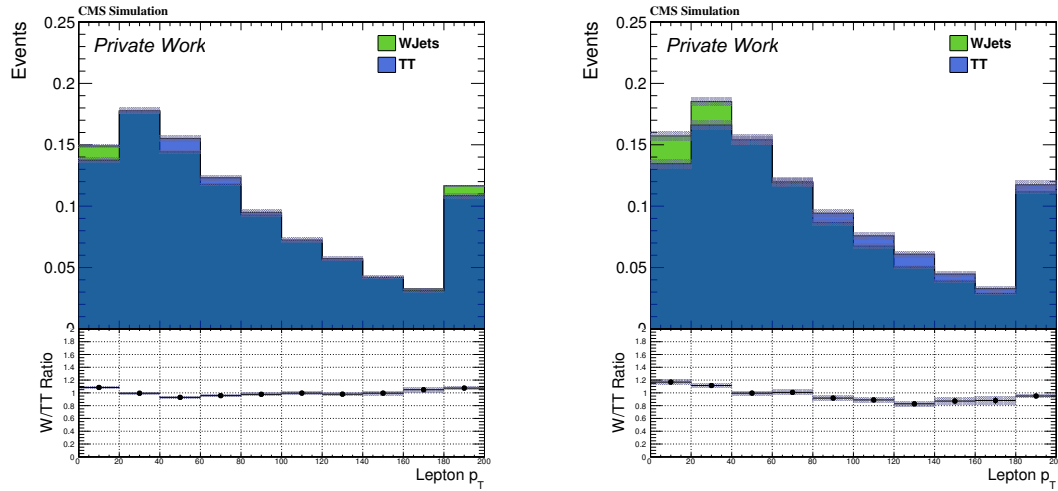


Figure 8.1: Superimposed lepton p_T distributions for $W + jets$ and $t\bar{t}$ for an inclusive selection corresponding to SR1 (left) and SR2 (right). The distributions are area-normalised.

Nonprompt Estimation

Given that the fake-rate method simultaneously estimates the different sources of non-prompt leptons, it is crucial to evaluate the underlying systematic uncertainties of the procedure described in Section 7.2.3. One of the main sources of the systematic uncertainty is the non-universality of the ϵ_{TL} due to the dependence on flavour and mother parton p_T of the jets originating the nonprompt leptons. In order to evaluate this systematic effect, the measurement was additionally performed when enriching or depleting the sample with b-tagged jets, by applying a b-veto or requiring at least one b-tagged jet. The systematic uncertainty on the ϵ_{TL} is determined based on these b-tag variations, with the values shown in Table 8.1.

p_T Bin	VL	L	M	H	CR
Sys.	20%	20%	30%	30%	50%

Table 8.1: Systematic uncertainties on the tight-to-loose ratio non-universality, based on measurements with b-tag variations. The first set of bins corresponds to the SR bins and the final bin corresponds to the CR with $p_T > 30$ GeV

Apart from the MC statistical uncertainty due to the prompt subtraction in both the

measurement and application regions, the systematic uncertainty related to the $W + jets$ and $t\bar{t}$ reweighting of the prompt samples is additionally propagated to the estimate. The latter systematic in the nonprompt-enriched measurement region is negligible relative to the non-universality systematic.

A Monte Carlo closure test of the procedure is performed in the signal regions, where the estimate using the ϵ_{TL} determined using simulation only is compared to the MC yield in the SR. For this test, the regions with low relative fake contribution (low- m_T bins a and b) are merged into one bin (ab), corresponding to $m_T < 95$ GeV, for simplicity. Moreover, the region with $p_T > 30$ GeV, corresponding to the high- p_T control region, is merged into one bin after estimation. The MC closure is shown in Figure 8.2. The uncertainties shown contain both the statistical and propagated systematic uncertainty related to the non-universality of the ϵ_{TL} .

Additional non-closure systematic uncertainty is assigned based on the level of the closure in individual bins and is shown in Tables 8.2 and 8.3.

Bin	sr1vlab	sr1lab	sr1mab	sr1hab	cr1lab	sr1vlc	sr1lc	sr1mc	sr1hc	cr1c
Sys.	20%	30%	80%	100%	200%	25%	–	–	–	100%

Table 8.2: Additional nonprompt estimation non-closure systematic uncertainties in SRI and CRI based on the combined MC closure, assigned per individual bin.

Bin	sr2vlab	sr2lab	sr2mab	sr2hab	cr2ab	sr2vlc	sr2lc	sr2mc	sr2hc	cr2c
Sys.	–	–	–	30%	150%	30%	30%	30%	30%	30%

Table 8.3: Additional nonprompt estimation non-closure systematic uncertainties in SR2 and CR2 based on the combined MC closure, assigned per individual bin.

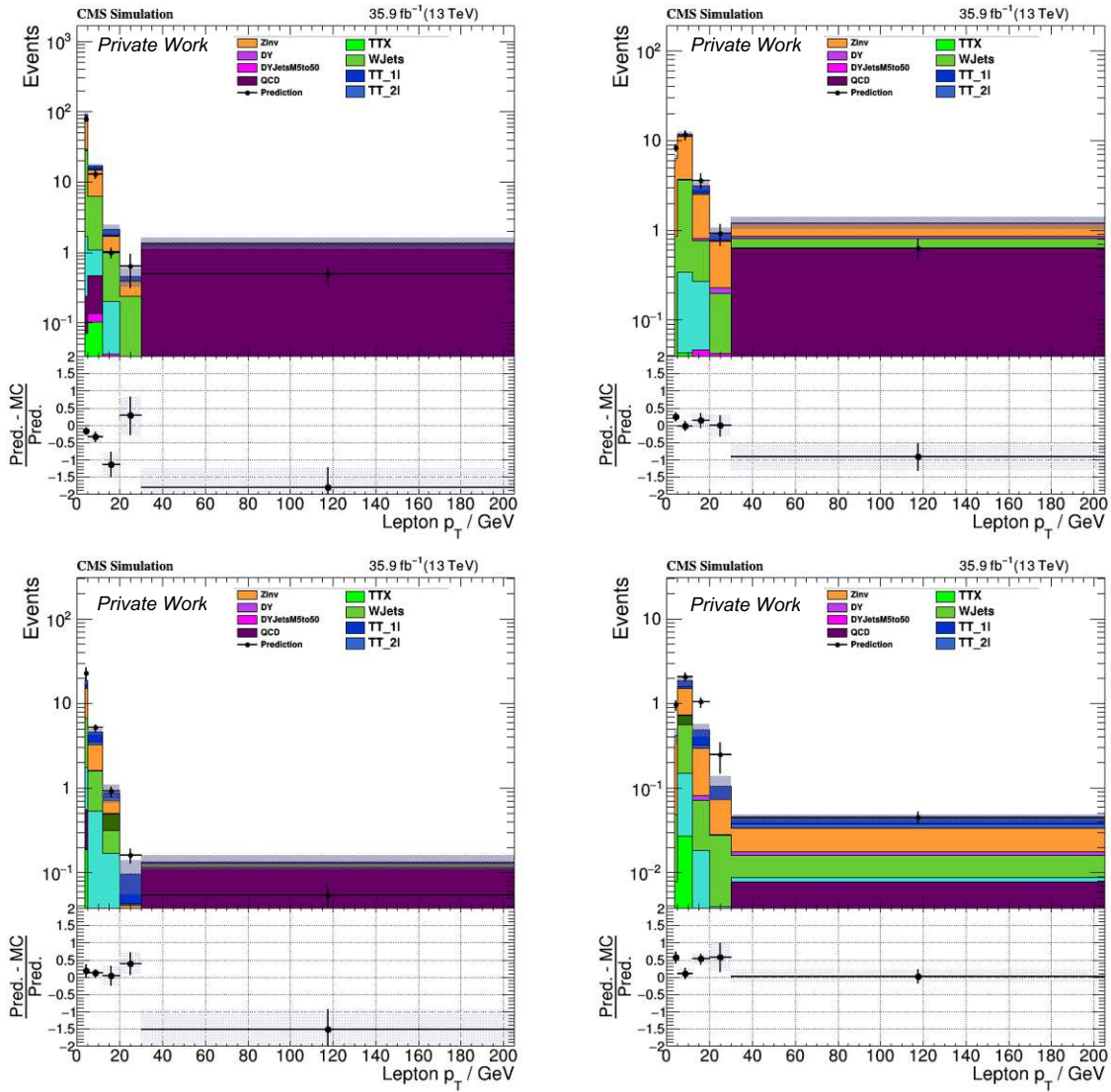


Figure 8.2: Nonprompt estimation MC closure in SRI (top) and SR2 (bottom) for m_T -regions ab : $m_T < 95$ GeV (left) and c : $m_T > 95$ GeV (right). Error bars represent both the statistical uncertainties and ϵ_{TL} non-universality systematic.

Summary

A summary of all the relevant systematic uncertainties in the search is presented in Table 8.4, relative to the total background and signal prediction in the main SRs. The largest systematics come from the ISR reweighting for both backgrounds and signal. The relative uncertainty related to the nonprompt background can be relatively significant for specific bins.

Systematic uncertainty	Background		Signal
	SR1	SR2	
ISR ($t\bar{t}$ and signal)	0.1 – 0.5	0.1 – 0.8	5 – 7
ISR (W+jets)	4.5 – 10.2	1.9 – 4.4	–
Lepton efficiency	1.0 – 1.8	1.0 – 1.5	3
Trigger efficiency	0 – 0.1	0 – 0.1	1
b-tagging	0.1	0.1 – 1.0	1 – 3
JES	1.2 – 2.1	0.1 – 1.4	3 – 4
JER	0.1 – 0.5	0.1 – 1.1	0 – 1
Pileup	0.1 – 1.8	0.1 – 2.0	1
Prompt estimation	0.1 – 1.6	0.1 – 2.2	–
Nonprompt estimation	1.0 – 4.6	1.0 – 9.5	–
Renormalisation and factorisation scales	–	–	2 – 3
p_T^{miss} FASTSIM modelling	–	–	2 – 3
Luminosity	–	–	2.5

Table 8.4: Typical ranges for relative systematic uncertainties (in %) on the total background or signal prediction in the main SRs. The dash indicates that the uncertainty source is not applicable.

8.2 Results

The total number of predicted background events in the SRs comprises the prompt and nonprompt background estimations together with the simulated rare processes. The predicted yields in the SRs are compared to the observed data yields in each of the 44 SRs, as shown in Figure 8.3. As indicated in Figure 8.3, the main sources of sensitivity of the search are the regions in the high m_T tail (Figure 6.9), as well as

the very low lepton p_T bin (Figure 6.7). Essentially, there exists a trade-off between sensitivity coming from low lepton p_T and high m_T . The results are consistent with the SM expectations, as no significant deviation is observed.

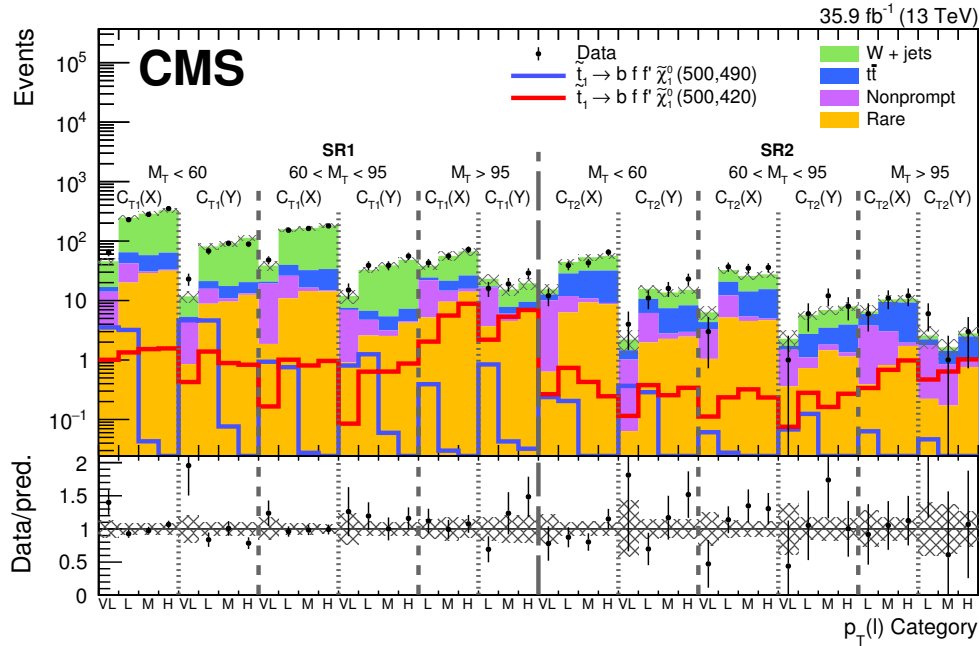


Figure 8.3: Summary of observed data and expected background yields in the SRs, with two T2tt benchmark signal points with extreme Δm [227].

For a BSM signal (s) and background (b) described by Poissonian distributions, the likelihood function, given an observed data yield (n), has the following form:

$$\mathcal{L}(\text{data}|r, \theta) = \prod_j^{\text{regions}} \text{Pois}(n_j | r \cdot s_j(\theta) + b_j(\theta)) \cdot \prod_i^{\text{nuisances}} p(\tilde{\theta}_i | \theta_i) \quad (8.1)$$

where r is a signal strength modifier, where $r = 0$ corresponds to a background-only hypothesis and $r > 0$ corresponds to a signal+background ($s + b$) hypothesis. The systematic uncertainties are included in the form of nuisance parameters θ_i , modelled by log-normal distributions with positive real values. The SR and CRs are treated simultaneously in the likelihood fit, accounting for potential small signal contamination in the CRs (which can be relevant for large Δm).

To distinguish between the two hypotheses, a test statistic is constructed in the form of

a profile likelihood ratio in the asymptotic limit [269], condensing the relevant information (expected/observed yields, systematics) of a BSM search into a single number:

$$\tilde{q}_r = -2 \ln \left(\frac{\mathcal{L}(\text{data}|r, \hat{\theta}_r)}{\mathcal{L}(\text{data}|\hat{r}, \hat{\theta})} \right) \quad \text{where } 0 \leq \hat{r} \leq r \quad (8.2)$$

where $\hat{\theta}_r$ is the conditional maximum likelihood estimator of nuisance parameter θ for a given signal strength modifier r , and \hat{r} and $\hat{\theta}$ denote the signal strength and nuisance that provide the global maximum likelihood. The constraint of \hat{r} is physical to ensure positive signal yields and avoid that potential upwards fluctuations are used as evidence against the signal. The comparison of the observed test statistic with that of the background-only and background+signal hypotheses can be then used to set limits on the signal strength modifier, or consequently the signal cross sections.

Given that no significant deviation from the SM is observed and the data is compatible with a background-only hypothesis, limits are set on the stop production cross section using a modified frequentist approach, known as the CL_s method [270, 271, 272]. The CL_s confidence level (CL) criterion is defined as a conditional probability of a BSM signal ($s+b$) observation for a test statistic $X (\equiv \tilde{q}_r)$, expressed in Equation 8.3. A common scenario in SUSY searches is that the expected signal and background yields are very small, especially in the case of depleted signal region bins where one has little sensitivity. Therefore, the criterion is designed to cover such cases in a more conservative way by normalising the probability or CL to that of the background-only hypothesis.

$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{P_{s+b}(X \leq X_{obs})}{1 - P_b(X \leq X_{obs})} < \alpha \quad (8.3)$$

The upper limits are defined to be greater than the true value of the test statistic with a probability at least equal to the CL, chosen to be at 95 % ($\alpha = 0.05$). Therefore, the upper limit is the largest value of the signal strength or cross section that is not excluded. As opposed to the observed limits, the expected limits are computed without data, as a pseudo-experiment using an Asimov dataset (background-only hypothesis),

free of fluctuations and gives an overall measure of the sensitivity of the analysis. Upwards deviations of the observed limit with respect to the expected limits would be indicative of a signal, under the assumption that the backgrounds are well-modelled. The results are interpreted in terms of the four-body decay and chargino-mediated scenarios, under the assumption of a 100 % branching ratio for the chosen decay channel where the decays are assumed to be prompt. Upper limits are set on the stop pair-production cross section at 95% CL in the plane of the stop masses versus Δm , as shown in Figure 8.4. The temperature plot corresponds to the upper limits σ_{obs}^{UL} , whereas the contour lines correspond to $\sigma_{pred}^{theo}/\sigma_{obs}^{UL} = 1$. The signal mass points to the left of the contour lines are excluded, for stop masses up to 560 GeV. The observed limit is weaker than the expected limit, especially at the intermediate to large mass splittings, due to a modest excess in SR1c (and SR2), which corresponds to slightly more than one standard deviation.

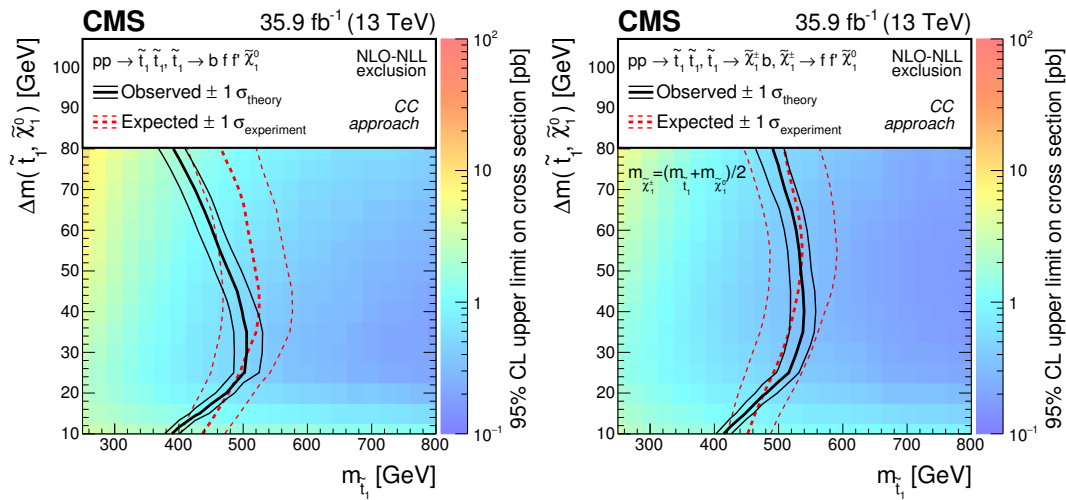


Figure 8.4: Exclusion limits at 95% CL on four-body stop decays T2tt (left) and chargino-mediated decays T2bW (right) as a function of $m_{\tilde{t}_1}$ and Δm . A 100 % branching ratio for the chosen decay channel is assumed and the decays are assumed to be prompt. The colour shading corresponds to the observed limit on the cross section. The solid black (dashed red) lines represent the observed (expected) limits, derived using the expected top squark pair-production cross section. The thick lines represent the central values and the thin lines the variations due to the theoretical or experimental uncertainties. [5]

Additionally, a statistical combination with the results of the all-hadronic analysis [242] (red) has been performed [252], due to the complementarity of the searches. The hadronic limits are stronger due to the much higher branching ratios with respect to the leptonic channels. The combination of the limits shown in Figure 8.5, yielding the most stringent limits in this region of the SUSY parameter space with 2016 data.

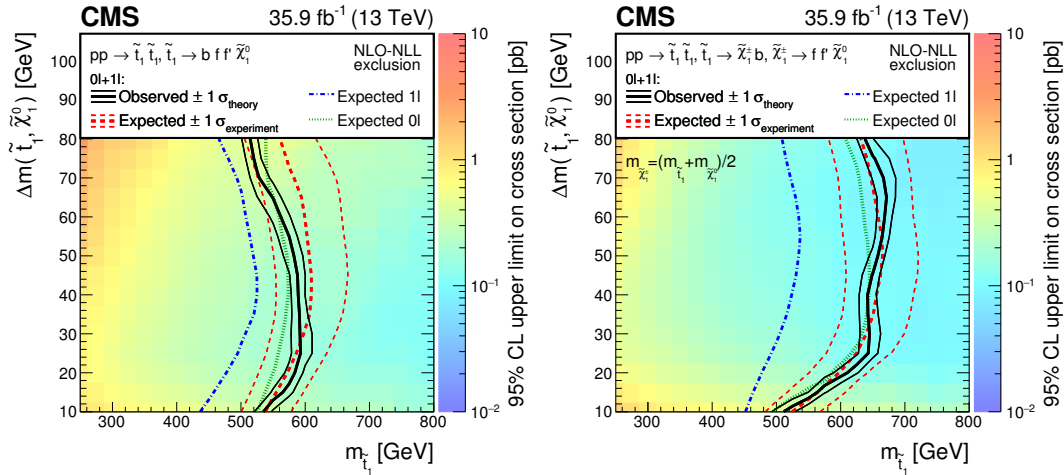


Figure 8.5: Exclusion limits at 95% CL on four-body stop decays $T2t\bar{t}$ (left) and chargino-mediated decays $T2bW$ (right) as a function of $m_{\tilde{t}_1}$ and Δm . A 100 % branching ratio for the chosen decay channel is assumed and the decays are assumed to be prompt. The colour shading corresponds to the observed limit on the cross section. The solid black (dashed red) lines represent the observed (expected) limits of a statistical combination of the single-lepton (blue) and all-hadronic channel analysis (red), derived using the expected top squark pair-production cross section. The thick lines represent the central values and the thin lines the variations due to the theoretical or experimental uncertainties. [5]

Part III

Soft Trigger

Additional Compressed Models

Charginos $\tilde{\chi}^\pm$ and neutralinos $\tilde{\chi}^0$ are collectively referred to as electroweakinos (EWKinos). As discussed in Section 2.2, they are mass eigenstates formed by mixtures of charged and neutral components of winos \tilde{W} , binos \tilde{B} and higgsinos \tilde{H} . Depending on the mass parameters (M_1 , M_2 , μ , $\tan \beta$) of the MSSM, which dictate the mixing of the mass eigenstates, they could have a wino, bino or higgsino-like nature.

Even though EWKino production has lower cross sections with respect to strong production at the LHC (as discussed in Section 5.1.2), it is of particular interest with strong theoretical motivations, where light higgsinos are favoured by naturalness arguments along with cosmological motivations [273, 274, 275]. Compressed SUSY searches have become sensitive to compressed EWK models, with larger datasets and more tailored analyses targeting clean multilepton signatures and low hadronic activity.

There are a number of simplified EWKino models that are considered, such as the production of chargino-neutralino ($\tilde{\chi}_1^\pm - \tilde{\chi}_2^0$), chargino-chargino ($\tilde{\chi}_1^+ - \tilde{\chi}_1^-$), neutralino-neutralino ($\tilde{\chi}_1^0 - \tilde{\chi}_2^0$). The indices correspond to masses in ascending order, with number one being the lightest. In most cases, the produced EWKinos are assumed to be mass-degenerate, and specific assumptions are made on their nature being wino- or higgsino-like.

The produced electroweakinos could have a wino and/or bino nature. This scenario considers the assumptions of $M_1 < M_2 \ll |\mu|$, discussed in Section 2.2, where typically the LSP ($\tilde{\chi}_1^0$) is bino-like and the NLSP ($\tilde{\chi}_2^0$) is wino-like. In such wino/bino

scenarios, the bino LSP is a strong DM candidate to match the observed relic density in compressed scenarios, as discussed in the context of light stops in Section 2.2.2.

Another scenario considers higgsino-like electroweakinos, where the lightest particles form a triplet of higgsino-like states $(\tilde{\chi}_2^0, \tilde{\chi}_1^\pm, \tilde{\chi}_1^0)$. It has a similar particle spectrum in which the mass splitting Δm between the states is partially determined by the magnitude of the wino (M_1) or bino (M_2) mass parameters relative to $|\mu|$. Light higgsinos are likely to be nearly degenerate in mass. Such a configuration is motivated by naturalness arguments that suggest that $|\mu|$ should be near the weak scale while M_1 and M_2 can be larger [123, 124, 274].

The production cross sections of wino/bino $\tilde{\chi}_1^\pm\text{-}\tilde{\chi}_2^0$ and $\tilde{\chi}_1^+\text{-}\tilde{\chi}_1^-$ production, as well as inclusive higgsino-like $\tilde{\chi}\tilde{\chi}$ production with a degenerate triplet, are shown in Figure 5.3. The wino/bino $\tilde{\chi}_1^\pm\text{-}\tilde{\chi}_2^0$ mode has slightly higher cross sections than the higgsino-like production of EWKinos, followed by wino-like chargino production.

Preliminary sensitivity studies of the different EWKino models indicated the analysis to potentially have limited sensitivity to two EWKino models [252]. At very small Δm , there is potential sensitivity to the chargino-neutralino ($\tilde{\chi}_1^\pm\text{-}\tilde{\chi}_2^0$) production shown in Figure 9.1, with a pure wino production cross section, calculated at NLO+NLL (Section 5.1.2). The $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$ are assumed to be mass degenerate and decay to the LSP via an off-shell $W * Z^*$ in the compressed mass region. The chargino $\tilde{\chi}_1^\pm$ (neutralino $\tilde{\chi}_2^0$) masses are scanned between $100 < m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} < 500$ GeV in steps of 25 GeV and the mass splittings with respect to the LSP are chosen between $3 < \Delta m < 50$ GeV. For $\Delta m > 10$ GeV, the steps are in 10 GeV, whereas, for $\Delta m < 10$ GeV the steps are in 3, 5, 7 GeV, to have higher granularity for the lowest mass splittings.

Another model to which the analysis indicated potential (albeit very limited) sensitivity is a compressed pMSSM-inspired model with a higgsino LSP [244, 245]. It is inspired by the phenomenological MSSM (pMSSM) [276] (19 free parameters), with additional constraints applied, to only consider several relevant parameters ($M_1, M_2, M_3, \mu, \tan \beta$). Trilinear couplings are not considered, and the gluino mass parameter M_3 is set very high and essentially decoupled. To reduce the dependence on $\tan \beta$, it is

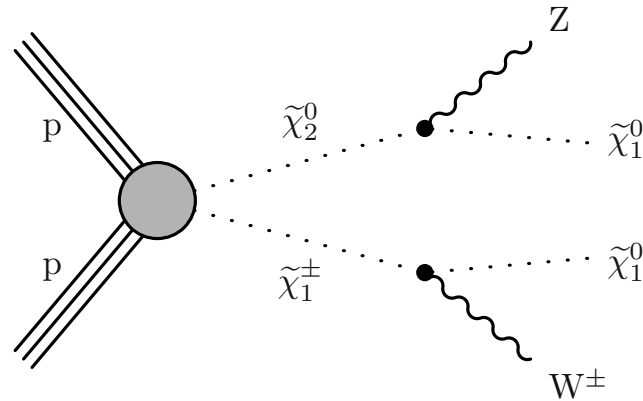


Figure 9.1: Simplified SUSY signal models for chargino-neutralino ($\tilde{\chi}_1^\pm\tilde{\chi}_2^0$) production (TChiWZ: $\tilde{\chi}_1^\pm\tilde{\chi}_2^0 \rightarrow W^*Z^*\tilde{\chi}_1^0\tilde{\chi}_1^0$).

set to a large value of 10. In this model, the common assumption that $M_1 \approx M_2/2$ is made, which is motivated by gaugino unification at the GUT scale. Ultimately, this results in two free parameters that are considered: the wino (M_1) and higgsino (μ) mass parameters. Henceforth, this model will simply be referred to as Higgsino pMSSM.

M_1 is scanned from 300 GeV-1200 GeV and μ is scanned from 100-240 GeV. In the pMSSM, larger μ values roughly correspond to larger masses for the parent sparticles, while larger values of M_1 correspond to smaller values of the Δm . The choice of parameters in the scan translates to values for a $\Delta m = \Delta m(\tilde{\chi}_1^\pm, \tilde{\chi}_1^0)$ mass difference ranging from 4-28 GeV. Cross sections are independently calculated for each model point in the Higgsino pMSSM space using the PROSPIN02 computational package [233]. The EWKino mass spectra, branching fractions and decay rates are determined from the Higgsino pMSSM parameters via additional computational tools described in [244, 245].

Soft Trigger Algorithms

10.1 Introduction

The choice of the trigger strategy is a fundamental element of an analysis, defining the specific data sample that is considered. An improvement of the trigger strategy can thereby extend the available phase-space and acceptance of the targeted signals. As discussed in Sections 3.3.6.4 and 6.4.2, a number of SUSY searches rely on an online trigger selection on pure p_T^{miss} , which plateaus ≈ 250 GeV. Lowering the p_T^{miss} trigger thresholds would further increase the acceptance for compressed signatures, and it can provide a window to less accessible SUSY models beyond the reach of the search, such as compressed EWKinos, presented in Section 9. Furthermore, with the increases in instantaneous luminosity and harsher pileup conditions at the LHC, discussed in Section 3.1.2, it is crucial to mitigate any potential increases in trigger thresholds with improvements in the trigger strategies.

10.2 Trigger Design

There are a number of considerations that need to be taken into account in the design of a new and improved trigger strategy. Firstly, it is important to define the targeted phase-space in terms of maximising the acceptance of the signal under consideration, while simultaneously finding a balance in terms of suppressing the backgrounds. Both elements are reflected in the expected trigger rates, which need to comply with the re-

requirements of the trigger systems and allocated bandwidth. The performance of the trigger needs to be evaluated in terms of efficiency in order to understand the viability and compatibility with the final analysis selection. Its stability and robustness need to be taken into account against the dynamic instantaneous luminosity and PU conditions of the LHC. Other considerations, such as the versatility, detector improvements and the interface between L1T and HLT algorithms, are key factors in the final design. Ultimately it comes down to finding an optimal balance between signal acceptance, background rejection, performance and versatility.

The *baseline*¹ analysis uses online trigger selection on pure p_T^{miss} , compatible with a compressed SUSY signature, as discussed in Section 6.4.2. The development of a new trigger strategy proceeds with an evaluation of the original trigger strategy as a baseline for improvement.

10.2.1 Signal Acceptance and Background Rejection

The trigger selection can be considered equivalent to an analysis cut in the context of a *cut & count* analysis, as discussed in Section 6.4.1, where the corresponding *acceptance* \times *efficiency* can be defined as:

$$\mathcal{A}_{trg} \times \epsilon_{trg} = \frac{N_{trg}^X}{N_{prod}^X} \quad (10.1)$$

where $N_{prod}^X = \sigma_X \int \mathcal{L} dt$ is the expected number of produced events and N_{trg}^X is the number of events accepted by the trigger for the signal or background processes (X). The target is to maximise the acceptance of the signal of interest $\mathcal{A}_{trg}^{sig} \times \epsilon_{trig} = N_{trg}^{sig}/N_{prod}^{sig}$, while simultaneously finding a reasonable balance in maximising the background rejection, defined as:

$$\mathcal{R} = 1 - \left(\frac{N_{trg}^{bkg}}{N_{prod}^{bkg}} \right) \quad (10.2)$$

¹Baseline refers to the 2016 analysis with 35.9 fb^{-1} data, presented in Part II.

Similarly, as with a general analysis approach, this can be performed by finding corners of parameter space where the signal-to-background ratio is maximised, with a focus on the dominant backgrounds.

To understand how the trigger strategy can extend the acceptance of the signals, it is important to re-examine the complete phase-space of the compressed SUSY signatures. As discussed in 5.2, a typical compressed SUSY signature comprises at least one high momentum jet, moderate p_T^{miss} and H_T , and soft leptons and jets. The ISR requirement circumvents the issue with very soft decay products but also significantly reduces the signal acceptance. Due to the boosted signal topology from ISR, there is a strong correlation between the p_T^{miss} and p_T^{ISR} , while the lepton p_T remains low. This is demonstrated in Figure 10.1, which shows the 2D distributions between the three main physics objects in the analysis, for the strong T2tt and EWKino TChiWZ signals, with masses close to the sensitivity limits and a medium Δm . The corresponding plots for the other T2bW and Higgsino pMSSM signals are shown in Appendix C. The general kinematics are relatively similar between the different targeted signals.

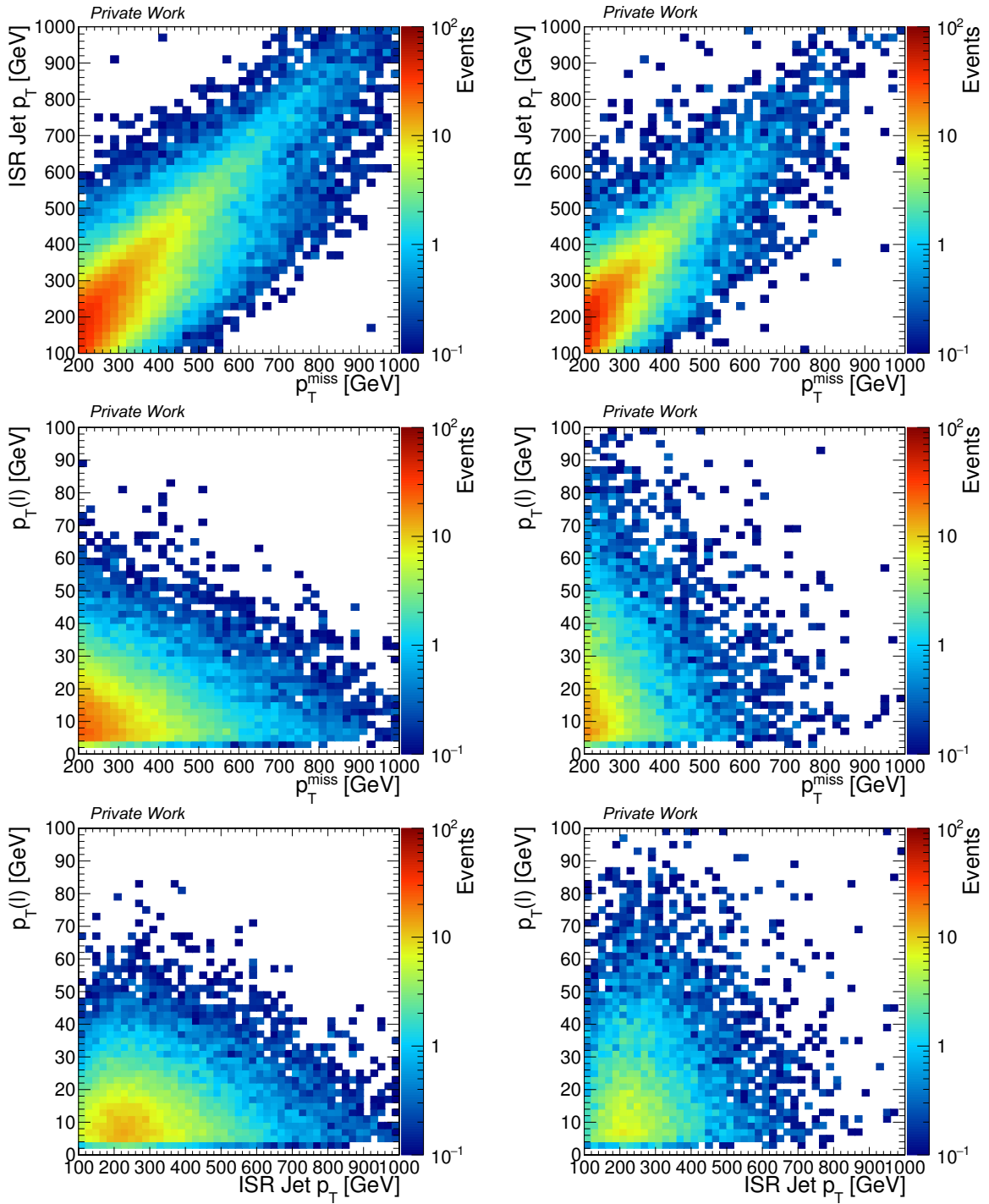


Figure 10.1: 2D signal distributions of p_T^{miss} , p_T^{ISR} and $p_T(\mu)$ for the strong T2tt : $\tilde{t}_1 \rightarrow b f' f' \tilde{\chi}_1^0$ (left) and EWKino TChiWZ : $\tilde{\chi}_1^+ \tilde{\chi}_2^0 \rightarrow W^* Z^* \tilde{\chi}_1^0 \tilde{\chi}_1^0$ (right) signals with masses close to the sensitivity limits and a medium Δm values.

Figure 10.1 indicates that despite the ISR boost of p_T^{miss} to moderate levels, the bulk of the different signals concentrates at even lower p_T^{miss} values than the level of the plateau

of the lowest unrescaled pure $p_T^{\text{miss}} + H_T^{\text{miss}}$ trigger that is used ($p_T^{\text{miss}} \approx 250$ GeV). In effect, a significant amount of signal acceptance is already lost at the level of the trigger selection. Ultimately, one can motivate the new trigger, especially by the EWKino signal having visibly smaller p_T^{miss} . Therefore, the prime target in developing a new trigger is to reduce the p_T^{miss} threshold in order to extend the probed phase-space.

10.2.2 Rates

Trigger rates are the main limiting factors in the context of the trigger systems with hard limits and need to be kept at a reasonably low level. Both signal acceptance and background rejection are reflected in the total rates of a given trigger.

The production rate of a given process (X) is defined in terms of the cross section σ_X and instantaneous luminosity \mathcal{L} :

$$R_{\text{prod}}^X = \frac{dN^X}{dt} = \mathcal{L} \times \sigma_X \quad (10.3)$$

In MC simulation, where one has the cross section data available from theoretical models, the rate of a given trigger is defined as:

$$R_{\text{trig}}^X = \frac{dN_{\text{trg}}}{dt} = \mathcal{L} \int \frac{d\sigma_X}{dp_T} \times \epsilon_{\text{trg}}(p_T) dp_T \quad (10.4)$$

where $\epsilon_{\text{trg}}(p_T)$ is the efficiency of the trigger in process X , expressed as a function of p_T acceptance. MC can be used to simulate new data-taking conditions of the detector, however, differences may arise if the simulation does not accurately reflect the data. Another complication is that one needs to run over a very large sample and include all different types of background processes.

Therefore, there is a preference to estimate the rates directly from data. It has the advantage that it automatically includes all processes and is more reliable than simulation. The trigger rates are proportional to the instantaneous luminosity \mathcal{L} and can be estimated from data by directly using the accelerator parameters (discussed in Section 3.1.2). These include the effective bunch-crossing frequency $\overline{f_{\text{BX}}} = f_{\text{LHC}} * n_{\text{bb}} =$

31.6 MHz, where f_{LHC} is the orbit frequency and $n_{\text{bb}} = 2808$ is the number of colliding bunches. This is also known as the *ZeroBias* rate $R_{ZB} \equiv \overline{f_{\text{BX}}}$, with the rate measured directly from the Beam Pickup Timing for the eXperiment (BPTX) detector. A fraction of these *ZeroBias* events is stored in every run within a *ZeroBias* dataset, providing a completely unbiased data sample with the only requirement being that a bunch-crossing has occurred. Such an unbiased dataset is necessary if the new HLT path also requires a new LIT seed. Calculating the fraction of bunch-crossing events that are accepted by the trigger provides the corresponding rate:

$$R_{\text{trig}}^{\text{data}} = \frac{dN_{\text{trig}}}{dt} = \overline{f_{\text{BX}}} \times F_{\text{trig}} = R_{ZB} \cdot \left(\frac{N_{\text{trig}}}{N_{ZB}} \right) \quad (10.5)$$

where $F_{\text{trig}} = R_{\text{trig}}/R_{ZB} = N_{\text{trig}}/N_{ZB}$. The fraction F_{trig} can be determined by re-running the LIT and HLT over existing *ZeroBias* data and counting the number of times the trigger fired N_{trig} , while N_{ZB} is the sample size.

If the new HLT path does not require a new LIT seed, the measurement can be performed in an *HLTPHysics* dataset that collects a fraction of the data that is already pre-selected by the LIT menu and does not require any HLT path. In this case, an alternative approach is also possible:

$$R_{\text{trig}}^{\text{data}} = \frac{dN_{\text{trig}}}{dt} = \frac{N_{\text{trig}}}{N_{LS} \cdot T_{LS}} \quad (10.6)$$

where N_{LS} is the number of *luminosity sections* (LS), which is a section of data subdivided using an predefined reference time interval $T_{LS} = 23.31 \text{ s}$.

If a prescale is applied at any point on the dataset, then this would need to be taken into account by a multiplicative factor, as in the case of the heavily-prescaled *HLTPHysics* datasets.

One caveat is that the estimated rates correspond to the LHC and detector conditions in that the data was taken. Since the rate is directly proportional to luminosity \mathcal{L} , if one assumes a linear dependence, one can perform a simplified scaling to a given target luminosity. The luminosity \mathcal{L} is convoluted with pileup, both dependent on

the beam parameters, which breaks the linearity. Therefore, the dependence of rate as a function of pileup is an important consideration, especially in the context of the increasing instantaneous luminosity and PU conditions (discussed in Section 3.1.2).

10.2.3 Efficiency

The trigger efficiency is defined as the ratio of the number of events that pass a given trigger selection to the total number of selected events in a given sample:

$$\epsilon_{trg} = \frac{N_{offline \& \text{trg}}}{N_{offline}} \quad (10.7)$$

where the denominator $N_{offline}$ is the number of events passing an offline selection and the numerator $N_{offline \& \text{trg}}$ is the number of events passing an offline selection and firing the trigger under consideration (`HLT_MainTrigger`). Therefore, for an efficiency measurement of a given trigger, the denominator effectively defines the numerator. The offline selection indicates the parameter space where the measurement is performed in. It can strongly affect the result and requires careful consideration to provide an accurate and unbiased measurement. The offline selection should be as close as possible to the final analysis selection, in order to reduce any potential kinematical biases. Additional requirements can be imposed in order to reduce the effect of misidentified objects and ensure the purity of the sample.

There are a number of methods that can be used to perform an efficiency measurement, which have their advantages and disadvantages and are specific to the trigger and analysis under consideration. Simulation can be used to re-emulate the L1T and HLT reconstruction, however, similarly to the rate estimation, such a measurement can be inaccurate in terms of detector conditions. Therefore, there is a preference to perform the measurements directly using recent data, where the denominator (and consequently numerator) selection also comprises a reference trigger (`HLT_ReferenceTrigger`). In this case, the offline selection must additionally be made compatible with the reference trigger itself.

$$\epsilon_{trig} = \frac{N(\text{HLT_MainTrigger} + \text{HLT_ReferenceTrigger} + \text{offline selection})}{N(\text{HLT_ReferenceTrigger} + \text{offline selection})} \quad (10.8)$$

The reference trigger could be a prescaled looser version of the original trigger under consideration. Alternatively, events where objects are pair-produced, such as leptons on the Z-peak, can be used to measure the efficiency of a single-object trigger. This is a variation of the *tag-&-probe* method, discussed in Section 6.3.2, where the requirement of the *probe* corresponds to the offline selection and the *tag* corresponds to the triggering object. A common approach is performing the measurement in an orthogonal dataset, defined by the reference trigger, to ensure that the sample is unbiased. An example is the measurement of a H_T or p_T^{miss} trigger in a dataset populated by single-muon events. Naturally, all these methods assume sufficient knowledge of the efficiency of the reference triggers.

The trigger efficiency is typically measured as a function of p_T and η on the physics objects used in the trigger algorithms themselves. Its dependence on other relevant or analysis-dependent observables, such as pileup, is important as well and motivates a multi-dimensional study. The efficiency is typically measured with respect to offline reconstructed quantities, which are more accurate representations of the truth. Consequently, this enables a more precise evaluation of the effect of the trigger on the analysis and its corresponding systematics.

Typical trigger efficiency *turn-on* curves are shown earlier in Section 3.3.6.4 and later in Section 10.4.3. They have a characteristic shape that can normally be parametrised with a Gauss error function:

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (10.9)$$

The sharpness or width of the turn-on is a reflection of the resolution of the physics trigger objects. Impurities due to background contamination, inefficiencies and dif-

ferences with respect to offline quantities are also reflected in the shape of the turn-on. For an HLT path, the total efficiency incorporates the implicit LIT seed efficiency. Depending on the differences in trigger object resolutions between LIT and HLT and the differences in thresholds between the paths, it can happen that the LIT turn-on cuts into the total HLT efficiency. This would result in a double turn-on, which is more difficult to parametrise and account for. Therefore, it is important that the HLT paths and corresponding LIT seeds are simultaneously optimised not only in terms of rates but also efficiencies.

Any conceivable analysis selection is typically chosen to be beyond the “knee” of the turn-on on the “plateau” of constant efficiency, where the efficiency is expected to be independent on the threshold parameter. Therefore, trigger efficiency turn-ons need to be sufficiently sharp not to limit the parameter space of the analysis. Any inefficiencies need to be accounted for in the analysis systematics, which can be extrapolated from the parametrisation of the turn-on. Thus, the trigger efficiency must be precisely known for the correct estimation of the total efficiency of the analysis selection and the corresponding uncertainty systematics. As defined in Equation 6.2, the total efficiency can be expanded as:

$$\epsilon_{total} = \epsilon_{tracking} \times \epsilon_{reco} \times \epsilon_{LIT} \times \epsilon_{HLT} \times \epsilon_{vertex} \times \epsilon_{analysis} \quad (10.10)$$

where $\epsilon_{trg} = \epsilon_{LIT} \times \epsilon_{HLT}$ is the total LIT and HLT trigger efficiency, corresponding to the seed and path efficiencies, respectively.

10.2.4 Other Considerations

A pragmatic approach is the development of a dedicated cross-trigger, with several requirements on objects, that can significantly reduce the rate with respect to single objects and thus allow for the reduction of trigger thresholds. Key observables used in the analysis can serve as handles in constructing such a trigger. In order to maintain the versatility to different models, the focus is on the main physics objects in the targeted compressed SUSY signatures, discussed in Sections 5.2 and 10.2.1. The pres-

ence of a high p_T ISR jet is a common feature of compressed SUSY analyses and can be used as an additional handle. The strong correlation of p_T^{miss} and p_T^{ISR} in ISR-boosted topologies (shown in Figure 10.1) needs to be taken into account in the context of rate reduction. The requirement of a high p_T jet in an event can also help mitigate the high PU-dependence of p_T^{miss} triggers (Figure 3.19). To take advantage of very soft objects at the trigger level, one needs to push to the limits of detection and trigger systems. As discussed in Section 3.3, the experiment is able to efficiently identify muons down to 3 GeV [277] and electrons down to 4.5 GeV, while the analysis considers muons with $p_T > 3.5$ GeV and electrons with $p_T > 5$ GeV. The target is to go as low as possible in lepton p_T , taking into consideration that the rates are expected to be very high for low thresholds, which would also select on backgrounds, such as misidentified low p_T jets. Such a cross-trigger requires the development of both an HLT path and a LIT seed with corresponding requirements. The two need to be designed to be compatible with one another and be optimised simultaneously in terms of rates and efficiency. Taking into account that a realistic cross-trigger would necessitate several requirements on different objects, it would be categorised as a more exclusive trigger type. Given that additional requirements limit the targeted signatures and, therefore, analysis use-cases, this also limits the possible available bandwidth in the trigger menus. The ballpark for similar cross-triggers at peak luminosity and PU is less than 1 kHz for an LIT seed and an order of 10 Hz for the HLT path, for which the timing should be of the order of a few milliseconds. These are the target rates that have been used in the subsequent study.

The trigger should also demonstrate reasonable robustness against the dynamic luminosity and PU conditions at the LHC, which affect both the rates and efficiency. Specifically, the high rate versus PU-dependence of energy sums (p_T^{miss} , H_T) at the trigger levels (Figure 3.19) has been problematic and a focus of a number of improvements of the trigger systems and algorithms. Stability in the form of reduced PU-dependence would mitigate any potential unwanted increases in thresholds.

10.3 Trigger Development

The studies involved in the development of the trigger were performed both on collected data and privately-generated MC samples. The tools used in the development of the algorithms include the Trigger Menu Editor (TME) for the L1T seed development and the ConfDB GUI for the development of the HLT path, which are described in Appendix C.5; more details on the general internal structure of the L1T and HLT algorithms, including objects and requirements, are also described.

10.3.1 Cross-Trigger

As a first step, the viability of a cross-trigger can be estimated with the evaluation of the rates and efficiencies of single-object triggers. The performance of the main single-object triggers in terms of rates and efficiencies is presented in Section 3.3.6.4. The rates of the lowest object thresholds are shown in Tables 3.2 and 3.3 and summarised in Table 10.1 for objects characteristic of compressed SUSY signatures.

Object	L1T Algorithm	Rate (kHz)	HLT Path	Rate (Hz)
Muon	L1_SingleMu22	8	HLT_IsoMu24	235
Electron	L1_SingleLooseIsoEG28er2p5	27	HLT_Ele32_WPTight_Gsf	165
p_T^{miss}	L1_ETMHF100	3	HLT_PFMET120_PFMHT120_IDTight	33
Jet	L1_SingleJet180	3	HLT_PFJet500	11
H_T	L1_HTT360er	4	HLT_PFHT1050	10

Table 10.1: Lowest thresholds of single-object triggers and rates at a $\mathcal{L} = 1.8 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ ($\text{PU} \approx 50$) relevant to typical compressed SUSY signatures.

For most objects (muons, electrons and p_T^{miss}), the difference between the thresholds between L1T and HLT, dictated by the rate reduction and coherence of efficiency curves, is relatively small (2-20 GeV). Given the lowest unrescaled thresholds of single-object seeds, it is clear that any development of a new HLT cross-path with a significant reduction of p_T^{miss} would additionally require the development of a dedicated L1T cross-seed.

Based on the thresholds and rate values in Table 10.1 versus the target rates (1 kHz at L1T and 10 Hz at HLT), a realistic cross-trigger would comprise a combination of re-

quirements on a lepton, p_T^{miss} and/or jet/ H_T . Electrons are more difficult to reconstruct than muons and have lesser purity, especially at the LIT stage, which translates to very high rates. Internal studies indicated significant difficulty in designing a viable low- p_T electron trigger. Therefore, a choice was made to focus on the muon channel only. Due to the already very high thresholds on H_T , related to the high jet multiplicity in a hadron collider, alternatively, a choice was made to consider the leading jet p_T requirement, which is also more characteristic of a compressed SUSY signal. Therefore, the study focused on different combinations of a low p_T muon, high- p_T leading jet and p_T^{miss} , including additional requirements on those objects.

10.3.2 Rate Estimation

The rate estimation for both the LIT seeds and HLT paths is performed using RAW data collected in 2017. Given that the developed paths include new seeds, the data needs to be re-processed with a re-emulation of the LIT with a custom menu. The data used in the measurement comes from unbiased *ZeroBias* (ZB) datasets, which ultimately only have the requirement that a bunch-crossing has occurred. These are populated by a `HLT_ZeroBias` path, which is a heavily-prescaled pass-through of the `L1_ZeroBias` seed selecting only valid bunch-crossing events. The LIT algorithm is based purely on an external signal coming from the BPTX detectors.

Given that the majority of ZB events are contaminated by inelastic or background events, a large number of statistics is needed to receive a significant number of triggered events and, thus, a reasonable statistical precision. Therefore, rather than using a subset of the most recent data, the full 2017 dataset is considered. However, this effectively averages out the luminosity and pileup conditions that have varied throughout the year. Therefore, for a more realistic result, the rates must be normalised by a luminosity scale factor representative of the latest conditions. Assuming a linear scaling of the rate with luminosity, one can determine this factor using muon triggers, which are known to have clean signatures and a linear dependence of rate. Taking the core single muon `HLT_IsoMu27` trigger as reference, the factor is calculated as ≈ 1.5 ,

indicating a $\approx 50\%$ increase with respect to the original estimates, corresponding to a luminosity $\mathcal{L} = 1.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. Another factor of ≈ 1.37 would be required to scale to $\mathcal{L} = 2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, which was the benchmark for studies and corresponds to the peak luminosity for 2018.

10.3.3 LIT Seed and HLT Path Development

Preliminary studies indicated that the main limiting factor is the LIT rate requirements. Therefore, a focus was placed on first developing and measuring the performance of a potential LIT seed, which would then form the baseline for the corresponding HLT path.

LIT Seed

The development of the LIT seed proceeded using TME within a custom LIT menu. To better understand the rate dependency and its correlations, approximately 200 different algorithms were created with varying p_T thresholds on different objects (muon, jet, p_T^{miss}), as well as additional requirements, such as restrictions on the pseudorapidity $|\eta|$. Table 10.2 shows a subset of these seeds, indicating the dependency of the rate on various selections, and the necessary stringency of the cuts to achieve the target rate of the order of $\approx 1 \text{ kHz}$.

The simple requirement of a muon with a standard *single muon* quality cut (selection on number of muon detector layers), with an optimised selection on single muon events, and no additional requirements on acceptance results in an extremely high rate, as expected. Adding a p_T requirement of 3 GeV reduces the rate by almost half, which was chosen as the main reference point. This also corresponds to the minimum muon p_T cut in the barrel (BMTF), as lower p_T muons can lose sufficient energy in the calorimeters that they do not reach the second station of muon chambers, which is required for standard muon reconstruction. On the other hand, increasing the muon p_T threshold further would be undesirable due to the compressed signal mainly lying

in the lowest lepton p_T ranges, as shown in Figure 10.1.

LIT Seed	Thresholds	Rate (kHz)
L1_SingleMuOpen	μ with <i>single muon</i> quality (MU-QLTY_SINGL)	2077
L1_SingleMu3	$p_T(\mu) > 3 \text{ GeV}$	1211
L1_SingleMu3er2p1	$p_T(\mu) > 3 \text{ GeV}; \eta < 2.1$	987
L1_SingleMu3er1p5	$p_T(\mu) > 3 \text{ GeV}; \eta < 1.5$	535
L1_SingleMu3_BMTF	$p_T(\mu) > 3 \text{ GeV}; \eta < 0.8$	204
L1_SingleMu3Neg	$p_T(\mu) > 3 \text{ GeV}; Q(\mu) = -1$	619
L1_SingleMu3_SingleJet70	$p_T(\mu) > 3 \text{ GeV} + \text{jet } p_T > 70 \text{ GeV}$	40
L1_SingleMu3_SingleJet90	$p_T(\mu) > 3 \text{ GeV} + \text{jet } p_T > 90 \text{ GeV}$	10
L1_SingleMu3_SingleJet100	$p_T(\mu) > 3 \text{ GeV} + \text{jet } p_T > 100 \text{ GeV}$	6
L1_SingleMu3_SingleJet110	$p_T(\mu) > 3 \text{ GeV} + \text{jet } p_T > 110 \text{ GeV}$	3
L1_SingleMu3_SingleJet120	$p_T(\mu) > 3 \text{ GeV} + \text{jet } p_T > 100 \text{ GeV}$	2
L1_SingleMu3_SingleJet90_ETMHF30	$p_T(\mu) > 3 \text{ GeV} + \text{jet } p_T > 90 \text{ GeV} + p_T^{\text{miss}} > 30 \text{ GeV}$	7
L1_SingleMu3_SingleJet90_ETMHF50	$p_T(\mu) > 3 \text{ GeV} + \text{jet } p_T > 90 \text{ GeV} + p_T^{\text{miss}} > 50 \text{ GeV}$	4
L1_SingleMu3_SingleJet90_ETMHF60	$p_T(\mu) > 3 \text{ GeV} + \text{jet } p_T > 90 \text{ GeV} + p_T^{\text{miss}} > 60 \text{ GeV}$	3
L1_SingleMu3_SingleJet100_ETMHF30	$p_T(\mu) > 3 \text{ GeV} + \text{jet } p_T > 100 \text{ GeV} + p_T^{\text{miss}} > 30 \text{ GeV}$	5
L1_SingleMu3_SingleJet100_ETMHF50	$p_T(\mu) > 3 \text{ GeV} + \text{jet } p_T > 100 \text{ GeV} + p_T^{\text{miss}} > 50 \text{ GeV}$	2.5
L1_SingleMu3_SingleJet100_ETMHF60	$p_T(\mu) > 3 \text{ GeV} + \text{jet } p_T > 100 \text{ GeV} + p_T^{\text{miss}} > 60 \text{ GeV}$	2

Table 10.2: LIT Seed Rates at $\mathcal{L} = 1.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. All *SingleMu* seeds have a *single muon* quality (MU-QLTY_SINGL) selection applied. "er" stands for "eta-restricted".

The effect of various cuts on pseudorapidity η has also been investigated (e.g. $|\eta| < [2.1, 1.5, 0.8]$), motivated by the detector segmentation (Table 3.1). As shown in Figure 10.2, there is a much higher rate for low- p_T muons in the endcaps (EMTF). Therefore, restrictions to the barrel area of the detector would provide significant rate reduction while being compatible with the minimum p_T selection of BMTF. Such a selection is also motivated by the fact that in a back-to-back ISR topology, the decay products of the considered signal are more centrally produced than those of dominant backgrounds, as discussed in Section 6.4.4.4. A cut on the muon charge was investigated as well, as the requirement of a negative lepton charge could provide some signal-to-background separation, due to charge asymmetry in the production of $W + \text{jets}$, as discussed in Section 6.4.4.4. However, such a requirement is also rather exclusive and incompatible with the EWKino signals that are considered, so, ultimately, was not considered.

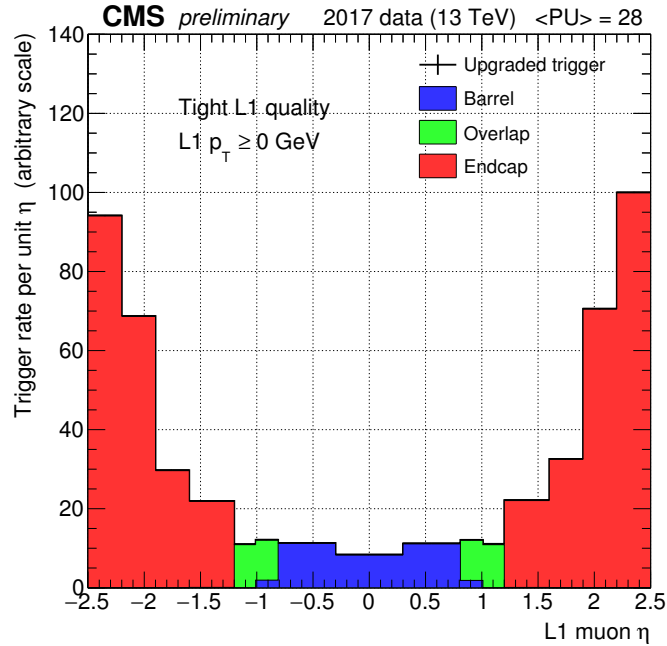


Figure 10.2: Distribution of muons in bins of η built by the three track finders (barrel, overlap, and endcap) in the upgraded L1 muon trigger (μ GMT) [278].

Consequently, a collection of the most realistic seeds with rates in the same ballpark as the target rate of ≈ 1 kHz is shown in Table 10.3. The η -restriction on the jet is motivated by the offline selection on jets, providing reduction of rate from the endcaps.

LIT Seed	Thresholds	Total Rate (kHz)
L1_SingleMu3_SingleJet110_ETMHF50	$p_T(\mu) > 3 \text{ GeV} + \text{jet } p_T > 110 \text{ GeV} + p_T^{\text{miss}} > 50 \text{ GeV}$	1.7
L1_SingleMu3er1p5_SingleJet110_ETMHF50	$p_T(\mu) > 3 \text{ GeV}; \eta < 1.5 + \text{jet } p_T > 110 \text{ GeV} + p_T^{\text{miss}} > 50 \text{ GeV}$	0.8
L1_SingleMu3er1p5_SingleJet110_ETMHF30	$p_T(\mu) > 3 \text{ GeV}; \eta < 1.5 + \text{jet } p_T > 110 \text{ GeV} + p_T^{\text{miss}} > 30 \text{ GeV}$	1.2
L1_SingleMu3er1p5_SingleJet110er2p4_ETMHF30	$p_T(\mu) > 3 \text{ GeV}; \eta < 1.5 + \text{jet } p_T > 110 \text{ GeV}; \eta < 2.4 + p_T^{\text{miss}} > 30 \text{ GeV}$	0.9
L1_SingleMu3er1p5_SingleJet110er2p4_ETMHF50	$p_T(\mu) > 3 \text{ GeV}; \eta < 1.5 + \text{jet } p_T > 110 \text{ GeV}; \eta < 2.4 + p_T^{\text{miss}} > 30 \text{ GeV}$	0.8
L1_SingleMu3_SingleJet120er2p4_ETMHF40	$p_T(\mu) > 3 \text{ GeV}; \eta < 1.5 + \text{jet } p_T > 120 \text{ GeV} + p_T^{\text{miss}} > 30 \text{ GeV}$	1.1

Table 10.3: Realistic LIT seed rates at $\mathcal{L} = 1.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. All **SingleMu** seeds have a single muon quality (MU-QLTY_SNGL) selection applied.

HLT Path

The corresponding HLT path is primarily based on the LIT seed definition, comprising the same types of physics trigger objects. Since the HLT has access to the complete detector information, HLT objects are reconstructed by taking advantage of some of

the standard reconstruction methods described in Section 4, tailored towards fast processing. The reconstruction of muons is performed in two steps: the first is equivalent to offline standalone muon reconstruction (L2), followed by a combination with hits in the inner tracking system (L3). The reconstruction at L3 is seeded by an L2 muon and follows an iterative track reconstruction. The jet reconstruction and calculation of energy sums are performed with the particle flow (PF) algorithm.

For the corresponding p_T^{miss} -like selection at HLT, the path follows the $p_T^{\text{miss}} + H_T^{\text{miss}}$ definition of the baseline analysis trigger (see Section 6.4.2, and Tables 3.3 and 3.4), which requires PF-computed p_T^{miss} (*PFMET*) and missing hadronic energy H_T^{miss} (*PFMHT*) above a common energy threshold, along with *Tight* identification criteria applied on the jets that are used to compute the H_T^{miss} .

As an alternative, *PFMETNoMu* and *PFMHTNoMu* variations of the paths are also investigated, which do not include muon energies in the energy sum calculations. This is achieved by taking the magnitude of the p_T^{miss} or H_T^{miss} , reconstructed with the PF algorithm at HLT, excluding the contribution of muon candidates to the vector sums. At the Level-1 trigger, jet and p_T^{miss} reconstruction is based purely on calorimeter quantities and does not include muons. Therefore, this aligns the definitions of p_T^{miss} and H_T^{miss} together at L1T and HLT and provides a better correspondence in terms of optimising L1T seeds to their corresponding HLT paths.

The specific order of the modules is shown in Figure A.11 in Appendix C.5: muon (L2, L3), jet and then H_T^{miss} & p_T^{miss} . The first requirement is that of a muon, which is a strong filter due to the electroweak nature of the particle. This is followed by reconstruction and filtering on simplified p_T^{miss} , H_T^{miss} and jet based on calorimeter quantities only. The path sequence ends with a more processing-heavy PF calculation of the energy sums and jet.

The final thresholds on the HLT physics objects are primarily chosen based on rate estimates, as well as efficiency measurements, and their compatibility with the L1T seeds, as well as the final offline selection. In some cases it was affordable to choose selections that are close to each other, due to sharp turn-ons and/or rate allowance.

10.3.4 Efficiency in Simulation

Efficiency measurements are crucial in assessing the viability of the potential triggers, especially for determining the final analysis selection compatible with the turn-on plateau. The measurements are important in the context of very low p_T muons, which are in a non-standard regime. Furthermore, the measurement of the p_T^{miss} leg of the trigger would provide a measure of the relative gain with respect to the pure p_T^{miss} triggers of the baseline analysis. Moreover, there must be a compatibility between the efficiencies of the LIT seeds and HLT paths, in order to make sure that the LIT efficiency does not cut into the HLT one. Therefore, the efficiency studies of both the LIT seeds and HLT paths were performed in parallel.

The studies are performed on MC signal samples that were generated privately by preparing a full generation chain (described in Section 4.3). The motivation for this is in order to perform efficiency studies on reconstructed quantities, as opposed to generated quantities, providing a more realistic performance measurement that takes detector resolutions into account. The $T2\tau\tau$ signal model with masses at the analysis sensitivity limits ($m_{\tilde{\tau}_1} = 500$ GeV, $\Delta m = 40$ GeV) is used in order to measure and optimise the trigger efficiencies on the considered signal in the phase-space of interest.

In order to integrate custom triggers as part of the event data, the private MC generation procedure proceeded in three major steps, each using the output of the previous step as input. The first generation and simulation stage only had to be done once to prepare the data in the correct format (GEN-SIM and ROOT). The latter two DIGI-RAW-Trigger and RECO steps needed to be repeated for each modification of the LIT seed or HLT path done using the TME or ConfDB. The final ROOT file was output in the AOD format, containing reconstructed quantities on which realistic studies could be performed.

For the measurement of the trigger efficiency, the offline selection primarily comprises a loose muon ID selection criterion, corresponding to the final muon object definition of the baseline analysis. This ensures a high-purity sample. Any remaining selections involve requiring to be on the plateau of the observables that are not being measured

and are reflected by the trigger selection itself (including the pseudorapidity).

Measurements of the final selected HLT paths (discussed below) are shown in Figure 10.3. The measurements of the corresponding LIT seeds are included in Figure A.5 in Appendix C.2. The shapes of the turn-ons for the different objects are compatible with the central efficiency measurements in Section 3.3.6.4 and the plateaus are reached at values ($p_T(\ell) > 6$ GeV, $p_T^{\text{ISR}} > 130$ GeV, $p_T^{\text{miss}} > 200$ GeV) that are reasonable for the final analysis selection, especially in the non-standard low- p_T regime for the muon leg.

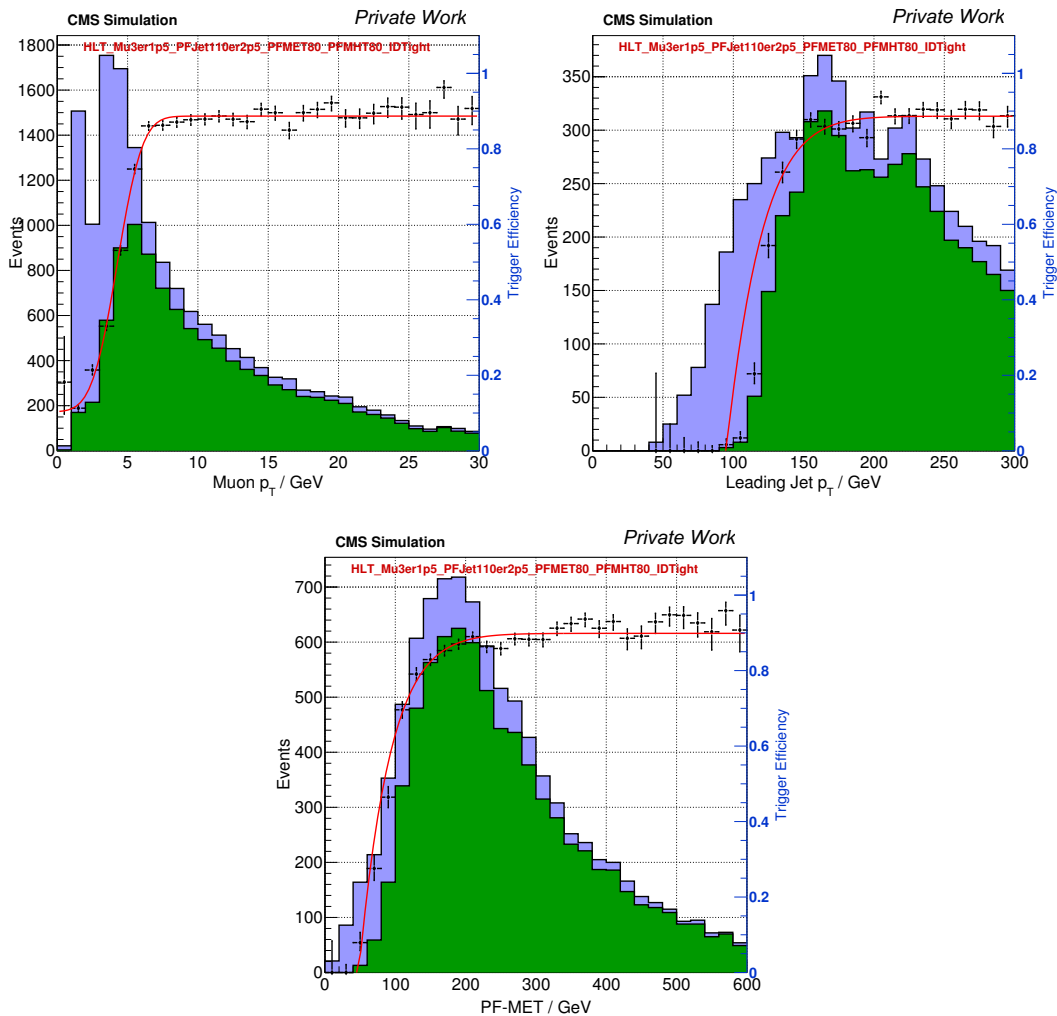


Figure 10.3: Efficiency (LIT+HLT) of the muon, jet and p_T^{miss} legs of a soft ($p_T > 3$ GeV) muon, high $p_T > 110$ GeV jet and p_T^{miss} cross-trigger measured in a simulated T2tt signal sample ($m_{\tau_1} = 500$ GeV, $\Delta m = 40$ GeV). The blue histogram corresponds to the denominator, whereas the green histogram corresponds to the numerator. The efficiency is parametrised with an error function (Equation 10.9).

10.3.5 Final Trigger

Ultimately, a single type of *soft μ + hard jet + moderate p_T^{miss}* cross-trigger was proposed to be included in the trigger menus, which contain the full set of LIT seeds and HLT paths, effectively encapsulating the entire physics programme of the experiment. The HLT paths with four different p_T^{miss} (H_T^{miss}) thresholds are shown in Table 10.4. The final threshold values were motivated primarily by the rates, as well as the efficiency turn-ons. Versions of the triggers with energy sums without the muon momenta (labelled *NoMu*) have also been included, as they are more aligned with the LIT definitions. All the HLT paths are seeded by a logical *OR* of corresponding seeds listed in Table 10.5.

Type	HLT Path	Rate (Hz)
Optimistic	HLT_Mu3er1p5_PFJet100er2p5_PFMET(NoMu)70_PFMHT(NoMu)70_IDTight	-
Baseline	HLT_Mu3er1p5_PFJet100er2p5_PFMET(NoMu)80_PFMHT(NoMu)80_IDTight	8 (13)
Backup	HLT_Mu3er1p5_PFJet100er2p5_PFMET(NoMu)90_PFMHT(NoMu)90_IDTight	3 (6)
Backup	HLT_Mu3er1p5_PFJet100er2p5_PFMET(NoMu)100_PFMHT(NoMu)100_IDTight	-

Table 10.4: Final HLT path proposal with two (Mu/NoMu) versions of four paths with PF- p_T^{miss} and PF- H_T^{miss} thresholds ranging from 70-100 GeV.

Each path is seeded by the following *OR* of dedicated LIT seeds, shown in Table 10.5.

Type	LIT Seed	Total Rate (kHz)	Pure Rate (kHz)
Baseline	L1_SingleMu3er1p5_SingleJet100er2p5_ETMHF40	1.3 ± 0.2	0.7 ± 0.1
Backup	L1_SingleMu3er1p5_SingleJet100er2p5_ETMHF50	-	-

Table 10.5: Final LIT seed proposal with two seeds with two thresholds on p_T^{miss} (ETMHF) of 40 and 50 GeV.

The selected thresholds and η -restrictions find a reasonable balance in the trade-off between rate reduction and signal acceptance. The inclusivity and versatility of the trigger were taken into account, in order to not be too restrictive and thus usable by other analysis groups targeting ISR-like leptonic signatures.

The tables list the estimated pure rates on top of simplified LIT and HLT menus containing the core paths and seeds, using data collected at a $\text{PU} = 57$, $\mathcal{L} = 1.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. Comparing the rate results to the target rates, trigger thresholds of $\text{PFMET} (p_T^{\text{miss}}) > 80 \text{ GeV}$ and $\text{PFMHT} (H_T^{\text{miss}}) > 80 \text{ GeV}$ at HLT and $\text{ETMHF} (p_T^{\text{miss}})$

> 40 GeV at LIT have been proposed as the baseline values. An “optimistic” version of the HLT path with a 10 GeV lower *PFMET*/*PFMHT* threshold has also been included as an alternative. Higher threshold triggers (90, 100 GeV at HLT and 50 GeV at LIT) have been included as backup, in case the LHC or detector conditions cause the rates to increase.

The final threshold choice on the p_T^{miss} leg for both the HLT path and LIT seed is dependent on several factors, such as the LHC conditions, which have increased luminosity and pileup with respect to previous years. In this context, improvements for p_T^{miss} objects are being developed in the form of pileup mitigation techniques. Furthermore, the overlap with new similar paths (e.g. soft dimuon, soft VBF + μ and hard jet + p_T^{miss} triggers) dictates the actual pure rate with respect to simplified menus.

Following dedicated approval procedures, the LIT seeds and HLT paths were included in the first versions of the 2018 trigger menus. Furthermore, a dedicated data quality monitoring (DQM) [176] module was developed to perform efficiency measurements online during data-taking. The trigger algorithms were deployed online to collect data at the start of the 2018 data-taking period of LHC *Run 2* and are currently collecting data during the on-going *Run 3*. The proposed baseline thresholds were accepted as the lowest unprescaled ones. Due to the higher rates of the *NoMu* versions of the paths with respect to the baseline paths, they ended up being prescaled.

Ultimately, one can consider

`HLT_Mu3er1p5_PfJet100er2p5_PFMET80_PFMHT80_IDTight`

seeded by

`L1_SingleMu3er1p5_SingleJet100er2p5_ETMHF40`

as the main trigger. The final ConfDB implementation of the HLT path is shown in Figure A.11 in Appendix C.5.

10.4 Performance

As described above in Sections 10.2.2 and 10.2.3, trigger performances are primarily assessed in terms of rates and efficiencies. For HLT paths, it is also vital to assess the timing of the trigger.

10.4.1 Online Rate

The rates versus pileup dependence of the main *soft μ + hard jet + moderate p_T^{miss}* path from selected CMS runs in 2018 are shown in Figure 10.4, as measured online by the CMS Rate Monitoring tools [279] during online data-taking.

At a luminosity of $\mathcal{L} = 2.0 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ and $\text{PU} \approx 50$, the total HLT rate for the main path is ≈ 12 Hz, which adds ≈ 3 Hz to the full HLT menu (pure rate). Correspondingly, the LI rate ≈ 1.5 kHz, which has a pure rate of ≈ 700 Hz. Therefore, the actual total rates are within the same ballpark as estimates and extrapolations from 2017 data.

As anticipated for p_T^{miss} -type triggers, the PU-dependence is non-linear. However, it is more linear than pure p_T^{miss} triggers and the gradient is lower, so the rate is under better control. This is partially due to the additional muon requirement, which has a linear dependence on PU, while the presence of a high- p_T jet is known to reduce the PU-dependence of triggers due to the boosted topology.

10.4.2 Timing

Apart from the rate, the timing (or execution time) of the HLT paths must be kept under control. Separate measurements of individual paths are required to ensure that the total timing budget of ≈ 450 ms (Figure 3.23) is not exceeded. Exclusive paths typically have an average timing of the order of a few ms and this is the ballpark target value for the HLT path.

The measurement is performed on benchmarked computing nodes, where one is able to run timing studies in an environment similar to that of standard data-taking conditions. The measurement proceeds by executing the HLT configuration file for the processing of prepared input data, running the modules of the path.

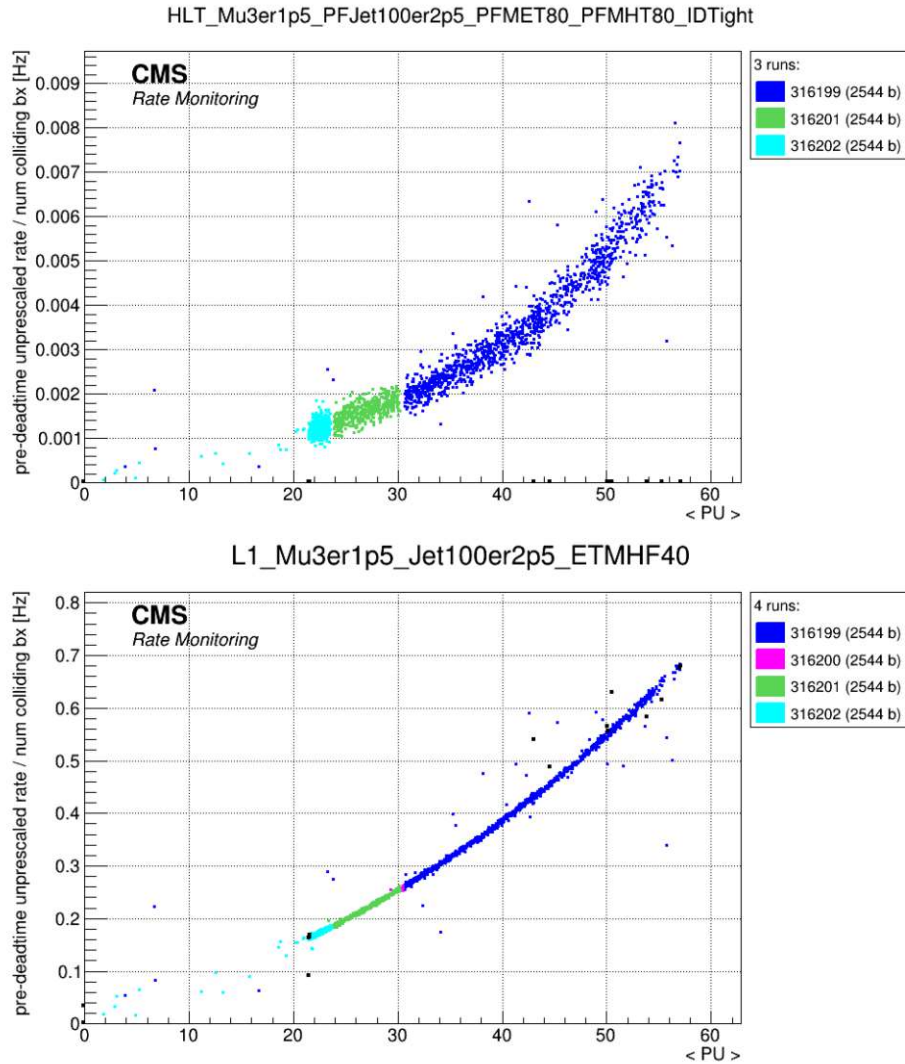


Figure 10.4: Rate vs PU plots of the main *soft μ + hard jet + moderate p_T^{miss}* HLT path (top) and corresponding L1T seed (bottom) produced by the online CMS Rate Monitoring tools [279], in CMS runs corresponding to $\approx 400 \text{ pb}^{-1}$.

The timing measurements are extracted with a special client, outputting a histogram of the processing time, as shown in Figure 10.5.

The first peak indicates early rejections, where events were discarded at the beginning of the processing steps. The second visible peak indicates fast algorithms, which in this case, corresponds to the reconstruction of the muon and jets. The slow tail normally corresponds to more processing-heavy algorithms (e.g. tracking or PF), which is almost negligible in this case.

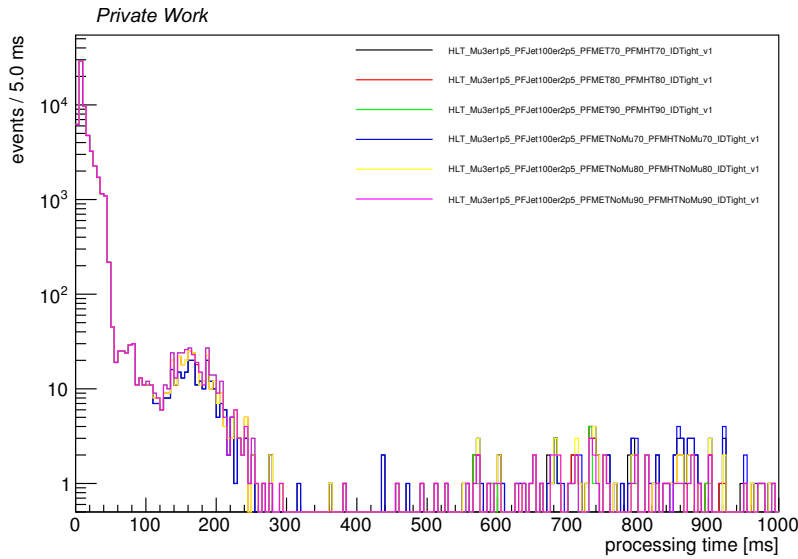


Figure 10.5: Timing measurement results of the *soft μ + hard jet + moderate p_T^{miss}* HLT paths, indicating the number of events versus processing time.

The average total timing of the main path is ≈ 4.6 ms, which falls within the expected budget.

10.4.3 Efficiency in Data

The efficiency measurement of the *soft μ + hard jet + moderate p_T^{miss}* path is performed directly using data collected by the trigger, corresponding to 59.7 fb^{-1} recorded by the CMS detector in 2018 [11]. The measurement is split between the three legs of the cross-trigger: muon, jet and p_T^{miss} , as a function of the offline muon p_T , jet p_T or p_T^{miss} quantities, respectively. The measurements are done in orthogonal control samples collected using reference triggers. Additional requirements are applied to ensure that the measurement is done in a parameter space in which the other two legs of the cross-trigger are well on the plateau of the efficiency turn-on curves of the other two legs, in order to minimise any potential biases in the measurement. Ultimately, the efficiency measurements performed in data (presented in the Figures 10.6 and 10.7 below), are compatible with the measurements performed in simulation shown in 10.3.4, which were used to optimise the trigger design.

Muon Leg

The muon leg efficiency measurement in Figure 10.6 (left) is performed in orthogonal control samples collected using reference triggers based on pure p_T^{miss} , requiring $p_T^{\text{miss}} > 120$ GeV and $H_T^{\text{miss}} > 120$ GeV. Additional requirements on the offline leading jet $p_T > 150$ GeV and $p_T^{\text{miss}} > 250$ GeV are applied to ensure that the measurement is done on the plateau of the efficiency turn-on curves.

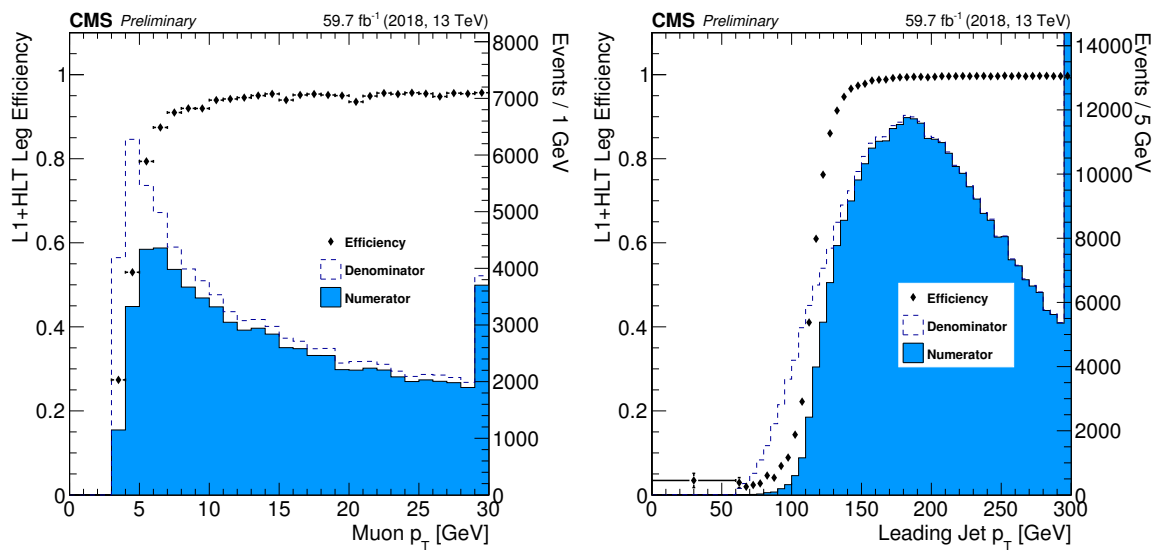


Figure 10.6: Efficiency (L1+HLT) of the muon (left) and jet (right) leg of a soft ($p_T > 3$ GeV) muon, high $p_T > 100$ GeV jet and p_T^{miss} cross-trigger measured in a sample based on pure p_T^{miss} .

Jet Leg

The jet leg efficiency measurement in Figure 10.6 (right) is performed in orthogonal control samples collected using reference triggers based on a single-muon condition, requiring an isolated high p_T muon with $p_T > 24$ GeV, as well as triggers based on pure p_T^{miss} , requiring $p_T^{\text{miss}} > 120$ GeV and $H_T^{\text{miss}} > 120$ GeV, in order to minimise any potential biases in the measurement. Additional requirements on the offline muon $p_T > 30$ GeV and $p_T^{\text{miss}} > 100$ GeV are applied to ensure that the measurement is done on the plateau of the turn-on curves of the other two legs and the reference trigger.

Missing Momentum Leg

The p_T^{miss} leg efficiency measurement in Figure 10.7 is performed in orthogonal control samples collected using reference triggers based on a single-muon condition, requiring an isolated high p_T muon with $p_T > 24$ GeV. Additional requirements on the offline leading jet $p_T > 150$ GeV and muon $p_T > 30$ GeV are applied to ensure that the measurement is done on the plateau of the efficiency turn-on curves of the other two legs. In order to obtain a cleaner sample, offline jets that are within $\Delta R < 0.4$ (where $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$) with respect to leptons are not considered. In order to make the measurement less susceptible to the differences between the p_T^{miss} definitions at L1 and HLT, an upper threshold of 40 GeV is applied to the muon p_T .

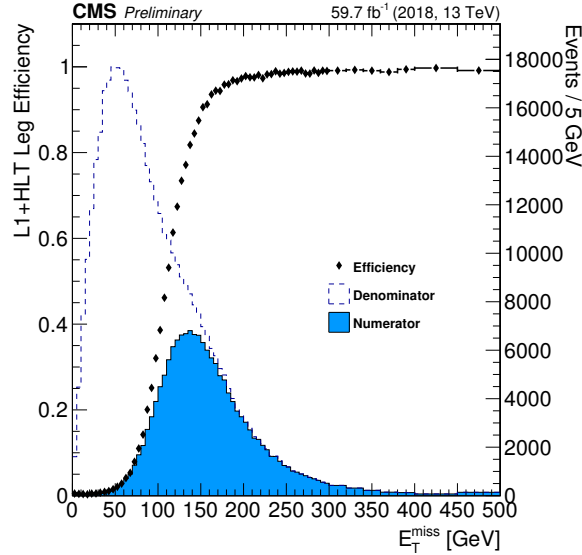


Figure 10.7: Efficiency (L1+HLT) of the p_T^{miss} leg of a soft ($p_T > 3$ GeV) muon, high $p_T > 100$ GeV jet and p_T^{miss} cross-trigger measured in a sample based on single-muon events.

10.5 Sensitivity

In order to incorporate the new *soft μ + hard jet + moderate p_T^{miss}* trigger into the analysis, one must introduce a strategy on how to take advantage of the additional phase-space that is opened up. It is crucial to discern the behaviour of the SUSY signal with respect to the dominant backgrounds in this region in the context of a sensitivity study. The considered signals are the two strongly-produced signals (T2tt, T2bW) and two EWKino signals (TChiWZ, Higgsino pMSSM), introduced in Section 9. The study is performed using simulated nanoAOD samples generated with 2018 conditions, corresponding to an integrated luminosity of $L = 59.8 \text{ fb}^{-1}$, with simplified reweighting and systematics. Given that the new trigger selects on muons, only the muon channel is considered, which is also the more sensitive channel of the analysis.

10.5.1 Low-MET Region

Based on the trigger efficiency curves in Section 10.4.3, a reasonable choice of the offline selection defined by the trigger plateaus is $p_T^{\text{miss}} > 200 \text{ GeV}$, muon $p_T > 6 \text{ GeV}$ with $|\eta| < 1.5$ and leading jet $p_T > 130 \text{ GeV}$ with $|\eta| < 2.4$. Therefore, the additional phase-space lies in the range $200 < p_T^{\text{miss}} < 300 \text{ GeV}$, with the upper limit defined by the $C_T > 300 \text{ GeV}$ selection of the signal regions, together with the additional requirements on the muon and leading jet. It is subsequently labelled as the *low-MET* region Z.

In order to better understand the signal and background kinematics of this extended region, a detailed signal significance study of the main observables and handles in the analysis is performed, based on a Figure-of-Merit (FoM) in the form of the signal-to-background ratio (s/\sqrt{b}):

- kinematic variables: $p_T^{\text{miss}}, H_T, p_T^{\text{ISR}}, C_T$
- leptonic properties: $p_T(\ell), \eta(\ell), \phi, W\text{-}p_T, m_T, Q(\ell)$
- lepton identification: IDs, isolation and impact parameters ($d_{xy}, d_z, \sigma_{d_{xy}}, \sigma_{d_z}$)
- jet multiplicities: $N_{\text{jets}} (N_{\text{jets}}^{\text{soft}}, N_{\text{jets}}^{\text{hard}}), N_{\text{b-jets}} (N_{\text{b-jets}}^{\text{soft}}, N_{\text{b-jets}}^{\text{hard}})$

- ISR topologies:

- recoil of objects against ISR, with ratios between p_T^{ISR} , H_T , p_T^{miss} , $p_T(\ell)$, $W-p_T$
- angular correlations ($\Delta\phi$, ΔR) between the ISR jet, p_T^{miss} and lepton

In terms of the kinematic selection, a pragmatic approach would be to lower the C_T requirement and add an extra general signal region defined by $200 < C_T < 300$ GeV, following the characteristic “L” shapes in the 2D plane, as shown in Figure 10.8.

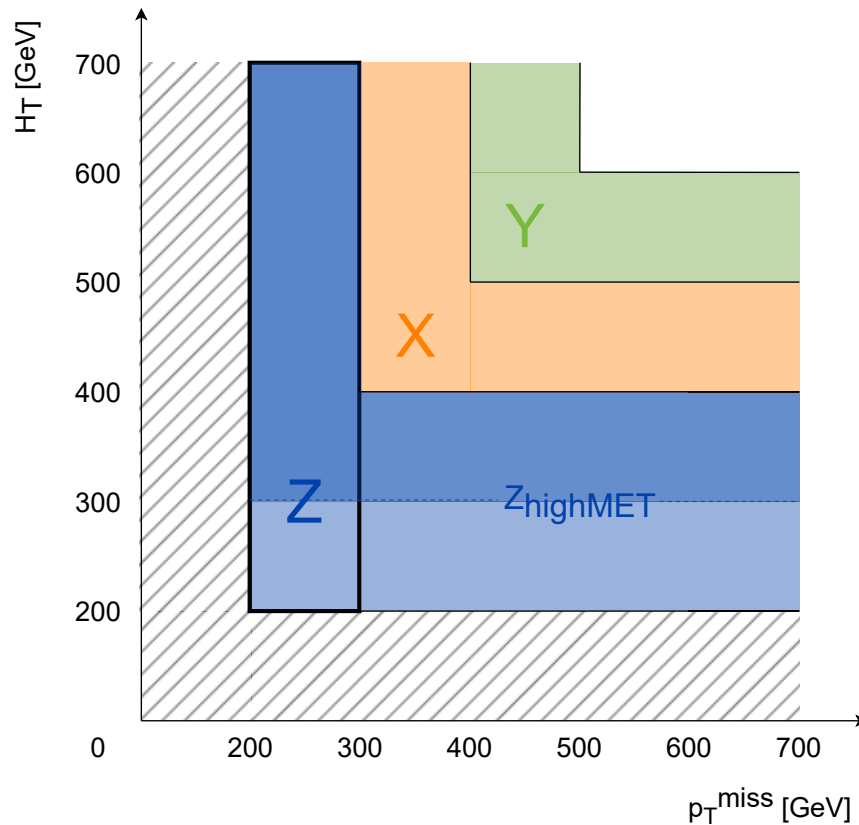


Figure 10.8: Schema of the signal region splitting in the p_T^{miss} vs H_T plane, including the *low-MET* region Z . Region Z_{highMET} indicates a potential extension of the region for higher p_T^{miss} values, as originally selected by the C_T variable.

To understand the viability of this simplified approach, a first step is to deconstruct and revisit the elements of the kinematic C_T variables (introduced in Section 6.4.4.3). Due to the back-to-back ISR topology, p_T^{miss} has a strong positive correlation with p_T^{ISR} , and,

consequently, H_T . Given that H_T is essentially a proxy for p_T^{ISR} , one can focus mainly on H_T to understand the correlation with p_T^{miss} , which is used for SR1. Furthermore, the b-jet requirement of SR2 is incompatible with the EWKino signals. Therefore, the study focuses on SR1, which is also the more sensitive region in the analysis.

As discussed in Section 6.4.4.3, the discriminating power of the C_T variables lies in the different correlation strengths between signal and main backgrounds. This can be seen in Figure 10.9, which presents 2D distributions of the deconstructed variables (p_T^{miss} and H_T) for the different signals plotted against the two main backgrounds, $W + jets$ and $t\bar{t}$, for a general SR1 selection. For the signals, the masses are chosen close to the sensitivity limits and medium Δm values. Going from lower to higher Δm , the shape of the signal spectrum shifts upwards in the plane towards higher H_T values due to the additional available energy.

To paint a more accurate picture, the 1D projections of the 2D distributions are shown in Figure 10.10, incorporating all the relevant backgrounds². They provide a measure of the signal significance as a FoM in the form of s/\sqrt{b} . When imposing higher requirements on H_T , there is a significant drop in the signal spectrum and, consequently, the significance.

The distributions show that the bulk of the signals concentrates both at lower p_T^{miss} and H_T values, as also demonstrated in Figure 10.1 in the context of the trigger design. Taking into account Figure 10.8, it is clear that a C_T requirement in its original definition would effectively discard a significant portion of the signals. Therefore, in the *low-MET* region, it is preferable to drop the simultaneous C_T selection in its present form. Additionally, the level of the H_T cut is aligned with that of $p_T^{\text{miss}} > 200$ GeV, while the p_T^{ISR} cut is kept at its minimum of 130 GeV, also in SR2. This adds additional acceptance in the parameter space of the hadronic energy observables. The motivation behind this modified approach is further justified by Figure 10.11, which shows the ratio distributions between p_T^{miss} ($W-p_T$), H_T and p_T^{ISR} , that peak near unity, as expected from ISR-like topologies.

²N-1 distributions are shown, removing the cut on the variable that is being plotted.

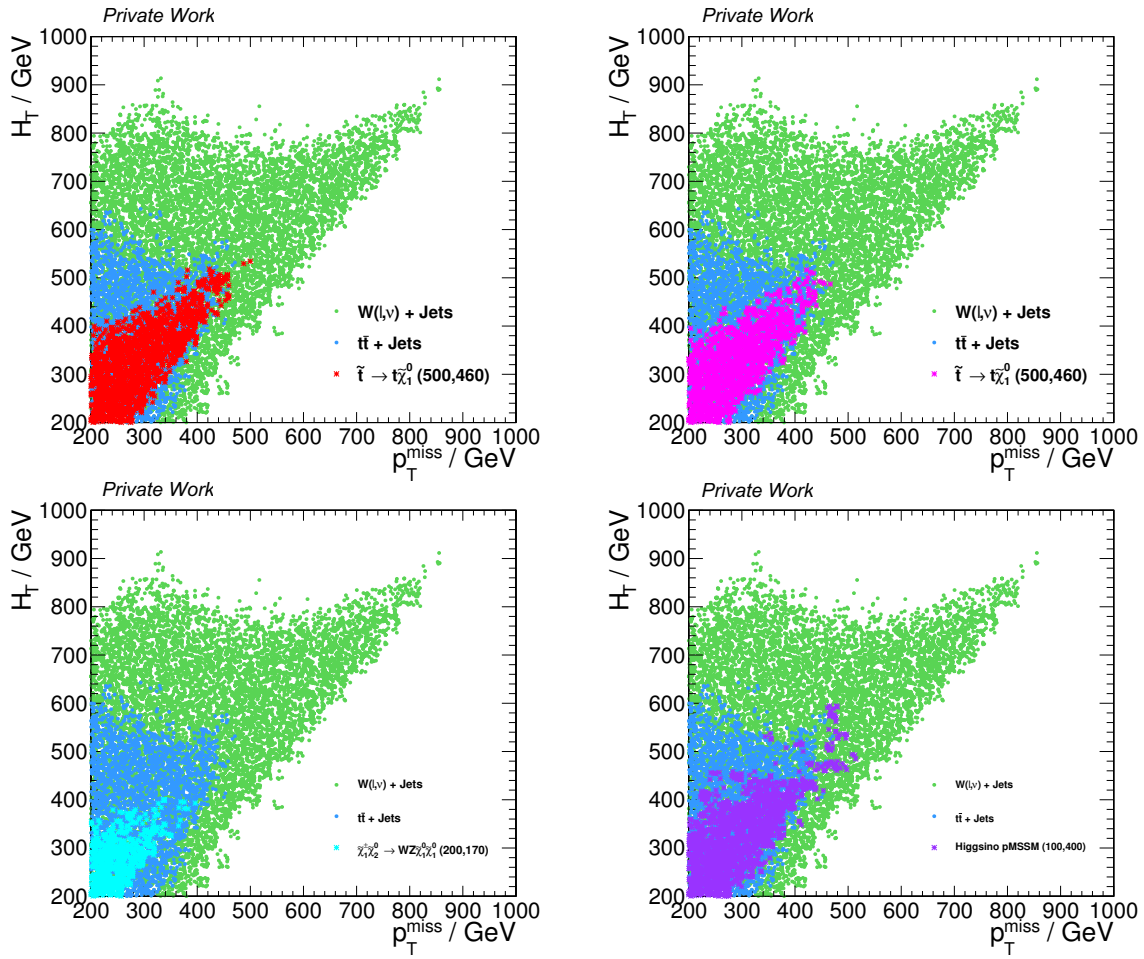


Figure 10.9: 2D distributions of p_T^{miss} versus H_T for the (a) T2tt: $\tilde{t}_1 \rightarrow b f f' \tilde{\chi}_1^0$, (b) T2bW: $\tilde{t}_1 \rightarrow b \tilde{\chi}_1^\pm \rightarrow b f f' \tilde{\chi}_1^0$, (c) TChiWZ: $\tilde{\chi}_1^\pm \tilde{\chi}_2^0 \rightarrow W^* Z^* \tilde{\chi}_1^0 \tilde{\chi}_1^0$ (d) Higgsino pMSSM, respectively, plotted against the two main backgrounds, $W + jets$ and $t\bar{t}$. A general SRI selection (inclusive in lepton $p_T < 30$ GeV, m_T and $Q(\ell)$) is applied with relaxed $H_T > 200$ GeV and $p_T^{\text{miss}} > 200$ GeV requirements. The signal masses are chosen close to the sensitivity limits and a medium Δm .

With the bulk of the dominant backgrounds also concentrating at lower values, one also loses the original discriminating power of the C_T variables due to the reduced differences in the correlation strengths. An alternative simultaneous selection is considered that takes advantage of observables directly related to the recoil of the system against ISR. While the angular distributions ($\Delta\phi$, ΔR) between the ISR jet, p_T^{miss} and muon show some potential in the depleted tails of the distributions, the momentum ratios in Figure 10.11 indicate that one should be able to achieve significant separation from backgrounds, especially in the region near unity, which is slightly off-diagonal in the 2D plane. The ratio shapes also show a dependence on the Δm , which is more

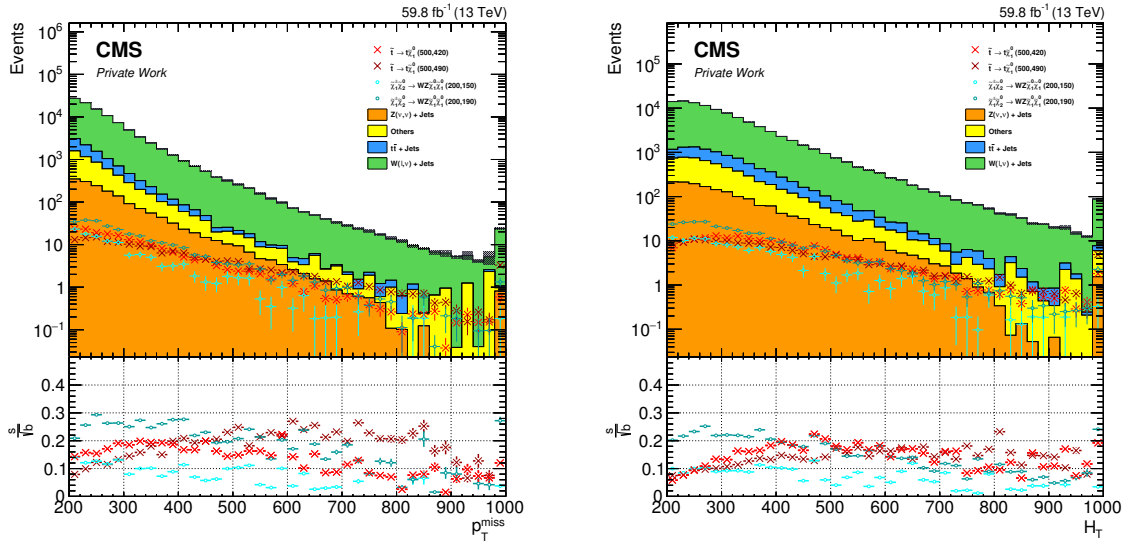


Figure 10.10: Signal and background distributions of p_T^{miss} (left) and H_T (right) with a general SRI selection (inclusive in lepton $p_T < 30$ GeV, m_T and $Q(\ell)$) with relaxed $H_T > 200$ GeV and $p_T^{\text{miss}} > 200$ GeV requirements. Distributions include benchmark signal points for strong (T2tt) and EWKino (TChiWZ) production, with masses chosen close to the sensitivity limits and high and low Δm values. Backgrounds are taken directly from simulation, with QCD events considered negligible and thus removed.

prominent for p_T^{miss}/H_T than in the $p_T^{\text{ISR}}/p_T^{\text{miss}}$ or $p_T^{\text{ISR}}/W-p_T$ observables. Therefore, a new discriminator is introduced in the form of the ratio between p_T^{miss} and H_T . Such a variable is also exploited by other analyses to suppress QCD backgrounds (e.g. [244]). In order to isolate these regions near unity and maintain sensitivity to a range of Δm values, a fine binning of

$$[0, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, \infty]$$

in p_T^{miss}/H_T is implemented.

To summarise, in terms of the kinematic selection, the *low-MET* region is defined by $200 < p_T^{\text{miss}} < 300$ GeV and $H_T > 200$ GeV (marked as Z in Figure 10.8), split into bins of p_T^{miss}/H_T . This *low-MET* region can additionally be extended by the low H_T and high p_T^{miss} region, as originally selected by the C_T variable (marked as Z_{highMET} in Figure 10.8), which is unused by the baseline analysis.

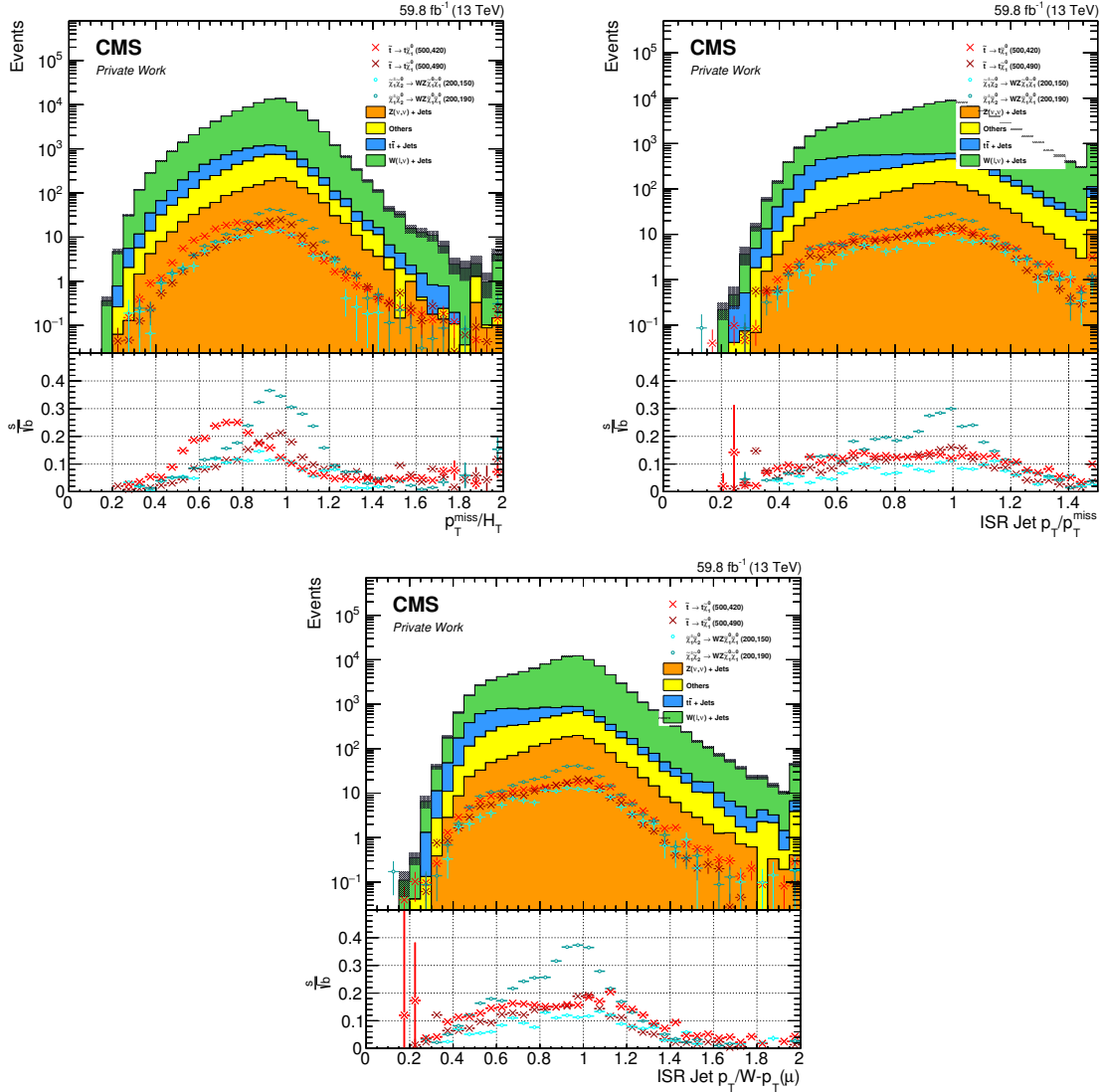


Figure 10.11: Signal and background distributions of the ratio of p_T^{miss}/H_T and the recoil of the leading ISR jet against p_T^{miss} ($p_T^{\text{ISR}}/p_T^{\text{miss}}$) and W - p_T (p_T^{ISR}/W - p_T) with a general SR1 selection (inclusive in lepton $p_T < 30$ GeV, m_T and $Q(\ell)$) with relaxed $H_T > 200$ GeV and $p_T^{\text{miss}} > 200$ GeV requirements. Distributions include benchmark signal points for strong (T2tt) and EWKino (TChiWZ) production, with masses chosen close to the sensitivity limits and high and low Δm values. Backgrounds are taken directly from simulation, with QCD events considered negligible and thus removed.

However, preliminary studies have shown it not to contribute significantly to the sensitivity, mainly due to the preference of the signal to higher H_T values with respect to the p_T^{miss} , as discussed above in Section 6.4.4.3 in the context of the C_T selection.

10.5.1.1 Background Suppression

An effective background suppression strategy must be implemented to extract the most sensitivity out of the *low-MET* region. In the extended low p_T^{miss} and low H_T regions, there is a significant increase in acceptance of both the signal and backgrounds due to the exponential shape of the distributions. Therefore, one can afford to employ a more aggressive approach and cut very tightly on discriminators in order to reduce the dominant backgrounds.

Due to lepton universality in electroweak interactions, a third of the leptonic W-boson decays are expected to proceed via an intermediate tau τ , which is reflected in a physical separation from the interaction point, expressed by the impact parameters (IPs) of the leptons. The distributions of the IPs (d_{xy} , d_z) together with their uncertainties ($\sigma_{d_{xy}}$, σ_{d_z}) are shown in Figure 10.12, with the ranges corresponding to the baseline analysis lepton definitions (Section 4.2).

One can see that the signal is prompt around the origin, while a sizeable portion of the backgrounds lies at higher absolute values. Therefore, a significant fraction of the dominant backgrounds can be removed by strongly tightening the IP, selecting events only at the peaks of the distributions. Furthermore, it can be seen that higher uncertainties are associated with backgrounds, and therefore maximum requirements on the IP error σ are applied as well. The exact values are chosen by a numerical maximisation of the signal significance:

- transverse IP $|d_{xy}| < 0.0015 \text{ cm}$ and $0.0012 < \sigma_{d_{xy}} < 0.0022$
- longitudinal IP $|d_z| < 0.003 \text{ cm}$ and $\sigma_{d_z} < 0.004$

In addition, the muon identification criteria have been tightened from the *loose ID* to the *tight ID* (Section 4.1) in order to help further reduce any residual nonprompt contributions.

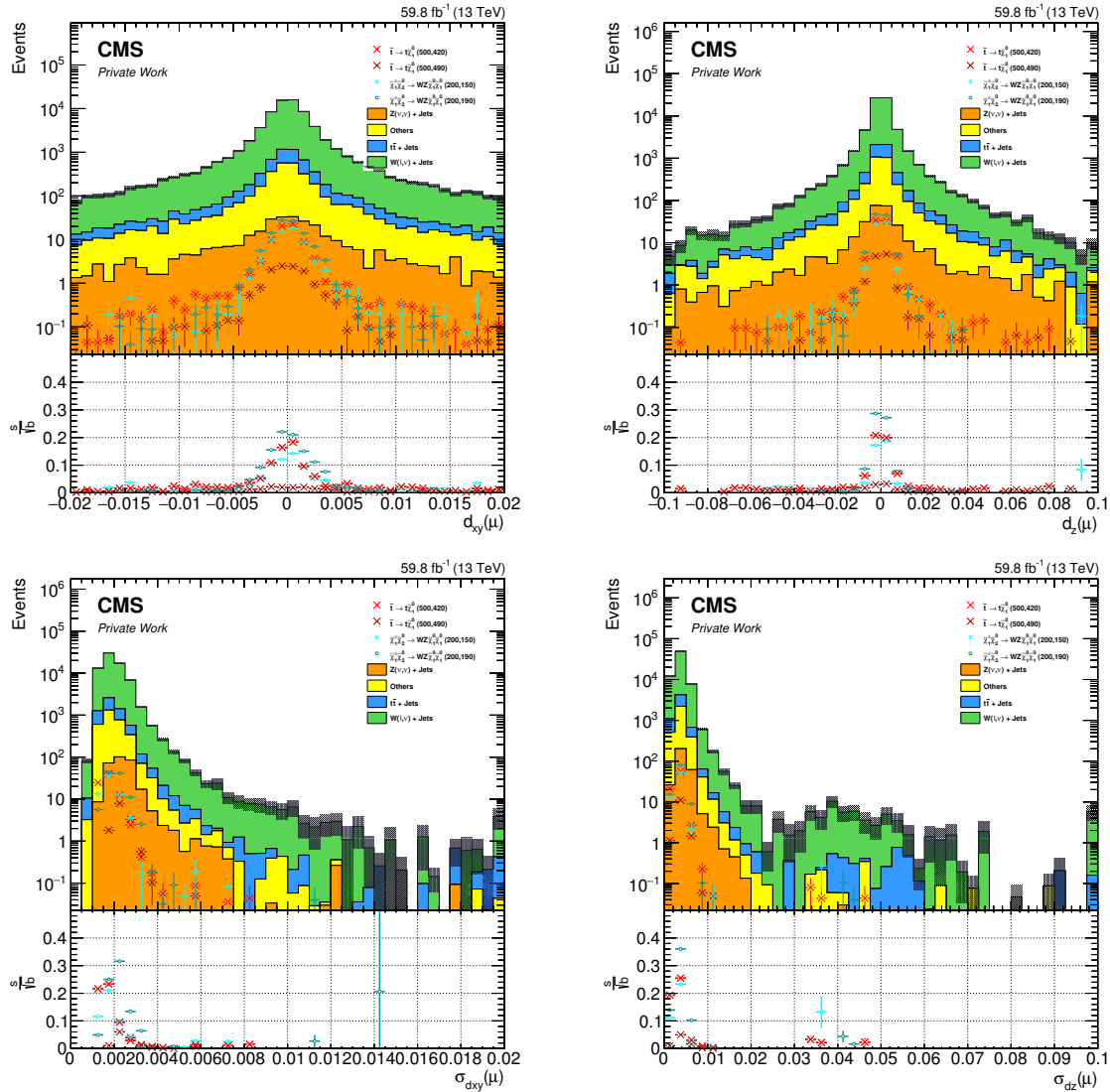


Figure 10.12: Signal and background distributions of the transverse and longitudinal lepton impact parameters (d_{xy} , d_z) and their corresponding uncertainties ($\sigma_{d_{xy}}$, σ_{d_z}) with a general SRI *low-MET* region selection with $200 < p_T^{\text{miss}} < 300$ GeV and $H_T > 200$ GeV (inclusive in lepton p_T (< 30 GeV), m_T and $Q(\ell)$). Distributions include benchmark signal points for strong (T2tt) and EWKino (TChiWZ) production, with masses chosen close to expected sensitivity limits and high and low Δm values. Backgrounds are taken directly from simulation, with QCD events considered negligible and thus removed.

10.5.1.2 Additional Splitting

The binning of the *low-MET* region is adopted from that of the baseline analysis signal regions. Revisiting the two main observables of the analysis (m_T , $p_T(\ell)$) in the context of the *low-MET* region, shown in Figure 10.13, indicate potential to improve the sensi-

tivity by splitting the regions further.

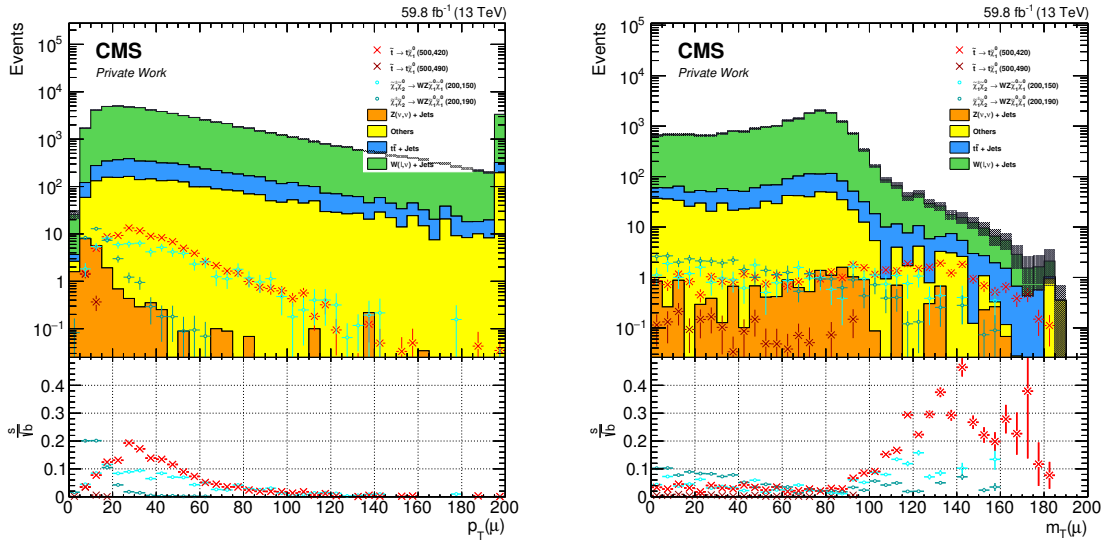


Figure 10.13: Signal and background distributions of the lepton p_T (left) and m_T (right) with a general SRI *low-MET* region selection with $200 < p_T^{\text{miss}} < 300$ GeV and $H_T > 200$ GeV (inclusive in lepton $p_T (< 30$ GeV), m_T and $Q(\ell)$). Leptons are required to have very tight IP and tight ID. Distributions include benchmark signal points for strong (T2tt) and EWKino (TChiWZ) production, with masses chosen close to expected sensitivity limits and high and low Δm values. Backgrounds are taken directly from simulation, with QCD events considered negligible and thus removed.

One could take further advantage of the m_T tail, shown in Figure 10.13, which is one of the main sources of sensitivity of the search, as discussed in Section 8. A feature of kinematic edges is that there is both depletion of backgrounds and signal, where one takes advantage of the difference in shapes, as discussed in Section 6.4.4.2. Similarly, as in the baseline analysis, there exists a trade-off between sensitivity coming from low lepton p_T and high m_T . In the particular case of the *low-MET* region, even though lower p_T^{miss} values favour lower m_T values, in special configurations with lower p_T^{miss} and higher $p_T(\ell)$, this can lead to a population of the very high m_T regions and provide extra sensitivity. To isolate this very high m_T region, an extra bin with $m_T > 130$ GeV is added, resulting in four m_T bins. The lepton p_T plot also indicates that for higher Δm values the signal acceptance can be increased by extending the general SR definition with a higher lepton p_T bin $30 \text{ GeV} < p_T(\ell) < 60 \text{ GeV}$ (VH). This is also motivated by the extra high m_T region. In turn, the control regions are re-defined by $p_T(\ell) > 60 \text{ GeV}$.

It is important to note that these considerations are also relevant for the higher p_T^{miss} regions of the baseline analysis as well and can help improve the sensitivity of the entire search.

10.5.1.3 Electroweakinos

In the baseline analysis, a selection on both charge $Q(\ell)$ and pseudorapidity $\eta(\ell)$ is implemented in SRI in order to suppress backgrounds with W-boson decays, as discussed in Section 6.4.4.4. The corresponding distributions in the *low-MET* region are shown in Figure 10.14.

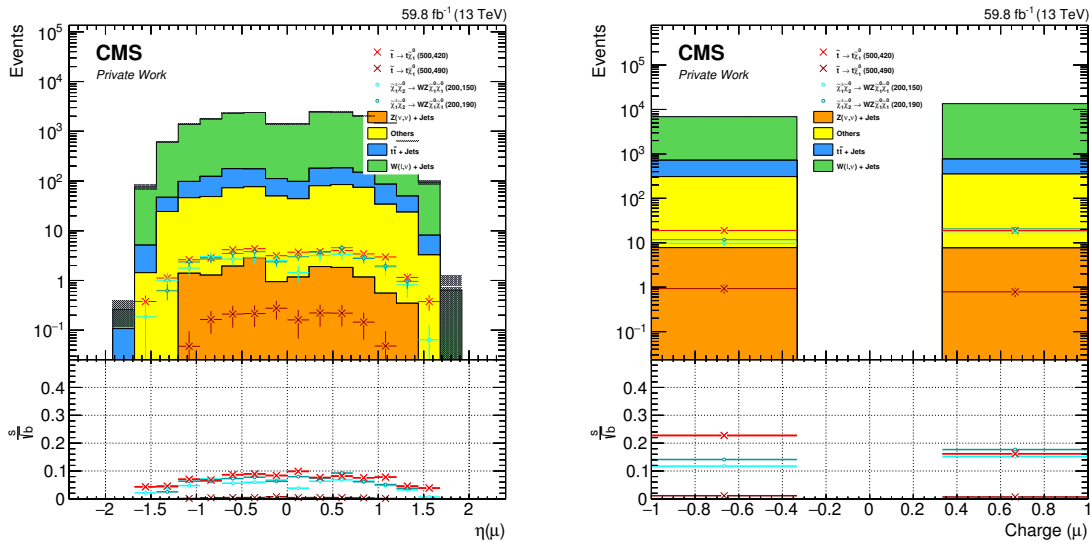


Figure 10.14: Signal and background distributions of the lepton pseudorapidity η (left) and charge $Q(\ell)$ (right) with a general SRI *low-MET* region selection with $200 < p_T^{\text{miss}} < 300$ GeV and $H_T > 200$ GeV (inclusive in lepton p_T (< 30 GeV), m_T and $Q(\ell)$). Leptons are required to have very tight IP and tight ID. Distributions include benchmark signal points for strong (T2tt) and EWKino (TChWZ) production, with masses chosen close to expected sensitivity limits and high and low Δm values. Backgrounds are taken directly from simulation, with QCD events considered negligible and thus removed.

With the additional acceptance of the *low-MET* region, in general, further tightening or splitting of the $\eta(\ell)$ could provide some gains in sensitivity in the context of the strongly-produced signals. However, in the context of EWKinos, due to the same electroweak nature of the production mechanisms as the dominant backgrounds (Section 9), additional selections are not considered. Therefore, for the *low-MET* region, there

is no further tightening of $\eta(\ell)$ and the charge cut is dropped in all low $m_T < 95$ GeV regions.

10.5.1.4 Summary

To summarise, the *low-MET* region Z is generally defined by kinematic cuts on $200 < p_T^{\text{miss}} < 300$ GeV and $H_T > 200$ GeV, marked as Z in Figure 10.8. It could additionally be extended by the low H_T and high p_T^{miss} region Z_{highMET} , unused by the baseline analysis. Only muons with $p_T > 6$ GeV with $|\eta| < 1.5$ are considered, and the minimum requirement on the p_T^{ISR} is increased to 130 GeV, in compliance with the trigger conditions.

The signal region binning of the *low-MET* region follows that of the baseline analysis with the following modifications:

- 8 bins in p_T^{miss}/H_T (replacing the C_T binning): $[0, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, \infty]$
- very tight lepton IP:
 - $|d_{xy}| < 0.0015$ cm and $0.0012 < \sigma_{d_{xy}} < 0.0022$
 - $|d_z| < 0.003$ cm and $\sigma_{d_z} < 0.004$
- tight lepton ID
- additional high m_T bin = 4 m_T bins: $[0, 60, 95, 130, \infty]$
- additional high lepton p_T bin = 4 lepton p_T bins: $[6, 12, 20, 30, 60]$
- charge-inclusive in all regions for EWKinos: $Q(\ell) = \pm$.

Ultimately, the *low-MET* region with the additional binning results in an extra 256 SRs (and correspondingly 64 CRs) on top of the baseline analysis regions, labelled as

$$\text{SR}[1, 2] [\text{L}, \text{M}, \text{H}, \text{VH}] [\text{a}, \text{b}, \text{c}, \text{d}] \text{Z}$$

indicating the extra m_T and $p_T(\ell)$ bins. The labels are additionally appended by a suffix indicating the upper bound of the p_T^{miss}/H_T bins (e.g. 0p9 for $0.8 < p_T^{\text{miss}}/H_T < 0.9$).

The background composition of the SR1 bins against a benchmark T2t τ signal point is shown in Figure 10.15, while the composition of SR2 and the CRs are shown in Appendix C.3.

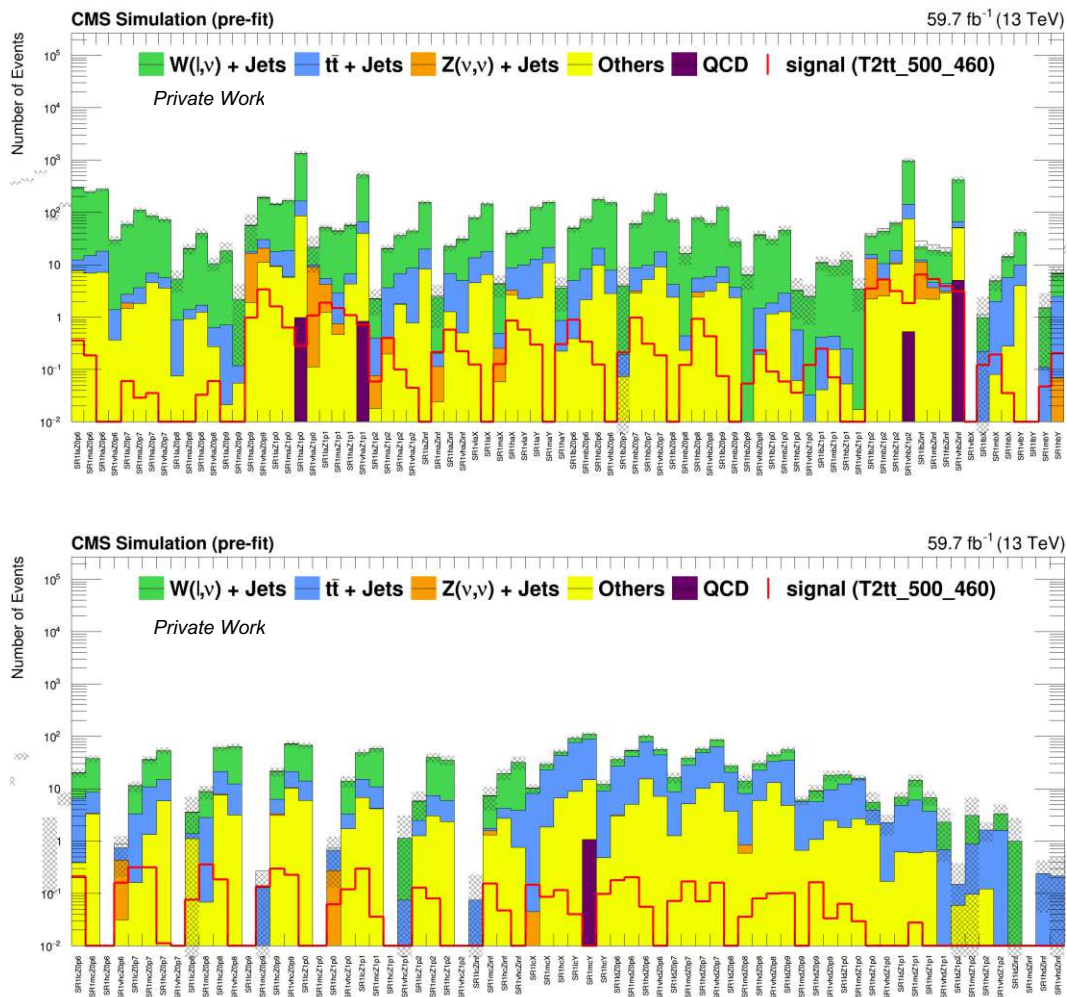


Figure 10.15: Background composition of all SR1 bins in the baseline and *low-MET* SRs, against a benchmark T2t τ ($\tilde{t}_1 \rightarrow b f f' \tilde{\chi}_1^0$) signal point with a mass chosen close to the sensitivity limits and a medium $\Delta m = 40$ GeV value.

10.5.2 Expected Limits

In order to quantify the potential relative gain of the *low-MET* region to the sensitivity of the analysis, the expected limits (background-only hypothesis) are calculated using MC, incorporating the new set of SRs. The same procedure based on the CL_S criterion is implemented as in Section 8, with simplified reweighting and systematics. The expected upper limits are calculated at 95 % CL interval using simulated samples generated with 2018 conditions, corresponding to an integrated luminosity of $L = 59.8 \text{ fb}^{-1}$. Expected results are also calculated with a scaling of the simulated signal and background yields to $L = 260 \text{ fb}^{-1}$, with the relative statistical and systematic uncertainties assumed constant. This corresponds to the expected dataset collected with the new trigger, amounting to the cumulative luminosity between the 2018 dataset and the expected LHC *Run 3* luminosity of $L \approx 200 \text{ fb}^{-1}$.

The relative difference in the expected upper limits between the baseline analysis and the aggregated *low-MET* region are shown in Appendix C.4 for the four different strong and EWKino models. The expected improvements reach up to $\approx 50 \%$, depending on the model and its parameters, with the largest improvements for the TChiWZ model. Generally, there is a larger increase in sensitivity for the higher Δm regions, understood to be mainly driven by the high m_T regions. The increases in sensitivity are reflected in the expected exclusion contours, shown in the 2D \tilde{t}_1 versus $\tilde{\chi}_1^0$ plane in Figure 10.16 for the strongly-produced signals (T2tt and T2bW).

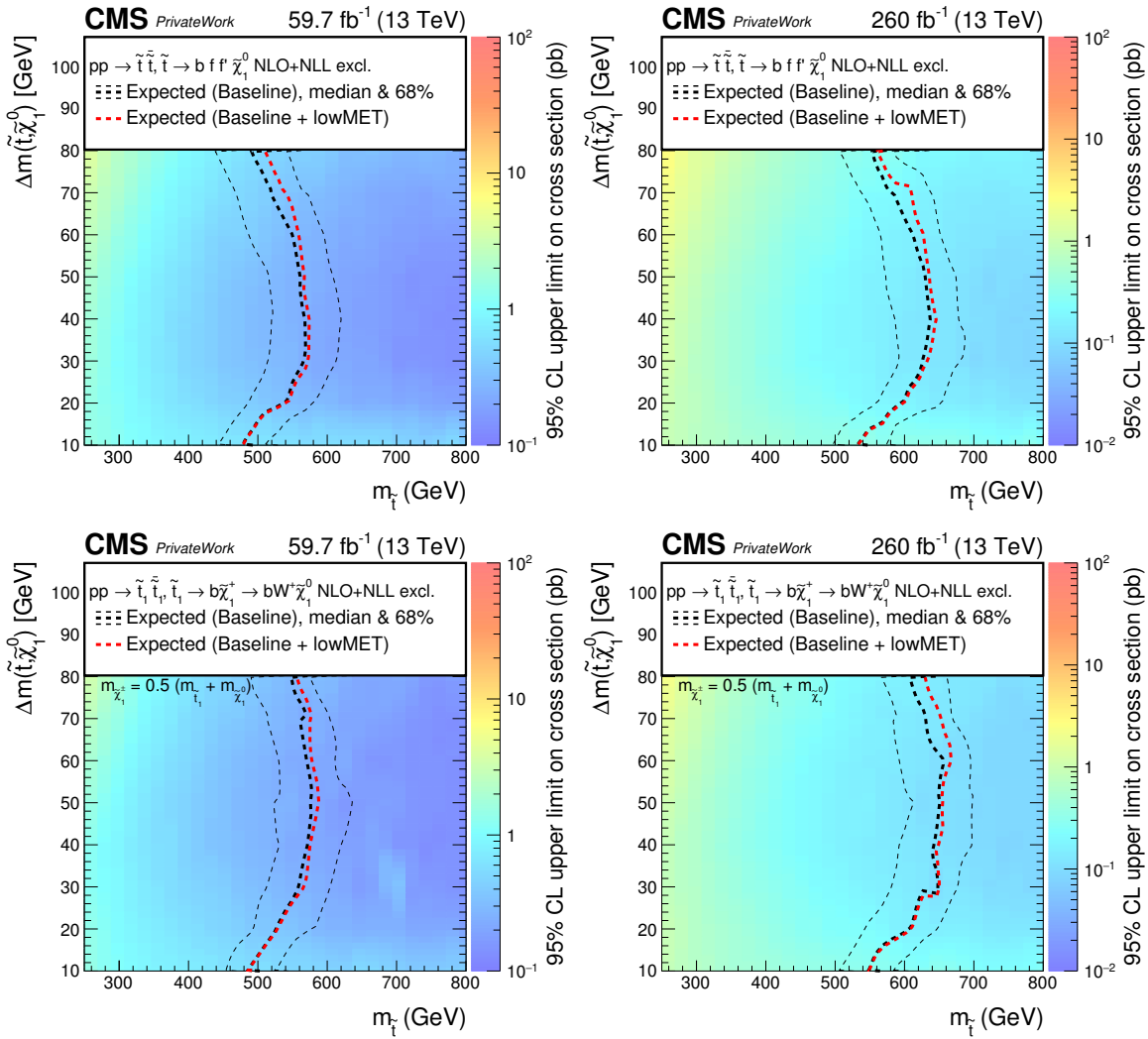


Figure 10.16: Comparison between the expected upper limits on cross section and exclusion contours for the baseline analysis (black) and the aggregated *low-MET* region (red) for the strongly produced signals ($T2tt: \tilde{t}_1 \rightarrow b f f' \tilde{\chi}_1^0$ and $T2bW: \tilde{t}_1 \rightarrow b \tilde{\chi}_1^\pm \rightarrow b f f' \tilde{\chi}_1^0$) in the 2D plane of \tilde{t}_1 masses vs. $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0)$, in the muon channel. The plots on the left show the expected limits for the 2018 luminosity $L = 59.8 \text{ fb}^{-1}$, whereas those on the right show the expected limits with yields scaled to $L = 260 \text{ fb}^{-1}$, corresponding to the cumulative expected dataset with LHC Run 3 data.

Earlier preliminary studies have shown that the baseline analysis has very limited sensitivity to EWKino models [252]. Therefore, in order to underline the relative gain of the *low-MET* region, the expected upper limits are shown for several mass configurations at the limit of sensitivity. The limits on the cross section are normalised to the SM background-only cross section (σ_{theory}), which can be interpreted as the signal strength modifier r (see Section 8). Therefore, the lines correspond to the exclusion contours where values below $r = 1$ indicate exclusion. For the TChiWZ model, the limits for the lowest mass points $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = [100, 125, 150]$ GeV are shown in Figure 10.17, as a function of $\Delta m(\tilde{\chi}_1^\pm, \tilde{\chi}_1^0)$. The plots also show a comparison between charge-inclusive regions, where the charge cuts are dropped both in the baseline and *low-MET* signal regions. For the TChiWZ model, it can be seen that the addition of the *low-MET* region improves the sensitivity for most Δm values, with the largest gain at higher Δm , as discussed above. Furthermore, the charge-inclusive selection indicates improvement in the lowest Δm values, even at nominal luminosity. With an increased integrated luminosity $L = 260 \text{ fb}^{-1}$, the combination of the *low-MET* region with the charge-inclusive selection could lead to exclusions both at very low Δm as well as higher Δm values.

Similarly, for the Higgsino pMSSM model, Figure 10.18 shows the limits for $\mu = [100, 120, 140]$ GeV as a function of M_1 . For the Higgsino pMSSM model, the expected gains in sensitivity from the *low-MET* region, increased luminosity, and charge-inclusivity are less significant, with some potential at low M_1 values, corresponding to higher Δm . The extrapolation to higher luminosity indicates that in the current analysis configuration, one could expect minimal exclusion at the lowest μ values only.

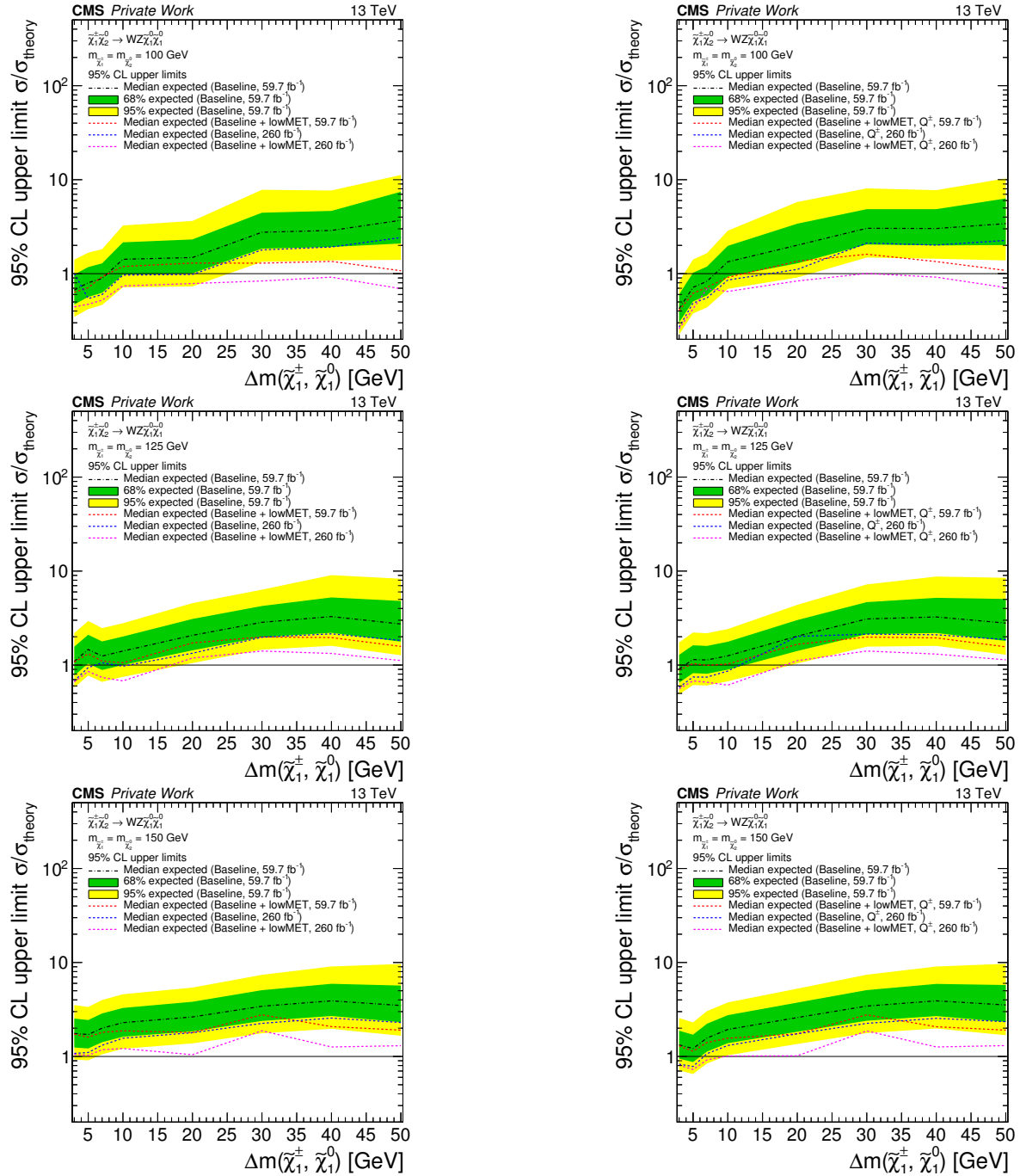


Figure 10.17: Comparison between the expected upper limits on cross section (normalised to the SM background-only cross section σ_{theory}) for the baseline analysis (black, blue) and the aggregated *low-MET* region (red, magenta) for the TChiWZ EWKino signal ($\tilde{\chi}_1^\pm \tilde{\chi}_2^0 \rightarrow W^* Z^* \tilde{\chi}_1^0 \tilde{\chi}_1^0$) for fixed mass points $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = [100, 125, 150]$ GeV as a function of $\Delta m(\tilde{\chi}_1^\pm, \tilde{\chi}_1^0)$, in the muon channel. The contours show expected upper limits both for the 2018 luminosity $L = 59.8 \text{ fb}^{-1}$ (black, red) and $L = 260 \text{ fb}^{-1}$ (blue, magenta), corresponding to the cumulative expected dataset with LHC Run 3 data. The plots on the right correspond to charge-inclusive (Q^\pm) signal regions.

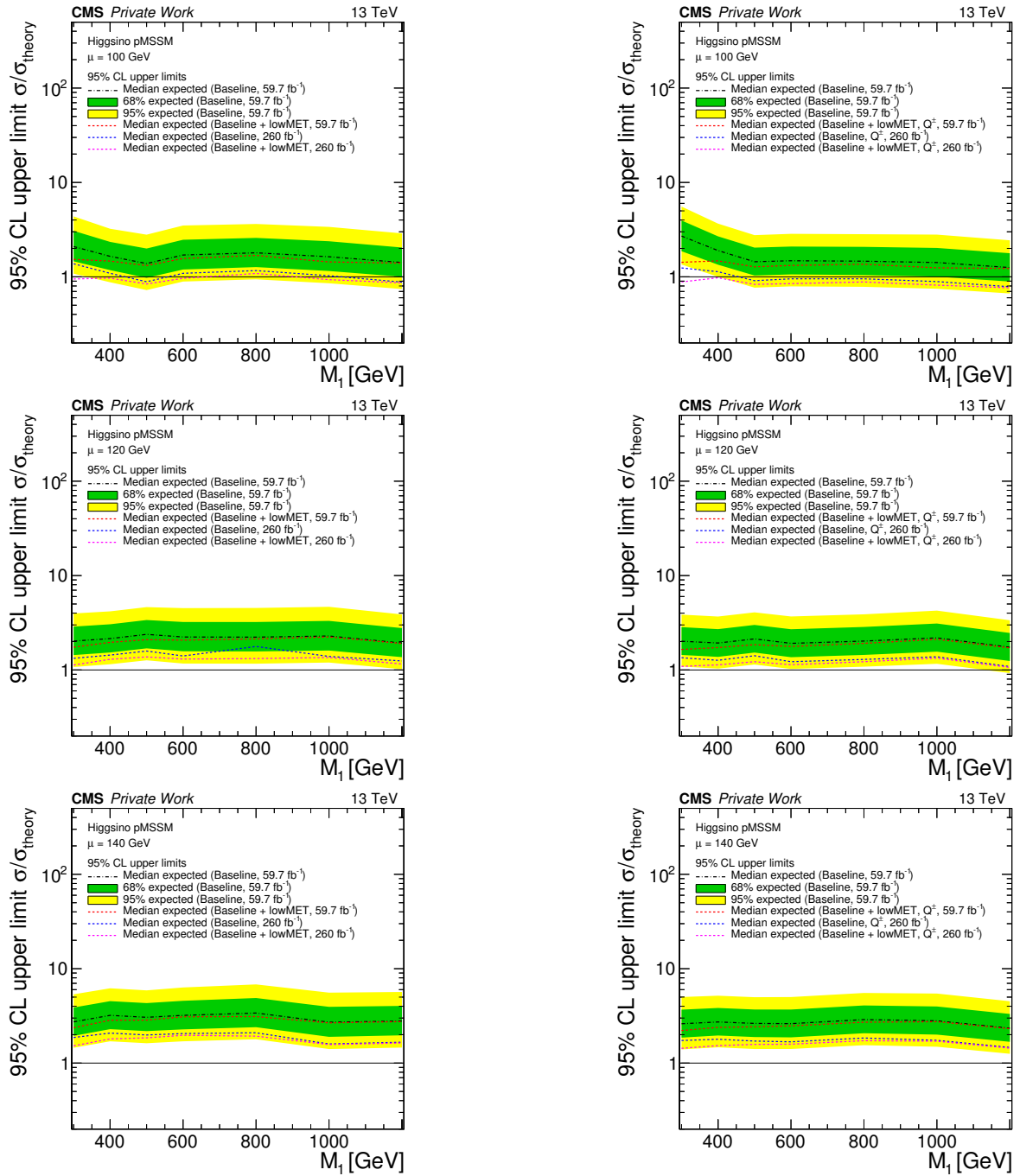


Figure 10.18: Comparison between the expected upper limits on cross section (normalised to the SM background-only cross section σ_{theory}) for the baseline analysis (black, blue) and the aggregated *low-MET* region (red, magenta) for the Higgsino pMSSM EWKino signals for fixed parameters $\mu = [100, 125, 150]$ as a function of M_1 , in the muon channel. The contours show expected upper limits both for the 2018 luminosity $L = 59.8 \text{ fb}^{-1}$ (black, red) and $L = 260 \text{ fb}^{-1}$ (blue, magenta), corresponding to the cumulative expected dataset with LHC Run 3 data. The plots on the right correspond to charge-inclusive (Q^\pm) signal regions.

10.5.3 Soft Multileptons

As discussed in Section 5, the soft multilepton analysis [244, 245] is a complementary compressed SUSY search that considers two or three leptons, with at least one oppositely charged lepton pair. The signal regions are mainly categorised in terms of p_T^{miss} and the invariant mass of the dilepton system $m_{\ell\ell}$. A dedicated dimuon, p_T^{miss} and H_T^{miss} cross-trigger

(HLT_DoubleMu3_DZ/DCA_PFMET50_PFMHT60) has been developed in order to expand the phase-space to a *low-MET* $p_T^{\text{miss}} < 200$ GeV region, increasing the acceptance and sensitivity of the analysis for EWKino models.

The *soft μ + hard jet + moderate p_T^{miss}* trigger was examined in the context of improving the sensitivity of the analysis by recording events with very soft sub-leading muons that are not reconstructed by the dimuon trigger [280, 281]. In the context of LHC *Run 3*, there is the consideration of extending the selection to lower invariant masses ($m_{\ell\ell} < 4$ GeV) and lower sub-leading muon momenta ($p_T < 5$ GeV) in the *low-MET* ($p_T^{\text{miss}} < 200$ GeV) region. A preliminary study indicates a significant increase in the acceptance of EWKino signals, which could potentially improve the sensitivity to low Δm points in the very low mass ($1 < m_{\ell\ell} < 4$ GeV) signal regions. Figure 10.19 shows the expected TChiWZ signal significance for the $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = \{150, 175, 200\}$ GeV and $\Delta m(\tilde{\chi}_1^\pm, \tilde{\chi}_1^0) = 8$ GeV, as a function of the sub-leading muon $p_T(\ell_2)$. There is the indication of possible gains in sensitivity without modifying the baseline selection, as shown in the final bin corresponding to $p_T(\ell_2) > 5$ GeV. There is further potential in exploiting the trigger if the sub-leading lepton p_T cut is relaxed down to 3 GeV. This addition is considered in the context of the *Run 3* analysis, pending a full sensitivity study with the relevant systematics, as well as optimisation of the final selection for background suppression.

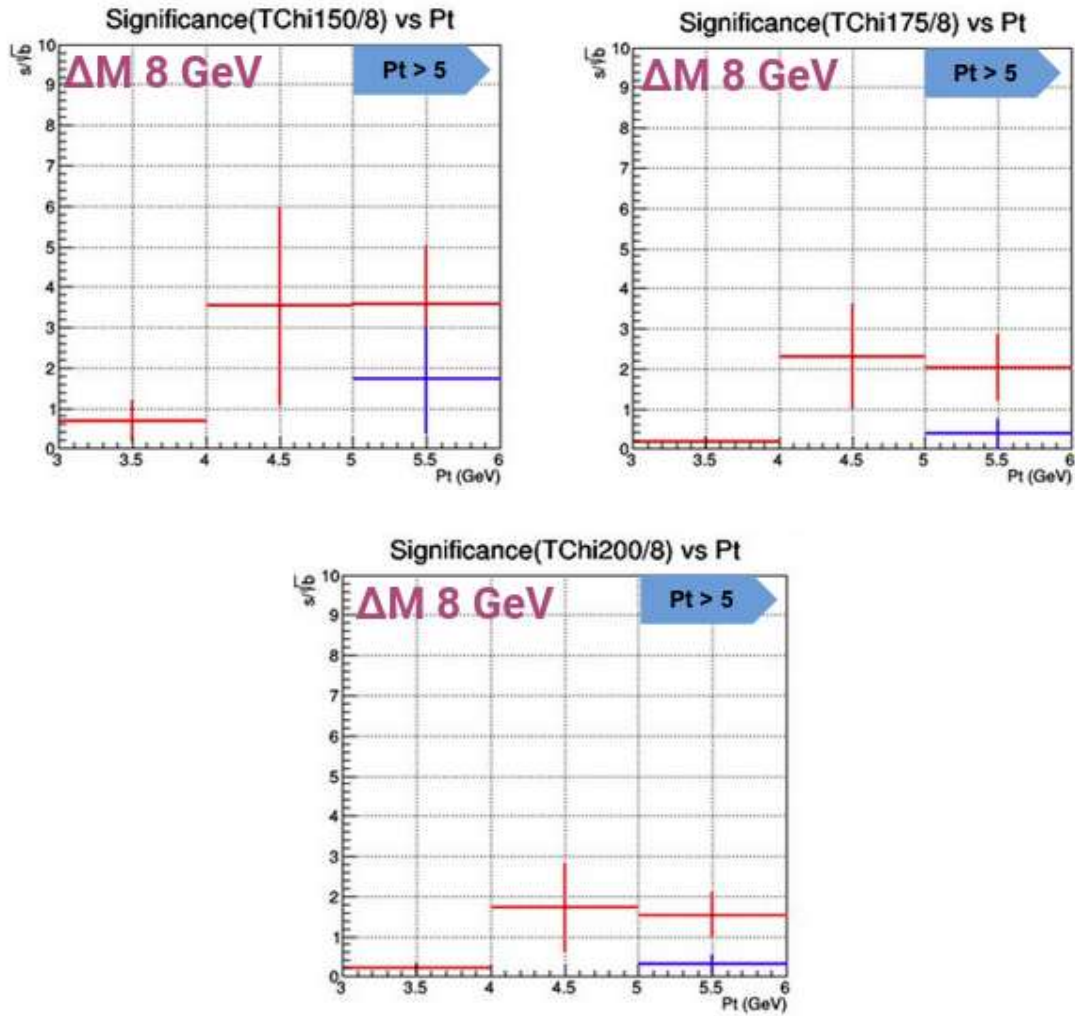


Figure 10.19: Expected TChiWZ signal significance for the $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = \{150, 175, 200\}$ GeV and $\Delta m(\tilde{\chi}_1^\pm, \tilde{\chi}_1^0) = 8$ GeV, as a function of the sub-leading muon $p_T(\ell_2)$, in the *low-MET* ($p_T^{\text{miss}} < 200$ GeV) and low invariant mass ($1 < m_{\ell\ell} < 4$ GeV) signal region. The blue lines correspond to a soft dimuon trigger selection and the red lines indicate a logical 'OR' together with the soft single muon trigger. The final bin $p_T(\ell_2) > 5$ GeV corresponds to the baseline signal region selection.

Outlook

Soft Trigger and Electroweakinos

One of the primary motivations for the development of the *soft μ + hard jet + moderate p_T^{miss}* trigger was opening the window for the less accessible compressed EWKino models, which is achieved in the form of a significant increase in signal acceptance. The largest expected gains in sensitivity are for the TChiWZ model, which was deemed the most promising model based on earlier preliminary sensitivity studies [252].

Considering the luminosity projections, it is reaffirmed that solely an increase in integrated luminosity is not expected to bring large gains in sensitivity. It is apparent that in order to reach sensitivity levels of the soft multilepton search [245, 244] (for 137 fb^{-1}), the *soft single-lepton analysis* would need to be re-optimised towards EWKino signals as a whole.

There are naturally certain caveats in performing a direct comparison. Firstly, the electron channel needs to be considered in the baseline analysis regions as well, which would improve the overall limits. On the other hand, the sensitivity study was performed with simplified reweighting and systematics, so one would expect the actual limits to be slightly weaker. Furthermore, there are differences in the signals due to additional reweighting and modelling improvements implemented in [245], leading to different dilepton invariant mass spectra and, ultimately, upper limits on the signals.

It is important to note that multilepton channels have a cleaner signature, with additional rejection of single-lepton backgrounds, especially $W + jets$. Furthermore, there are additional discriminating observables, such as correlations between two leptons, that can be taken advantage of. However, this comes at the cost of lower acceptance, which is higher in the single-lepton channel. A trade-off between signal acceptance and background contamination exists between the two channels. With a larger signal acceptance and, consequently, sample of events, there is additional potential to extract the signal with more sophisticated methods.

Given that in building the *low-MET* region, most of the main observables and handles in the soft single-lepton analysis were considered, the signal significance study underlines the limitations of a *cut-&-count* approach in separating the signal from the dominant backgrounds in a more difficult region, especially for very low Δm . This is also apparent for the strongly-produced signals, where the potential gains in sensitivity are rather limited, which can additionally be attributed to the analysis already being well-optimised for stop signals.

Nevertheless, with the extended parameter space, additional gains in sensitivity can be achieved by further optimisation in the separation of the signal against the backgrounds. This could be achieved with sophisticated analysis methods. An example is Machine Learning (ML), which potentially could pinpoint correlations or corners of phase-space that are missed in a standard *cut-&-count* analysis. Such methods are briefly discussed below, also in the context of the full analysis.

A healthy perspective is that the single-lepton and multilepton searches complement each other by targeting different decay channels, in an attempt to cover as much of the parameter space as possible. Therefore, one could foresee a combination of the single and multilepton search in the context of e.g. a multivariate (MVA) or Deep Neural Network (DNN) analysis. The synergies between the two are further underlined with the potential of using the *soft $\mu + hard jet + moderate p_T^{\text{miss}}$* as a complementary trigger strategy.

Several other factors need to be considered in the context of a full analysis that are not reflected in the performed study, such as fine-tuning the final offline selection with respect to the trigger systematics. This is evaluated by the measurements of the efficiency curves and is ultimately dependent on the way the trigger selection is applied (or parametrised) in data and simulation. Such considerations also open up the potential of using events on the efficiency turn-on in order to expand further into even lower p_T^{miss} regions. Therefore, one could consider the plateau selection in the signal sensitivity study as a rather conservative one.

A complementary approach to improvements in the analysis methods are potential further improvements to the trigger strategy itself. These are equally important in the context of mitigating the increase in thresholds due to the higher luminosity and harsher pileup conditions at the LHC. LHC *Run 3* is currently collecting data at an instantaneous luminosity $\mathcal{L} \approx 2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ and pileup $\text{PU} \approx 60$. Furthermore, test LHC fills have shown that there is potential to go up to $\mathcal{L} = 2.4 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ and $\text{PU} \approx 70$, which is being evaluated by the experiments. In such a scenario, one would need to really push the limits of the trigger systems. There is strong encouragement and support from the experiment to develop new trigger algorithms that not only keep rates under the limits but additionally open phase-space for new signals, as achieved by this trigger.

One can always consider imposing additional requirements in order to reduce rates and thresholds further. One could take advantage of the angular correlations in compressed SUSY signatures between the ISR-jet, p_T^{miss} and leptons. Other examples include requirements on lepton isolation at the trigger level; this is more challenging to achieve at the LIT level, where internal studies have shown that for high p_T muons, the rate reduction was small at the cost of efficiency in the plateau. However, one must also keep in mind that any additional requirements come at the cost of inclusivity and add additional layers of complexity to the final analysis treatment. Here, once again, one would need to make the choice between inclusivity and new potential users versus exclusivity and lower rates and thresholds.

Another possibility of the trigger systems that is becoming more widely used is data *parking* for delayed reconstruction, opening up resources for higher rates and consequently lower thresholds; this was used in 2018 in the context of low p_T ($6 < p_T < 18$ GeV) single-muon triggers, targeting B-flavour anomalies (R_K, R_{K^*}). Data *scouting* with reduced event contents and a simplified event description also allows for the reduction of thresholds, and there are significant efforts in its improvements, including LIT scouting at the full bunch crossing rate (40 MHz) [282], also with DNN [283]. A major upgrade for LHC *Run 3* is heterogenous processing with GPUs on HLT, offloading a significant amount of the bandwidth, especially for processing-heavy reconstruction algorithms (e.g. local tracking, ECAL and HCAL reconstruction) and making space for a high rate (≈ 30 kHz) scouting stream [181], and possibly ML for PF [284]. Furthermore, the viability of a low p_T electron trigger could be re-evaluated with the most recent improvements in electron reconstruction and trigger algorithms, discussed below. Finally, one could consider dedicated trigger algorithms in the search for a displaced stop, as discussed below. Ultimately, synergies with trigger strategies of other groups that have similar goals should be sought for.

Machine Learning

Advanced machine learning (ML) techniques are being developed and implemented in a wide variety of applications, which provide superior performance compared to standard approaches. ML methods can be exploited to construct discriminating observables in an analysis, which are then used to define the final signal region selection. They have the potential to better exploit correlations of observables that might be hidden in the standard optimisation of a sequential *cut & count* analysis. This can be especially advantageous for complex topologies, where there are subtle differences in the spectra of kinematic observables, such as compressed SUSY signatures.

As described in Section 6.1, parallel to the *cut & count* soft single-lepton analysis [4, 5], the data is also analysed using an MVA technique [249] based on boosted decision trees

(BDTs) [250]. Both methods rely on the same baseline event selection with only a few minor differences that are used to maximise the sensitivity. The set of input variables to the BDT consists of kinematic variables (p_T^{miss} and H_T), leptonic variables ($p_T(\ell)$, $\eta(\ell)$, $Q(\ell)$, m_T), jet variables (p_T^{ISR} , $p_T(\mathbf{b})$, N_{jets}), as well as quantities that take advantage of b-tagging algorithms ($N(\mathbf{b}^{\text{loose}})$, $\Delta R(\ell, \mathbf{b})$ and b-tagging discriminant $D(\mathbf{b})$). The BDT is trained using simulated samples of the T2tt signal model with four-body decays and $W+\text{jets}$, $t\bar{t}$ and Z_{inv} events for the background. Due to the large dependence of the event kinematics on the Δm , each training is done as a function of the Δm , averaging over the stop masses. The resulting BDT output response for two extreme Δm points in both data and simulation can be seen in Figure 11.1, where the differences in the distributions reflect the dependency of the signal event kinematics on the Δm and BDT training optimisation. The signal-to-background separation is achieved at high BDT values and the SRs are defined by requiring a minimum value on the BDT discriminant for each Δm point. The upper limits of the MVA approach are shown in Figure A.3; the limits between the MVA and *cut & count* (Figure 8.4) approaches are similar and compatible with one another, especially for low and medium Δm . The main improvement in sensitivity is for the high Δm regions, which can partially be attributed to the open lepton p_T cut in the training of the BDTs, additionally selecting leptons with $p_T > 30$ GeV, which are not considered by the baseline *cut-&-count* signal region selection. It is recently complemented by a full LHC Run 2 legacy result [251] with 138 fb^{-1} .

The application of an ML discriminator in the context of the *low-MET* region targeted by the *soft μ + hard jet + moderate p_T^{miss}* trigger could provide superior signal-to-background separation to the suggested strategy in the sensitivity study in this challenging region. One must consider that for each of the EWKino signals, separate training and optimisation of input variables are required, also with a mass-splitting granularity.

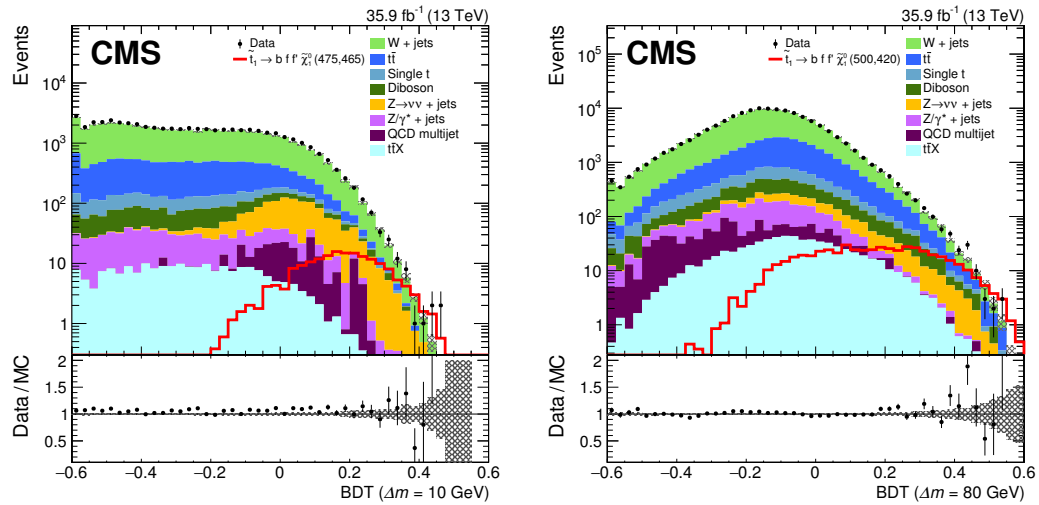


Figure II.1: Distributions of the BDT output at the preselection level in data and simulation for the T2 $t\bar{t}$ signal with Δm of 10 GeV (left) and 80 GeV (right) [227].

In the context of object reconstruction and identification, ML has been shown to provide improved efficiency and purity with respect to more standard methods, such as in the case of MVA and DNN b-tagging algorithms covered in Section 4.1 (e.g. DeepCSV, DeepJet, ParticleNet), which are also exploited for triggering purposes, as mentioned in Section 3.3.6.3. Therefore, one could explore potentially exploiting DNN for the reconstruction of soft b-jets at the HLT level, targeting the stop signals. Furthermore, there exist dedicated electron, photon and muon MVA identification, with dedicated training also performed for soft leptons (as discussed in Section 4.1) that could be taken advantage of. These topics are covered below in more detail.

Further advanced developments are being investigated and implemented, including sophisticated algorithms for more generic anomaly detection such as New Physics Learning Machine (NPLM) [285] or an autoencoder-based (AE) anomaly detection [286], which is implemented for data quality monitoring (DQM) and even at the L1 trigger level [287].

Soft Objects

Compressed SUSY signatures include very soft decay products, with their p_T s dependent on the Δm . Therefore, the ability of the detectors to efficiently reconstruct soft objects is crucial in the context of compressed SUSY signatures, and any improvements could translate into increased sensitivity.

As discussed in Section 4.2, the analysis already selects leptons with minimum thresholds that border on the limits of the detector design. The acceptance thresholds are guided by efficiency measurements of the final leptonic object selection (identification, isolation and impact parameter), which also reflect the corresponding systematic uncertainties. Therefore, it would be advantageous to target more efficient lepton reconstruction in the low- p_T regime, which could also allow to reduce the minimal thresholds on acceptance further, and, thereby, increase the signal acceptance. The exponential signal spectrum drops very rapidly (Figure 6.7) for the smallest Δm values, so even a relatively small reduction in thresholds could lead to gains in sensitivity.

As mentioned in the ML discussion above, there exist dedicated electron, photon and muon MVA identification, with dedicated training also performed for soft leptons that could be taken advantage of. The typical use cases are in the context of b-hadron decays and quarkonia analyses. A recent development is custom tracker-driven low- p_T electron reconstruction based on BDTs that has been developed to target measurements of B-flavour anomalies (R_K , R_{K^*}) [288, 289], which is used for the delayed reconstruction of the b-parking dataset. Its improvement in efficiency relative to standard electron reconstruction in simulation is shown in Figure 11.2 (left), which is currently being taken advantage of. Another example is the *LowPt* muon ID criterion [290] developed by ATLAS, which is used by the compressed EWKino analysis [291]. The *LowPt* selection accepts candidates composed of track segments in the inner detector matched to track segments from a single station of the muon spectrometer. As shown in Figure 11.2 (right), it has a selection efficiency of $\approx 90\%$ for $p_T < 5$ GeV and results in an improved signal efficiency and background rejection for muons with

$p_T < 10$ GeV.

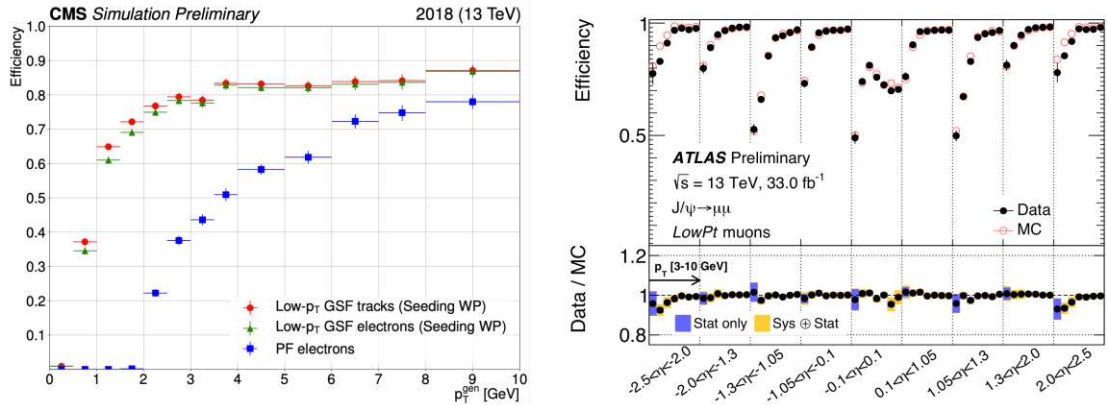


Figure 11.2: Left: Reconstruction efficiency curves for standard CMS electron candidates (blue squares) and for GSF tracks (red circles) and electron candidates from the custom low- p_T electron reconstruction algorithm (green triangles) as a function of the generator-level electron p_T [288, 289]. Right: Muon reconstruction efficiencies for the *LowPt* WP as a function of the muon η for $3 < p_T < 10$ GeV [290].

Therefore, one could revisit some of these reconstruction and identification algorithms in the context of compressed signatures. One could also tailor them for the particular scenario with a customised algorithm, which, for example, could be achieved by performing dedicated training of BDTs or DNNs on the compressed SUSY signals.

Despite possible improvements in lowering the acceptance thresholds as much as possible, a portion of the signal with extremely low lepton p_T will always lie below the acceptance of the detector. Therefore, in these cases where very low p_T leptons are not expected to be reconstructed, one could also take this into consideration. This could be done, for example, with a lepton veto, in an analogous way that b-jets are expected to be too soft to be detected and are therefore vetoed in SRI. Naturally, this type of selection would overlap with the all-hadronic analysis [242], which would need to be taken into account; A statistical combination of results (as presented in Section 8.2) would not be possible with such a selection. Other examples include the compressed vector boson fusion (VBF) search [292], where multiple signal leptons are not expected to be identified due to their extremely low p_T for signal regions targeting $1 < \Delta m < 10$ GeV. For these scenarios, another approach would be to take advantage of soft track signatures left in the detector. For example, the ATLAS compressed EWKino search [291]

uses a signal region based on a lepton and an opposite-sign isolated low- p_T track to increase the sensitivity for the lowest Δm . For these regions, the track is selected to be matched to a reconstructed lepton candidate with no identification requirements, allowing electron and muon candidates to be reconstructed with p_T s as low as 1 GeV and 2 GeV, respectively. Track information could also be taken advantage of to distinguish signal from background. In the stop signals, one would expect a significant amount of tracks coming from the ISR and the b-jet, whereas for $W + jets$, apart from ISR, one would not expect extra hadronic activity in a leptonic W-decay, leaving fewer tracks. One could employ track counting in specific regions of the boosted ISR-like topology. Such discrimination would be particularly useful in cases where the b-jet is too soft to be reconstructed, while the track information still remains.

Identification of b-tagged jets is important for a number of SM measurements and BSM searches, including SUSY signal models with stop decays. As discussed in Section 4.1, standard algorithms rely on calorimeter jet reconstruction with $p_T > 20$ GeV. Ultimately, higher mass splittings are targeted, such as in SR2 which requires at least one b-jet with $30 < p_T < 60$ GeV. The exploitation of machine learning techniques for b-tagging, such as deep machine learning (DeepCSV, DeepJet, ParticleNet), provides superior performance, which is also taken advantage of at the trigger level.

For the smallest mass splittings, many compressed analyses veto b-tagged jets [292, 244, 293, 294, 291], as they are expected to be too soft to be reconstructed by standard algorithms; In the *soft single-lepton analysis* this is the case in SR1. Several analyses [242, 295, 296, 297] take advantage of the sophisticated soft b-tagging algorithms [298, 299] that have been developed, which allow b-tagged jets to be identified with $1 < p_T < 20$ GeV. The ATLAS Track-based Low- p_T Vertex Tagger (T-LVT) [298] and CMS Inclusive Vertex Finder (IVF) [299] algorithms exploit the excellent tracking capabilities of the detectors in order to reconstruct soft b-tagged jets only from tracks or from secondary vertices without the strict requirement of association to a calorimeter jet. The performance of the T-LVT algorithm is shown in Figure 11.3. Therefore, the implementation of such tagging algorithms could help increase the sensitivity to stop

models, which is being performed via a selection on the secondary vertices. Furthermore, one could also potentially explore exploiting these types of algorithms (including DNN) for the reconstruction of soft b-jets at the HLT level. Naturally, one must take into account that the requirement of a b-jet is incompatible with the EWKino signals (Section 9).

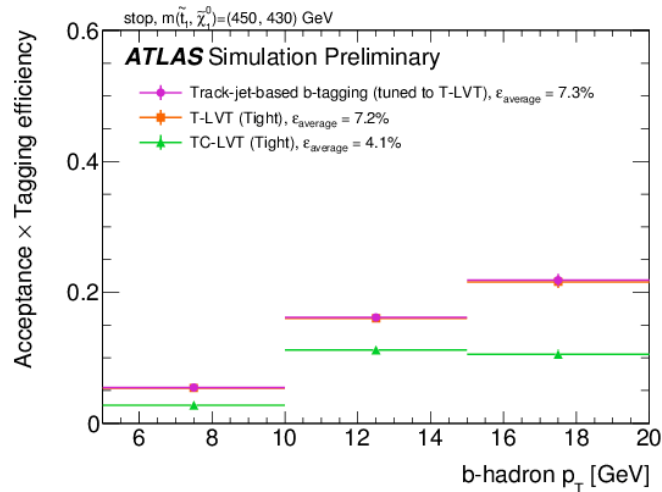


Figure 11.3: Acceptance times tagging efficiency for three algorithms after that of track-jet-based b-tagging as a function of b-hadron p_T measured in simulation [298].

Discriminating Observables

Apart from ML-based discriminators, there is always extra potential in optimising standard observables or exploiting new discriminators, which have not been used in the context of this analysis. Naturally, any new observables could, in turn, serve as additional input variables to ML training. Potential new observables that are promising in the context of compressed SUSY searches are discussed below.

Deconstructed m_T

As discussed in Section 6.4.4.2, the traditional transverse mass m_T observable allows one to take advantage of the difference in signal and background shapes in the kine-

matic edge of the distribution. The considerations in the sensitivity study in the context of the *low-MET* region (Section 10.5) attempt to take further advantage of the m_T tail, which is one of the main sources of sensitivity of the search. This is achieved in the form of additional binning in lepton p_T and m_T , which could also serve as potential improvements for the whole analysis. Due to the sensitivity of the m_T tail, one could also re-optimize the exact cut values (m_T^0) to maximise the signal-to-background ratio. Furthermore, there is also some potential for additional splitting at the low m_T values, dominated by low p_T leptons, corresponding to low Δm .

Nevertheless, a traditional m_T selection is most effective in distinguishing the dominant backgrounds from signals at higher p_T^{miss} values. Moderate or lower p_T^{miss} , characteristic of compressed signatures, generally favours lower m_T values that are largely populated by irreducible backgrounds, resulting in a very difficult region to explore. Moreover, for similar reasons, it is not as efficient for low Δm values due to the correlation with lepton p_T , which is also reflected in lower gains in sensitivity for the *low-MET* region.

A selection on traditional m_T discards some of the information of missing transverse momentum $p_T^{\vec{\text{miss}}}$, along with its correlations, that is used in its definition (Equation 6.3). It is possible to *deconstruct* the transverse mass variables (p_T^{miss} , $p_T(\ell)$) in order to preserve the maximal amount of information about the magnitude and direction of $p_T^{\vec{\text{miss}}}$ [300]. This is achieved by defining two independent variables:

$$\cos \phi = \frac{p_T(\vec{\ell}) \cdot p_T^{\vec{\text{miss}}}}{p_T(\ell) p_T^{\text{miss}}} \quad \text{and} \quad Q \equiv 1 - \frac{m_T^0{}^2}{2 p_T^{\text{miss}} p_T} \quad (11.1)$$

where $\cos \phi$ is the angle between the lepton p_T and $p_T^{\vec{\text{miss}}}$ in the transverse plane and m_T^0 is a fixed cut value chosen to maximise the signal-to-background separation. The choice of m_T^0 is typically beyond the Jacobian peak around the mass of the W-boson ($m_T > m_W > m_T^0$) and in the case of the *soft single-lepton analysis* would correspond to $m_T^0 = 95$ GeV.

The effectiveness of deconstructed m_T is demonstrated in [300], providing additional signal-to-background separation in the context of compressed SUSY searches. It can extend the accessible parameter space beyond the region probed by traditional methods, which can further be expanded to EWKino models.

Preliminary studies in the context of the *soft single-lepton analysis* indicated some potential in providing separation from the dominant $W + jets$ background. Given that m_T is one of the main variables that the signal regions are binned in, a possible implementation requires a substantial redefinition and re-optimisation of the final selection of the search, pending a full sensitivity study with the relevant systematics.

Initial-State Radiation

As discussed in the context of deconstructed m_T , the missing transverse momentum p_T^{miss} signature is often difficult to distinguish from one originating from SM backgrounds. One can take advantage of the correlations in the recoil of the LSPs against the ISR jet, which can provide additional indirect sensitivity and separation from the known backgrounds. There are several proxies that can quantify the recoil, such as the ratio of p_T^{miss} and the ISR jet p_T , $p_T^{\text{miss}}/\sqrt{H_T}$ or in the case of decay chains to a lepton $p_T^{\text{miss}}/W-p_T$ (where $W-p_T$ is the vectorial sum of the lepton p_T and p_T^{miss}), which have been studied in the context of the *low-MET* region in Section 10.5.

Another more sophisticated approach to taking advantage of the correlations in ISR-like topologies is Recursive Jigsaw Reconstruction (RJR) [301, 302]. It is a technique that allows recursively imposing a given decay tree on reconstructed events to yield observables that can exploit kinematic correlations. It sub-divides the event into simplified systems, calculated from the missing transverse momentum and four-momenta of reconstructed objects in their corresponding reference frames. In the case of compressed ISR signatures, RJR sub-divides the topology into an ISR system, as well as visible and invisible sparticle systems, yielding an estimator for the ISR recoil $R_{\text{ISR}} = \text{ISR-}p_T/p_T^{\text{miss}}$ [291, 295] in terms of the ISR system. Since the calculation resolves potential

kinematic and combinatoric ambiguities, such as the provenance of jets, this quantity is more accurate in events with multiple radiated jets.

ISR jets are predicted to originate predominantly from gluons rather than quarks. Based on theoretical principles and experimental measurements, jets initiated by gluons exhibit differences with respect to jets initiated by light-flavour quarks. Due to the different underlying QCD interactions and subsequent hadronisation, they exhibit properties depending on the flavour of the original parton. Gluon-initiated jets have higher particle multiplicity, a more uniform fragmentation function and are less collimated than quark jets, while light-flavour quark jets are more likely to produce narrow jets with hard constituents that carry a significant fraction of the energy [303, 304]. Therefore, gluon ISR jets can be *tagged* in order to distinguish them from certain backgrounds or signal quark jets. An example is using clustering with a larger-sized cone, yielding *large R* jets, and requiring them not to pass the b-tagging selection [242]. The Quark-Gluon Likelihood (QGL) discriminant [303, 305] has been trained to make use of the jet properties (jet energy sharing, multiplicity, angular spread) in order to distinguish gluon from light quark jets. ATLAS studies extracted data-driven templates for several discriminating variables for quark-gluon discrimination [304]. Furthermore, quark-gluon jet discrimination can not only help with ISR-tagging, but also help distinguish signals with quark jets from backgrounds with gluon-enriched jets. In the case of the soft single-lepton analysis, it could also be helpful in reducing the $W + jets$ and $t\bar{t}$ backgrounds. Moreover, as discussed above, quark-gluon tagging is also taken advantage of in the advanced heavy flavour tagging algorithms.

Long-Lived Particles

There is an increasing effort at the LHC for BSM searches for long-lived particles (LLP) [306], with some key topologies shown in Figure II.4. Such searches are very attractive not only because of the various theoretical motivations but also because they can be almost background-free. In parallel, there have been a number of developments in the

experiments to enable the reconstruction of displaced objects with higher efficiency. Apart from the tracker already being indispensable in reconstructing secondary vertices, other examples include improvements already at the trigger level. This involves a number of dedicated HLT algorithms free of beamspot constraints and related reconstruction bias. Furthermore, there are recent developments with FPGA implementations of a Kalman Filter in the BMTF (KMTF) or DNN training in the EMTF, enabling the possibility of triggering on displaced signatures at the L1 level.

As discussed in Section 5.1.2, the stop four-body decay can have a small width and, consequently, the stop can acquire a lifetime for the lowest mass splittings $\Delta m < 30$ GeV (Figure 5.7), which is long enough to observe displaced vertices in the detector. The total width of the stop is dependent on the possible mixing with the competing flavour-violating (FV) two-body decay process ($T2cc$), which is also expected to become dominant at the lowest mass splittings (Figure 5.6). Therefore, this very compressed parameter space could be either dominated by a large stop lifetime or FCNC interactions, which are ultimately dependent on the level of FV couplings in the MSSM. In effect, this means that the assumptions of prompt stop decays with zero lifetime might not be reflective of nature for the lowest Δm and should be taken into account in the modelling and interpretations.

This approach is currently being adopted in the context of a dedicated parallel search for a long-lived stop, where the signal models are extended with different assumptions on the lifetimes of the stop, as well as the mixing with two-body decays. The design of such an analysis requires removing the dependence on objects being compatible with the PV, which is primarily achieved by extending the thresholds on the impact parameter selection. Furthermore, a requirement of a large impact parameter can be used to suppress prompt backgrounds. Such considerations require a re-optimisation of the analysis in a number of aspects due to a rather different phase-space compared to prompt searches. The possible observation of displaced vertices would allow for conclusions on the flavour symmetry of the model.

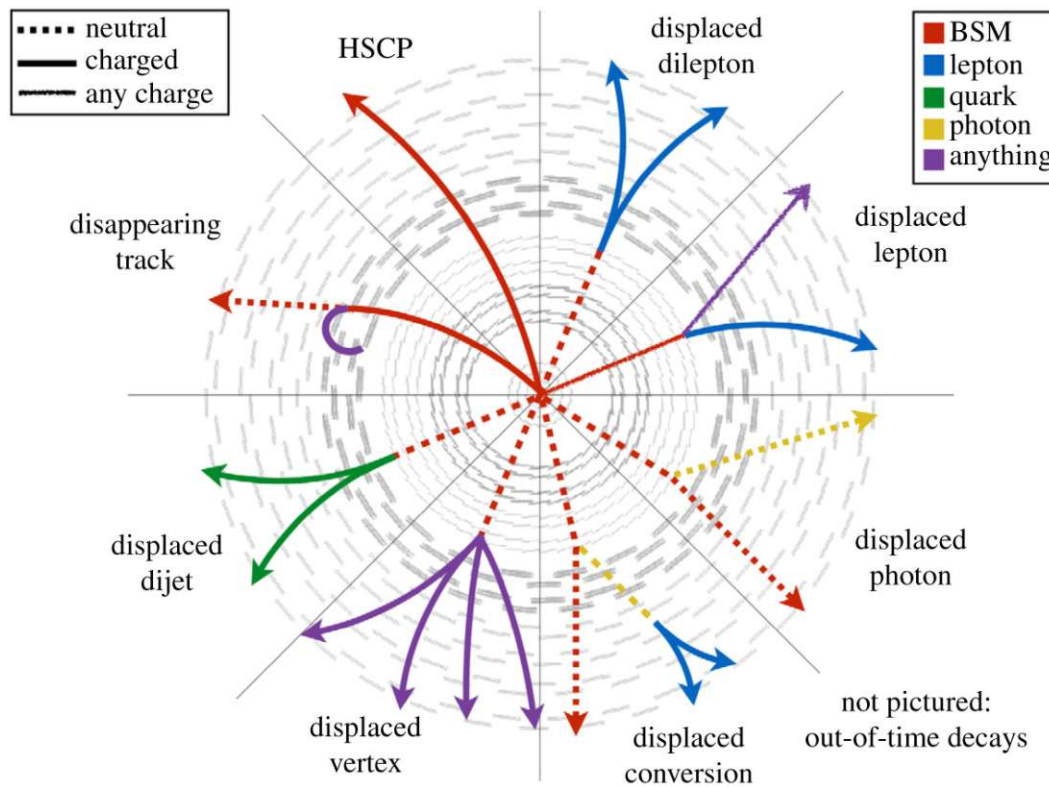


Figure 11.4: Key topologies searched for by LHC experiments in searches for long-lived particles (LLPs) [307].

11.1 Final Remarks

Searches for physics beyond the standard model form a central element of the physics programmes of the CMS and ATLAS, focusing on a large number of different SUSY models [308, 309]. These include ‘classical’ searches focusing on strong (gluino and squark) as well as electroweak (gaugino or slepton) production, targeting standard signatures. Such models were considered the most promising in terms of theoretical motivations and experimental reach in the LHC era, however, unfortunately, they did not lead to discoveries. In light of this, more challenging searches that target less-explored ‘crevasses’ of parameter space, such as models with highly compressed mass spectra considered in this thesis, have become more prominent. Summaries of recent select (and non-exhaustive) results targeting stop and EWKino production, including compressed models, are shown in Appendix D; They show the exclusion of wide regions on a number of SUSY models with sparticles being excluded even beyond 1 TeV.

Ultimately, this puts significant pressure not only on natural SUSY, which favours light superpartners, but SUSY as a whole.

An increased integrated luminosity from LHC *Run 3* and beyond is not expected to bring large gains in sensitivity for many searches. This is not only the case for classical types of searches but also for compressed SUSY searches (e.g. [310]). Therefore, there are strong efforts by the experiments to take advantage of the many innovative experimental and analytical techniques that are being developed. As discussed above, search capabilities are being expanded in previously unexplored directions with the use of sophisticated algorithms, including machine learning, new discriminating observables or improved trigger techniques. The accessible phase-space can be expanded to existing or new models, with sophisticated trigger algorithms, such as those reconstructing displaced vertices targeting exotic LLPs, which have become more sought after in the absence of recent discoveries at the LHC. Parking and scouting at the trigger level also provide useful alternatives to the standard trigger approaches. In parallel, improvements at the trigger level are necessary to mitigate the effects of the harsher LHC conditions due to the non-linearity of rates with PU. A choice that needs to be made in the nearest future between the experiments and LHC is whether the increase in luminosity justifies the reduced phase-space caused by higher trigger thresholds: a question of quality versus quantity.

On the theoretical front, there are equivalently a number of different considerations that must be taken into account in shaping future analyses and interpretations. Firstly, one must remember that the interpretations of the results are performed with simplified models that include a number of different assumptions. Apart from the general assumption that the sparticles in the decay chain are decoupled from the full theory, the 100 % branching ratio is not a realistic one. This is demonstrated in the case of four-body decays competing with the flavour-violating two-body decay process, which ultimately depends on the level of flavour-violation in nature. Furthermore, the assumption that the stop decays are prompt is also likely not reflective of nature due to the predicted long lifetime of the stop for the smallest mass splittings. Therefore, extending the signal models with additional parameters such as lifetimes or widths, as well as mixing with competing processes, are important modelling improvements.

Another more generalised approach to simplified models is that of the pMSSM, which instead of sparticle masses, considers the MSSM mass parameters as free parameters, which determine the mass spectra, branching fractions and cross sections. As with the higgsino pMSSM model, one also typically employs a phenomenological approach with assumptions on the hierarchy of the mass parameters and their values, based on theoretical constraints such as the unification of gauge couplings or cosmological motivations.

The MSSM itself is still a relatively economical model, where a possible soft SUSY breaking mechanism is motivated by naturalness considerations. Alternative SUSY breaking mechanisms have also been suggested, such as gauge-mediated SUSY breaking (GMSB) [126, 127], and there are a number of specific models that are being targeted by dedicated searches; In these models, the gravitino \tilde{G} is assumed the LSP and final states often contain photons [311]. Alternatively, stealth SUSY [312] models predict new light particles with weak-scale supersymmetric masses that feel SUSY breaking only through couplings to the MSSM. The MSSM is also defined to conserve R-parity due to its strong phenomenological motivations, however, from the theoretical point of view, it would not suffer any internal inconsistency if R-parity were violated. There-

fore, there are also searches for R-Parity Violating (RPV) models that do not impose this requirement [313].

Taking into account the different assumptions that are made with respect to an unconstrained SUSY theory, the exclusion limits set at the LHC could be viewed as weaker than they seem. This could, in effect, still keep the window open to natural SUSY. On the other hand, although naturalness has been the guiding principle of the design of most BSM theories, it might not be a necessary requirement. Some theorists do not support naturalness, attributing the anthropic principle to favour low values of the fundamental physical constants. Furthermore, the naturalness criterion appears to fail when applied to the cosmological constant. This motivated the building of a SUSY model that abandons the hierarchy problem and only uses unification and dark matter as the only guiding principles, known as split SUSY [314]. Of course, the least optimal (but not impossible) scenario for us would be that new physics is hidden at energy scales beyond LHC or even any conceivable future collider. Unfortunately, nature probably does not care about our feelings.

Naturally, to uncover this, we need to continue combing through different corners of the SUSY parameter space, with every new result scaling down the realm of possibilities. Consequently, we can certainly relax theoretical motivations or assumptions as well as increase the complexity of the models. On the other hand, the models should still remain feasible to be carried out on relatively short timescales of experimental data-taking periods. Ultimately, it is a matter of finding a pragmatic balance between covering the most phase-space possible while remaining phenomenologically viable.

This pragmatism can come from combining different searches and focusing on general signatures rather than single models. There is a substantial overlap between SM precision measurements and BSM searches, focusing on similar parameter spaces. Within the BSM area, there are also further overlaps of signatures that are sought after, such as monojet (ISR-like) searches for SUSY and exotic particles. Moreover, simplicity does open up the potential for re-interpretations of results, which help cover more models [315, 316]. We should take into account hints of new physics coming from other analy-

ses and experiments, such as various excesses or any other tensions within the SM that might arise. Any residual excesses observed during LHC *Run 2* have to be re-examined with the data arriving from *Run 3*, which will more than double our current dataset.

Nevertheless, if new physics does indeed turn out to be too heavy to be produced at the LHC, then we should still be able to see its effects indirectly with precision measurements of SM parameters at the electroweak scale. This can be probed by theoretical frameworks such as effective field theories (EFTs) [317], which are QFTs without the restriction of renormalisability and are chosen to be valid up to an energy scale that can be probed by LHC. Such approaches are rather different from standard searches for localised excesses in binned analyses, and differences could be very subtle, requiring high-precision measurements. Tensions with the SM could even be accounted for in terms of global EFT fits (e.g. [318] in the top sector), however, the underlying calculations are not trivial. In the absence of discoveries of exotic particles, it is not unlikely that we will enter an area of precision measurements for new physics searches.

All of these considerations will shape the way we perform BSM searches in the upcoming HL-LHC, HE-LHC, as well as future colliders (e.g. FCC, ILC or even muon colliders). A fundamental consideration for possible future colliders is essentially a broader question of quality versus quantity: whether we want to continue pushing to the highest energies and luminosities at hadron colliders with an increasingly difficult background environment (as is the current approach with hadron colliders, including the LHC) or whether we want to focus on cleaner electroweak signatures at lepton colliders at lower energy scales, for precision measurements. These decisions will need to be made in the nearest future and will be highly dependent on whether we find any hints of new physics in the results that are yet to come.

To conclude, starting from detector experts and operations teams forming the foundation to provide high-quality data, how we advance is ultimately a joint effort between analysts and theorists, with a bridge formed by phenomenologists. We need to take advantage of the latest sophisticated techniques and go beyond initial design considerations of the experiments to push to the limits of detection, analysis and theory.

Acknowledgements

Life is very different without you by our side. Your optimism, happiness, energy and especially love are what fuelled all our lives every single day. Thanks to you, I am the person who I am today. All you ever wanted was for us to be happy, and I know that you want us to continue being so - this is what helps me keep my head held up high every day. As promised, I am completing what I started and dedicate it all to your memory. Jestem z Ciebie dumny, Tato.

I thank my mother, Ewa, for everything she has done for me - all her unconditional love in a wonderful childhood. Despite the difficult times we were born into, we were given the opportunities that most do not have in this world. I could not have asked for anything more, and I am grateful for everything. Thank you to my loving grandparents, Babcia Krysia i Gienia and Dziadek Bronek, who shaped our lives - with their good hearts, they always wished for us to do well in life.

To the love of my life, Teresa: from day one, you have always been by my side, supporting me - through the good times as well as the tougher times. None of this would be possible without you by my side, and I am always grateful for that. I couldn't be happier to share my life with you, and I cannot wait to continue with the adventures that await us - especially with little Aurora, who is the light that brightens our lives - eres la luz de nuestra vida. (P.S. I hope you read this thesis when you're a bit older!) With you in my life, my family has grown much bigger and brighter, especially with Damelys, Paco and Sofi.

To my sisters, starting from my ‘younger’ twin, Monika: thank you for all the countless amazing memories and the many more that are yet to come! (Especially with David, where we can continue our philosophical discussions on the deterministic universe and free will, or lack thereof...). I just want to say that I am proud of you. Beata, I am glad to have an older sister like you, and I am always so happy to spend time with your wonderful family: little Lui, not-so-little-anymore Emi and Stefan. I am hoping for more Latin ‘*corazon sentimiento*’ parties or trips like Ischia, where time flows differently!

I would like to thank my supervisors, Manfred, Ivan and Wolfgang, for giving me this opportunity to work on one of the most magnificent and beautiful experiments in the world. Throughout all these years, I have learned more things than I could have imagined, especially entering an entirely new world of high-energy physics. This ranged from many-hour-long sessions debugging analysis code, carrying out weekend night shifts in the CMS control room as a DAQ shifter, to being fully responsible for deploying the L1T and HLT trigger menus, deciding which type of data CMS is going to collect! I cannot lie that it was very challenging at times. Your patience in this process has helped me shape my professional skillset and, taking your guidance as inspiration, at this point, I feel like I can tackle any problem thrown at me with confidence. Nevertheless, as contemplated by ancient philosophers, the more we learn, the more we realise we don’t know, so we should never stop learning.

I am grateful for all the friends and colleagues whom I had an opportunity to meet across HEPHY, DKPI, TU, and especially during my times at CERN, ranging from the SUSY to the L1T and TSG trigger groups and beyond. I would like to, however I cannot name you all (and I dare not miss anyone!), but you know who you are :) Nevertheless, I cannot resist making a special shoutout to the original DEMN group, with whom I shared my first (notorious) office at Wohllebengasse - those were definitely the most fun times! To all my other family and friends who were already in my life or whom I met later during my PhD journey, I thank you for being in my life.

List of Tables

2.1	Extra particle content in the MSSM with particles and their corresponding superpartners, in terms of gauge eigenstates or superfields [120]. . .	26
2.2	Supersymmetric mass eigenstates formed from the gauge eigenstates in the MSSM. The sfermion (squark and slepton) mixing for the first two families is assumed to be negligible. [120]	32
3.1	Pseudorapidity η segmentation of the main subsystems of CMS, including the tracker, ECAL, HCAL and muon detectors [136].	51
3.2	Lowest thresholds and approximate rates of benchmark LIT object seeds at a $\mathcal{L} = 1.8 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ (PU ≈ 50).	82
3.3	Lowest thresholds and rates of main unprescaled HLT paths at a $\mathcal{L} = 1.8 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ (PU ≈ 50). The rate uncertainties are of $\mathcal{O}(\text{Hz})$. [186]	84
3.4	Main $p_{\text{T}}^{\text{miss}}$ -type HLT paths in <i>Run 2</i> and their corresponding seeds. . .	87
4.1	Electron identification variables based on shower/cluster shape, track-cluster matching with ECAL, track vertex as well as rejection of photons conversion. A set of <i>Veto</i> , <i>Loose</i> , <i>Medium</i> and <i>Tight</i> WPs are presented. Upper bounds on the variables are listed for the barrel (endcap) regions with $ \eta_{SC} \leq 1.479$ ($ \eta_{SC} > 1.479$).	101
4.2	Definition of final object selection used in the <i>soft single-lepton analysis</i> presented in Part II.	113

6.1	Definition of preselection and general signal region requirements, indicating splitting in m_T with specific requirements on lepton η and charge $Q(\ell)$	157
6.2	Final signal region binning for both SR1 and SR2, indicating splitting in C_T , m_T and lepton p_T , together with the corresponding control region definition. The very low (VL) lepton p_T bin considers muons only. . . .	158
7.1	Correction factors $R_{ll} * R_{ll/\mu}$ for $N_{\nu\nu l}^{MC}$ for $C_{T1} > 75$ GeV, split into p_T bins. 170	
8.1	Systematic uncertainties on the tight-to-loose ratio non-universality, based on measurements with b-tag variations. The first set of bins corresponds to the SR bins and the final bin corresponds to the CR with $p_T > 30$ GeV	187
8.2	Additional nonprompt estimation non-closure systematic uncertainties in SR1 and CR1 based on the combined MC closure, assigned per individual bin.	188
8.3	Additional nonprompt estimation non-closure systematic uncertainties in SR2 and CR2 based on the combined MC closure, assigned per individual bin.	188
8.4	Typical ranges for relative systematic uncertainties (in %) on the total background or signal prediction in the main SRs. The dash indicates that the uncertainty source is not applicable.	190
10.1	Lowest thresholds of single-object triggers and rates at a $\mathcal{L} = 1.8 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ (PU ≈ 50) relevant to typical compressed SUSY signatures. 211	
10.2	LIT Seed Rates at $\mathcal{L} = 1.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. All <i>SingleMu</i> seeds have a <i>single muon</i> quality (MU-QLTY_SINGL) selection applied. "er" stands for "eta-restricted".	214

10.3	Realistic LIT seed rates at $\mathcal{L} = 1.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. All SingleMu seeds have a single muon quality (MU-QLTY_SINGL) selection applied.	215
10.4	Final HLT path proposal with two (Mu/NoMu) versions of four paths with PF- p_T^{miss} and PF- H_T^{miss} thresholds ranging from 70-100 GeV.	219
10.5	Final LIT seed proposal with two seeds with two thresholds on p_T^{miss} (ETMHF) of 40 and 50 GeV.	219
A.1	Simulated samples of SM background and signal processes. For the background samples, the cross sections used for normalisation are also quoted.	273
A.2	Level-1 Trigger UTM (Micro Trigger Menu) Grammar specification components.	281
A.3	Typical HLT path structure	285

Appendix

A Feynman Diagrams

A.1 Charm Decay

A possible competing process to the four-body stop decays involves the decay to a charm quark, via a flavour-violating (FV) neutral current (FCNC) interaction, as shown in Figure 5.4.

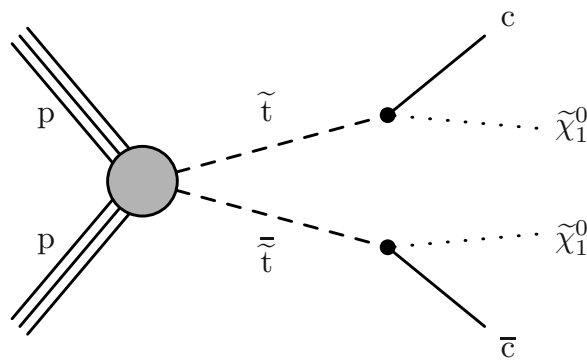


Figure A.1: Simplified signal models for flavour-violating (FV) neutral current (FCNC) stop decays to a charm quark ($\tilde{t}_1 \rightarrow c\tilde{\chi}_1^0$) [247, 242]

A.2 Stop Production

Possible Feynman diagrams contributing to four-body decays of the stop:

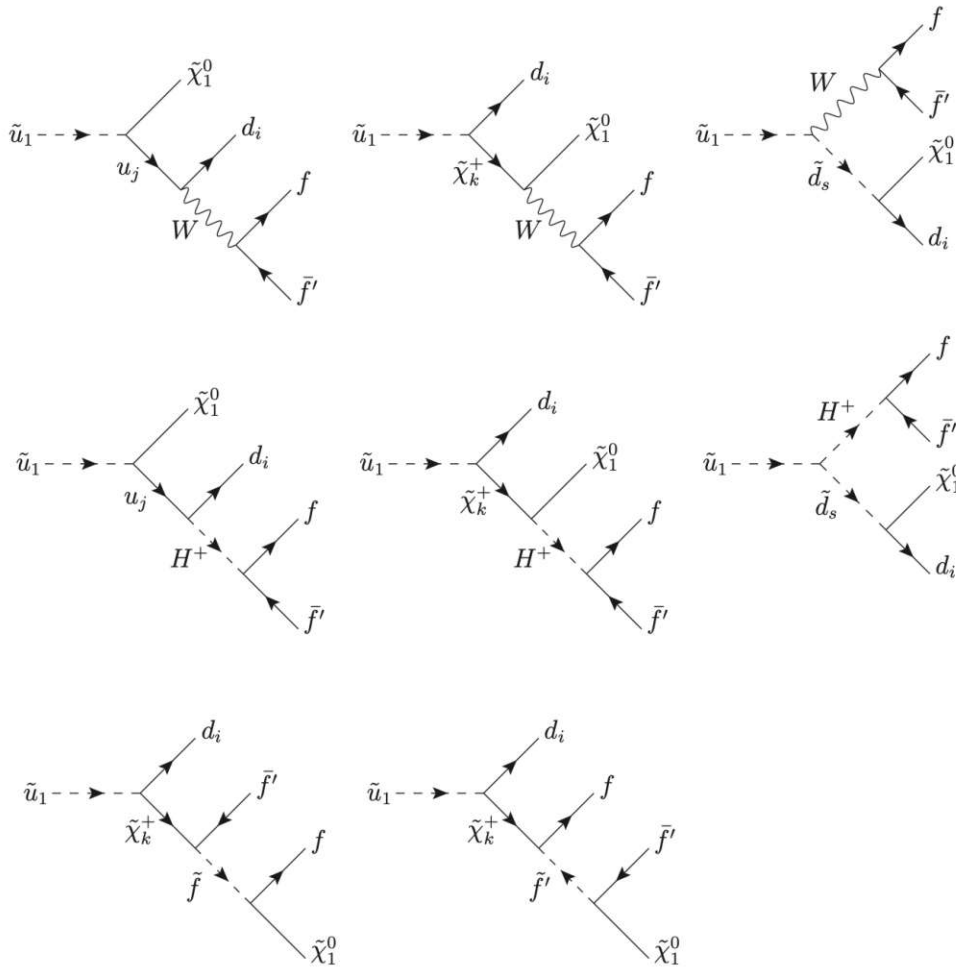


Figure A.2: Simplified signal models that contribute to stop pair-production with 4-body decays [246]

B Analysis

B.1 Samples

Process	Dataset	σ [pb]
W+jets: $H_T \in [70,100]$ $H_T \in [100,200]$ $H_T \in [200,400]$ $H_T \in [400,600]$ $H_T \in [600,800]$ $H_T \in [800,1200]$ $H_T \in [1200,2500]$ $H_T \in [2500,+\infty]$	/WjetsToLNu_HT-70To100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /WjetsToLNu_HT-100To200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /WjetsToLNu_HT-200To400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /WjetsToLNu_HT-400To600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /WjetsToLNu_HT-600To800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /WjetsToLNu_HT-800To1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /WjetsToLNu_HT-1200To2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /WjetsToLNu_HT-2500ToInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/	1596 1627 435.2 59.18 14.58 6.656 1.608 0.03891
t \bar{t} : Dilepton: Single lepton from t Single lepton from \bar{t}	/TTJets_DiLept_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /TTJets_SingleLeptFromT_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /TTJets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/	87.31 182.2 182.2
Z_{inv} : $H_T \in [100,200]$ $H_T \in [200,400]$ $H_T \in [400,600]$ $H_T \in [600,800]$ $H_T \in [800,1200]$ $H_T \in [1200,2500]$ $H_T \in [2500,+\infty]$	/ZjetsToNuNu_HT-100To200_13TeV-madgraph/ /ZjetsToNuNu_HT-200To400_13TeV-madgraph/ /ZjetsToNuNu_HT-400To600_13TeV-madgraph/ /ZjetsToNuNu_HT-600To800_13TeV-madgraph/ /ZjetsToNuNu_HT-800To1200_13TeV-madgraph/ /ZjetsToNuNu_HT-1200To2500_13TeV-madgraph/ /ZjetsToNuNu_HT-2500ToInf_13TeV-madgraph/	345 96.38 13.46 3.962 1.813 0.4411 0.01009
Diboson: WW WZ ZZ	/WW_TuneCUETP8M1_13TeV-pythia8/ /WZ_TuneCUETP8M1_13TeV-pythia8/ /ZZ_TuneCUETP8M1_13TeV-pythia8/	115 47.13 16.52
Z/γ^* , $m_{\ell\ell} \in [5,50]$: $H_T \in [100,200]$ $H_T \in [200,400]$ $H_T \in [400,600]$ $H_T \in [600,+\infty]$ Z/γ^* , $m_{\ell\ell} \in [50,+\infty]$: $H_T \in [100,200]$ $H_T \in [200,400]$ $H_T \in [400,600]$ $H_T \in [600,800]$ $H_T \in [800,1200]$ $H_T \in [1200,2500]$ $H_T \in [2500,+\infty]$	/DYjetsToLL_M-5to50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /DYjetsToLL_M-5to50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /DYjetsToLL_M-5to50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /DYjetsToLL_M-5to50_HT-600toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /DYjetsToLL_M-50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /DYjetsToLL_M-50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /DYjetsToLL_M-50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /DYjetsToLL_M-50_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /DYjetsToLL_M-50_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /DYjetsToLL_M-50_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /DYjetsToLL_M-50_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/	224.2 37.2 3.581 1.124 181.3 50.42 6.984 1.681 0.7754 0.1862 0.004385
Single t: t-channel W-associated Single \bar{t} : t-channel W-associated	/ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1/ /ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1/ /ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1/ /ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1/	136 35.6 80.95 35.6
$t\bar{t}+X$: $t\bar{t}+\gamma$ +jets $t\bar{t}+W$ to $l\nu$ $t\bar{t}+W$ to $q\bar{q}$ $t\bar{t}+Z$ to $ll/\nu\nu$, $M \in [1,10]$ $t\bar{t}+Z$ to $ll/\nu\nu$, $M > 10$ $t\bar{t}+Z$ to $q\bar{q}$	/TTGjets_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/ /TTWjetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/ /TTWjetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/ /TTZToLL_M-lto10_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ /TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8/ /TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/	3.697 0.2043 0.40620 0.0493 0.2529 0.5297
Signal:	/SMS-T2tt_dM-10to80_genHT-160_genMET-80_mWMin-0pl_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 /SMS-T2bW_dM-10to80_genHT-160_genMET-80_mWMin-0pl_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 /TChiWZ_genHT-160_genMET-80_TuneCP2_13TeV-madgraphMLM-pythia8 /MSSM-higgsino_genHT-160_genMET-80_TuneCP2_13TeV-madgraphMLM-pythia8	

Table A.1: Simulated samples of SM background and signal processes. For the background samples, the cross sections used for normalisation are also quoted.

B.2 MVA Analysis Limits

Results of the parallel MVA [249] BDT-based [250] analysis, with a common preselection, discriminating variables and background estimation methods [4, 5]:

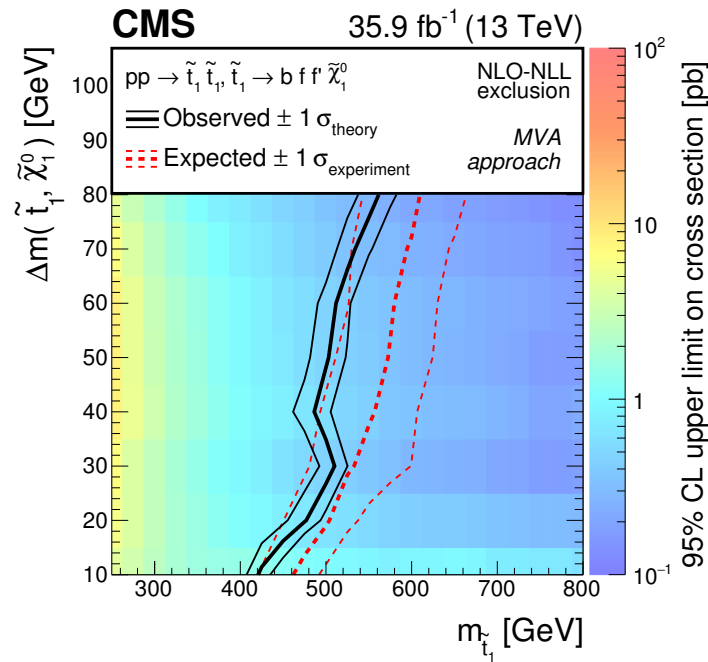


Figure A.3: Exclusion limits at 95% CL on four-body stop decays T2tt (left) and chargino-mediated decays T2bW (right) as a function of $m_{\tilde{t}_1}$ and Δm . A 100 % branching ratio for the chosen decay channel is assumed and the decays are assumed to be prompt. The colour shading corresponds to the observed limit on the cross section. The solid black (dashed red) lines represent the observed (expected) limits, derived using the expected top squark pair-production cross section. The thick lines represent the central values and the thin lines the variations due to the theoretical or experimental uncertainties. [4, 5].

C Soft Trigger

C.1 Signal Distributions

2D signal distributions for the strong T2bW : $\tilde{t}_1 \rightarrow b\tilde{\chi}_1^\pm \rightarrow bf'f'\tilde{\chi}_1^0$ and Higgsino pMSSM signals:

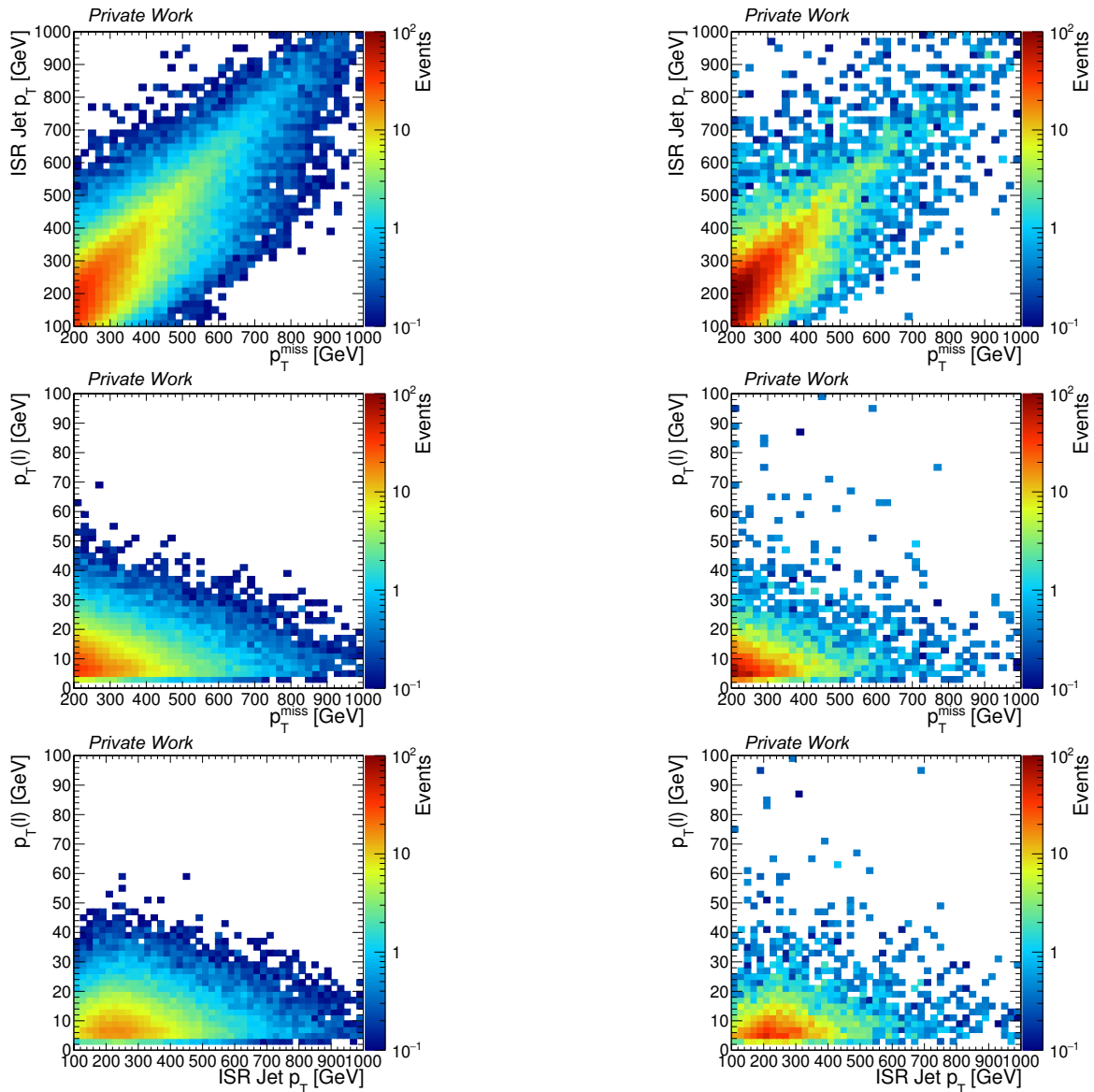


Figure A.4: 2D signal distributions of p_T^{miss} , p_T^{ISR} and $p_T(\mu)$ for the strong T2bW : $\tilde{t}_1 \rightarrow b\tilde{\chi}_1^\pm \rightarrow bf'f'\tilde{\chi}_1^0$ (left) and Higgsino pMSSM (right) signals with masses with a medium $\Delta m = 40$ GeV and μ and $M1$ parameters close to the sensitivity limits

C.2 LIT Seed Efficiency in Simulation

The LIT seed efficiency of the *soft* μ + *hard jet* + *moderate* p_T^{miss} measured in simulation:

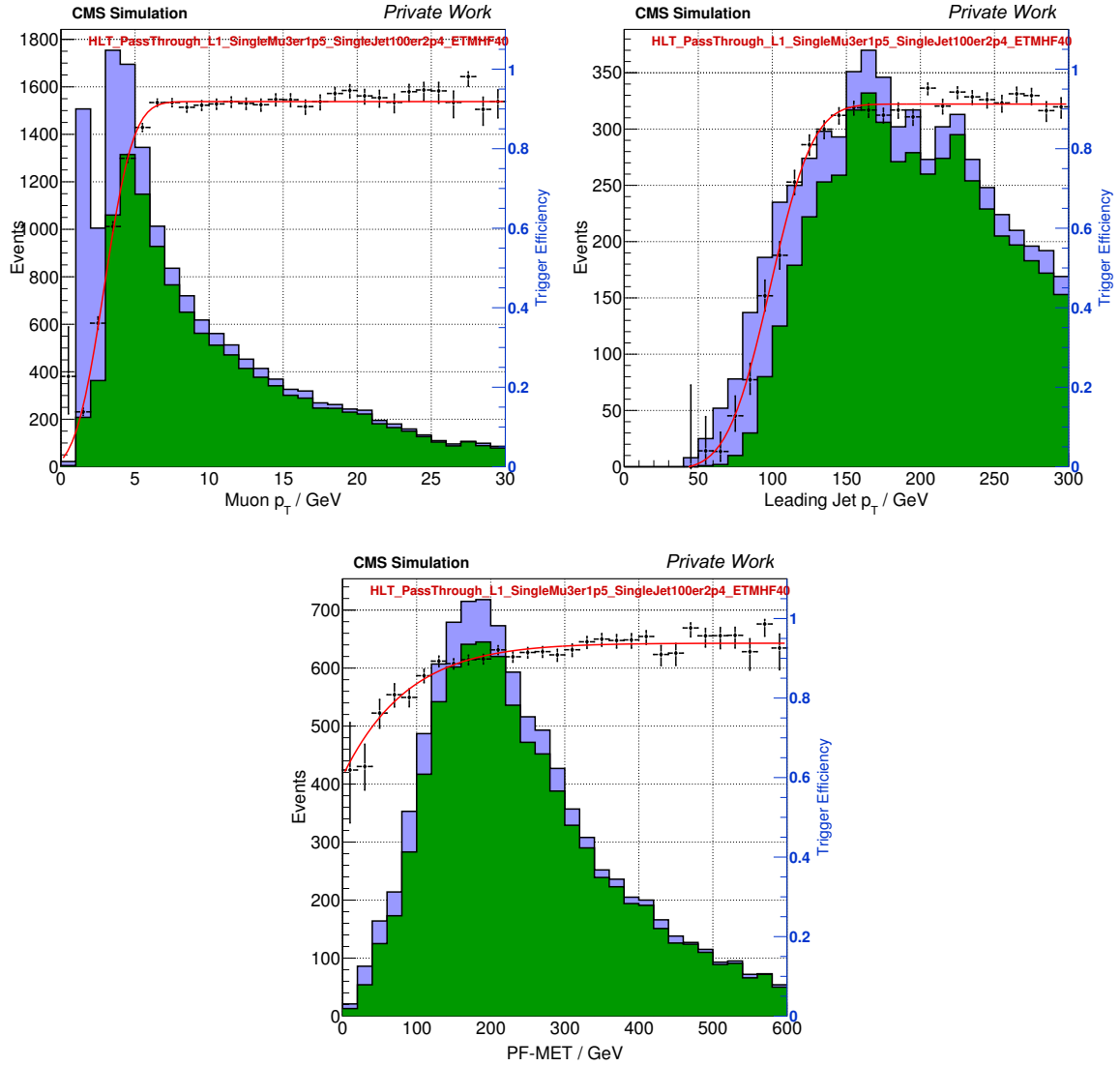


Figure A.5: Efficiency (LIT) of the muon, jet and p_T^{miss} legs of a soft ($p_T > 3$ GeV) muon, high $p_T > 100$ GeV jet and p_T^{miss} cross-trigger measured in a simulated T2tt signal sample ($m_{\tilde{t}_1} = 500$ GeV, $\Delta m = 40$ GeV)

C.3 Low-MET Regions

Summary of background composition of the *low-MET* SRs and CRs:

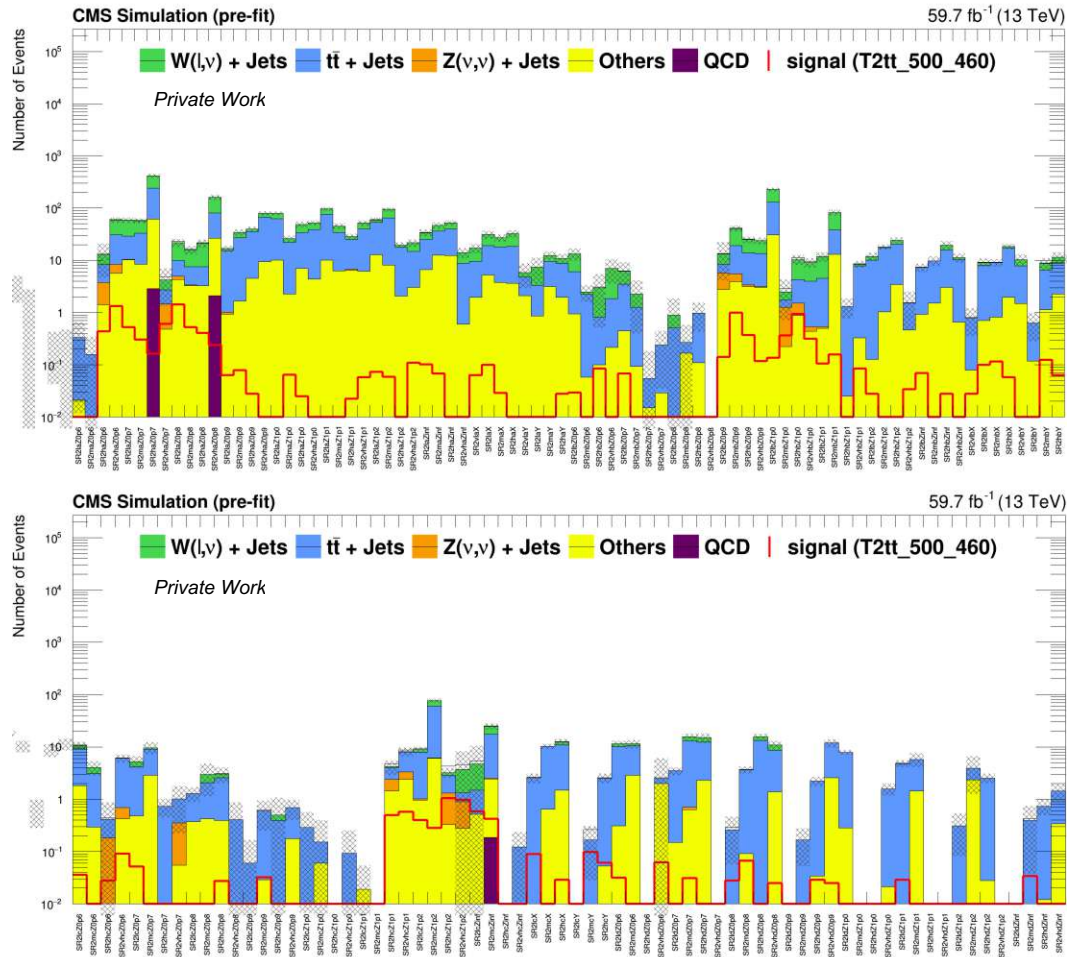


Figure A.6: Background composition of all SR2 bins in the baseline and *low-MET* SRs, against a benchmark $T2t\bar{t}$ ($\tilde{t}_1 \rightarrow b f' \tilde{\chi}_1^0$) signal point with a mass chosen close to the sensitivity limits and a medium $\Delta m = 40$ GeV value.

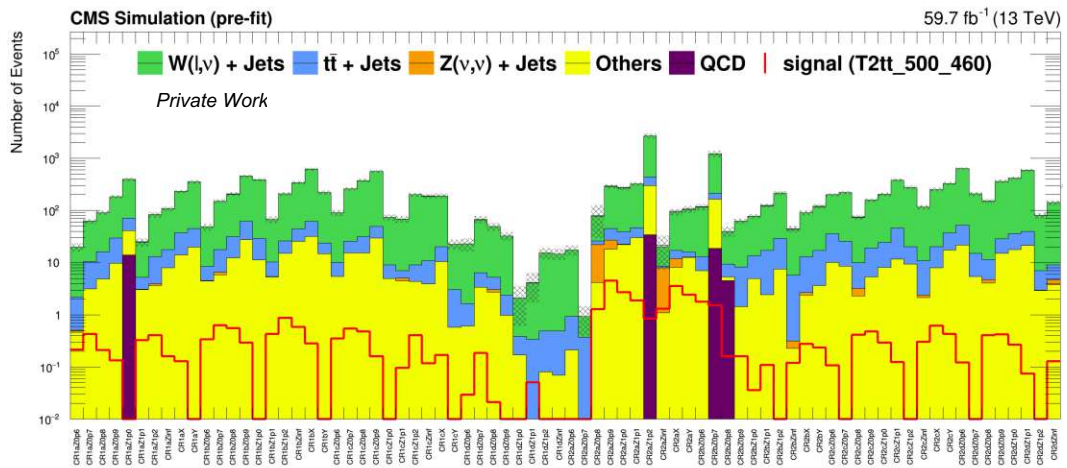


Figure A.7: Background composition of all bins in the baseline and *low-MET* CRs, against a benchmark T2ττ ($\tau_1 \rightarrow b f' f' \tilde{\chi}_1^0$) signal point with a mass chosen close to the sensitivity limits and a medium $\Delta m = 40$ GeV value.

C.4 Sensitivity

Relative difference between the expected upper limits for the combined baseline analysis and *low-MET* region:

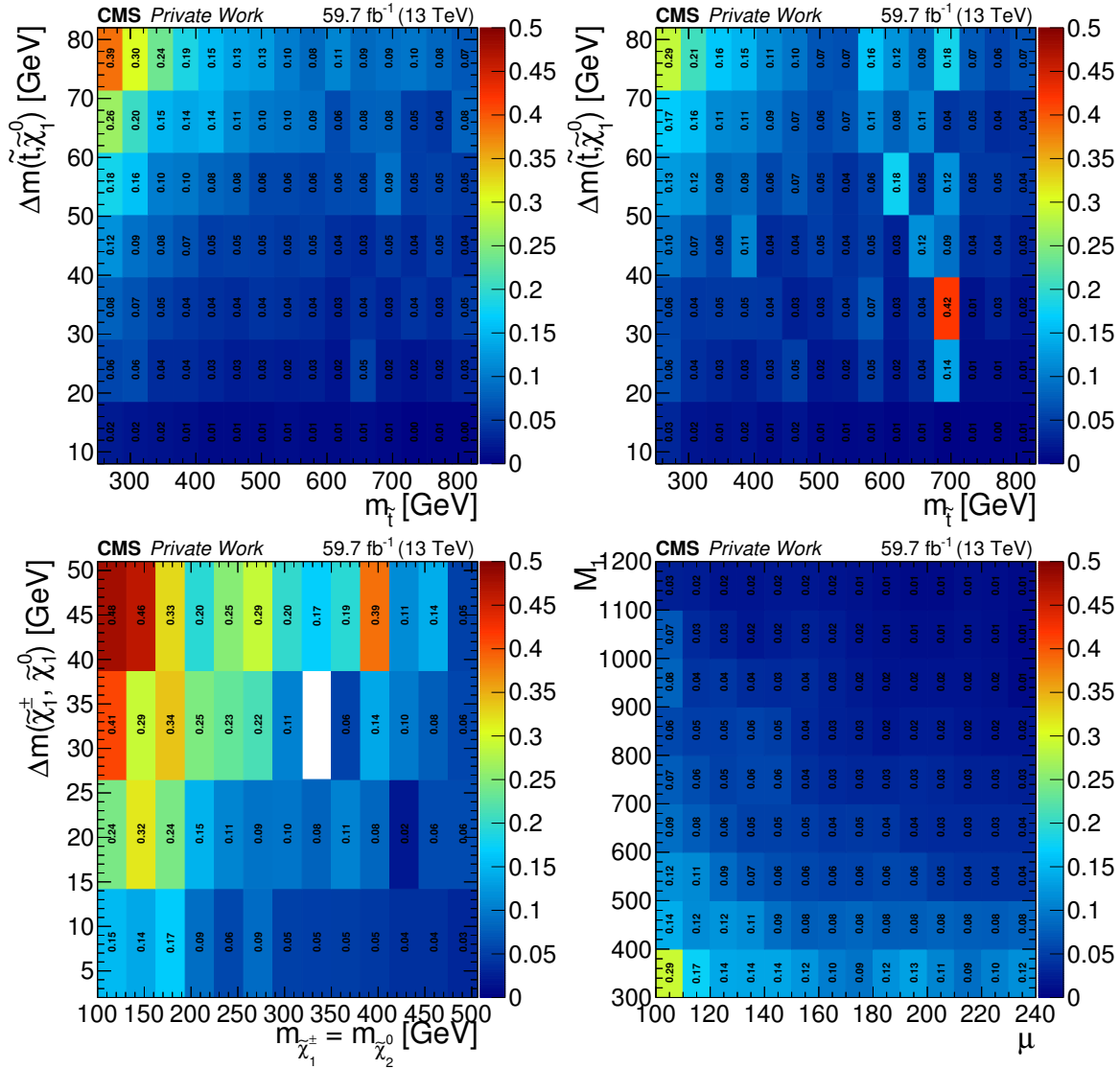


Figure A.8: Relative difference between the expected upper limits for the combined baseline analysis and *low-MET* region for the strongly-produced ($T2tt:\tilde{t}_1 \rightarrow bff'\tilde{\chi}_1^0$ and $T2bW:\tilde{t}_1 \rightarrow b\tilde{\chi}_1^\pm \rightarrow bff'\tilde{\chi}_1^0$) and EWKino ($T\text{ChiWZ}:\tilde{\chi}_1^\pm\tilde{\chi}_2^0 \rightarrow W^*Z^*\tilde{\chi}_1^0\tilde{\chi}_1^0$ and Higgsino pMSSM) signals, calculated using simulated samples corresponding to the 2018 conditions and luminosity $L = 59.8 \text{ fb}^{-1}$.

C.5 Trigger Tools

The development of the seed and corresponding path is performed using the same tools that are used by trigger experts to create the final trigger menus for CMS.

Trigger Menu Editor (TME)

For the Level-1 Trigger, the Trigger Menu Editor (TME) is a graphical user interface (GUI) for creating and modifying LIT menu XML files [319]. The created XML files are used for creating new trigger firmware implementations for the LIT Global Trigger (μGT), as well as generating new emulator code and trigger studies. The main user interface of the tool is shown in Figure A.9.

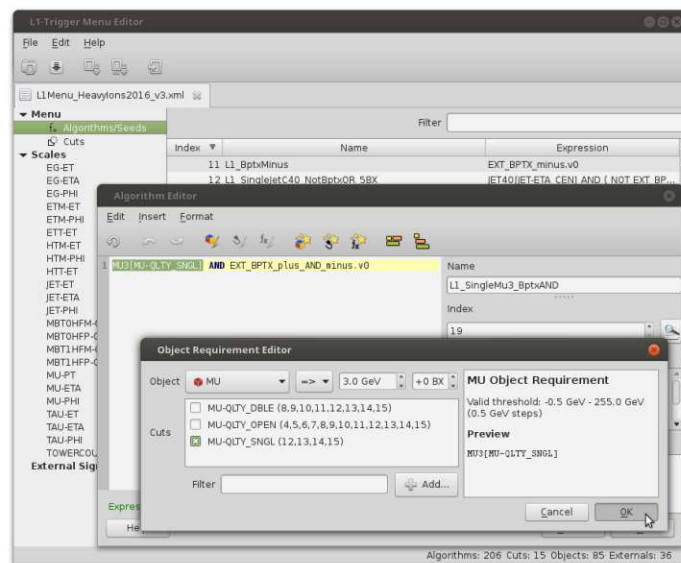


Figure A.9: Screenshot of Trigger Menu Editor (TME), a graphical user interface for creating and modifying LIT menu XML files.

The software application developed privately by CMS and is based on Python3/PyQt5 and makes use of the UTM (Micro Trigger Menu) C++ library. The algorithms are expressed using a purpose built language based on a specialised *Grammar* specification. The UTM grammar contains the pre-defined set of LIT objects (stored in $\pm 2BX$), functions, cuts, external signals and logical operators, as shown in Table A.2.

Type	Description	Example		
Object	physics object	particle	scalar	vector
		MU (μ)	ETT (E_T)	ETM[HF] (p_T^{miss})
		EG (e/γ)	HTT (H_T)	HTM (H_T^{miss})
		JET (jet)		
		TAU (τ)		
Function	logical or mathematical computation based on objects	comb	dist	mass
		coincidence	correlation	invariant mass
Cut	condition applied on objects or on return value of a function	ETA (η)	DETA ($\Delta\eta$)	MASS
		PHI (ϕ)	DPHI ($\Delta\phi$)	ISO
			DR (ΔR)	QLTY
			CHGCOR	CHG
External signal	pre-defined set of binary input signals to the μGT	EXT		
Operators	logical operators	NOT	AND	(X)OR

Table A.2: Level-1 Trigger UTM (Micro Trigger Menu) Grammar specification components.

The grammar is dependent on pre-defined scales and bit assignments for the μGMT and Calorimeter Layer 2 inputs to the μGT . This defines the data structures of the objects, including the granularity and ranges of the possible cuts, as described in a dedicated interface note [320]¹.

An LIT algorithm can be defined in the special expression language as follows:

$$\langle \text{object} \rangle [\langle \text{operator} \rangle] \langle \text{threshold} \rangle [\langle \text{bunchcrossingoffset} \rangle] \quad (\text{A.1})$$

where the object is a physics object, the operator defaults to greater or equal to and threshold in units of GeV, with decimal points expressed as p . Since the μGT stores information in neighbouring bunch crossings ($\pm 2BX$), there exists the possibility of setting a bunch crossing offset, which defaults to zero. The μGT allows for special functionalities, such as the requirement of topological correlations such as $\Delta\eta$, $\Delta\phi$, ΔR and invariant or transverse mass calculations. Different objects can be combined using the *AND* logical operator or using the *comb* function in case of same object types.

¹The versions of the μGT interface note are updated as the system evolves

To demonstrate some examples of algorithm expressions, in increasing order of complexity:

- [L1_SingleMu15er2p1_BX1](#): hypothetical algorithm requiring a single muon with $p_T > 15$ GeV and $|\eta| < 2.1$ and standard 'single muon' quality cuts and in BX +1)

```
MU15 [MU-QLTY_SNGL, MU-ETA_2p1] +1
```

- [L1_DoubleIsoTau32er2p1](#) from one of the core seeds in [Table 3.2](#)

```
comb{TAU32 [TAU-ETA_2p13, TAU-ISO_0xE], TAU32 [TAU-ETA_2p13, TAU-ISO_0xE]}
```

- [L1_DoubleMu4p5er2p0_SQ_OS_Mass_Min7](#)

```
mass_inv{MU4p5 [MU-QLTY_SNGL, MU-ETA_2p0], MU4p5 [MU-QLTY_SNGL, MU-ETA_2p0]}
```

```
[CHGCOR_OS, MASS_MIN_7]
```

- [L1_TripleMu_5_3p5_2p5_DoubleMu_5_2p5_OS_Mass_5to17](#)

```
comb{MU5 [MU-QLTY_DBLE], MU3p5 [MU-QLTY_DBLE], MU2p5 [MU-QLTY_DBLE]}
```

```
AND mass_inv{MU5 [MU-QLTY_DBLE], MU2p5 [MU-QLTY_DBLE]}
```

```
[MASS_MASS_5to17, CHGCOR_OS]
```

ConfDB GUI

As described in 3.3.6.3, HLT menus or configurations are stored within a Configuration Database (ConfDB) system [321]. The HLT configuration contains information on the relations between all paths, sequences, modules and their corresponding parameters, as well as services such as prescale schemes. There exist offline databases for development purposes and online databases for data-taking that are only accessible from the internal network (*.CMS*) of the experiment. The databases are based on the Oracle database management system.

The configurations are created, edited and stored using a dedicated Java application known as the ConfDB GUI. These populate a template that is dependent on a particular release version of the CMSSW software. A converter is then used to translate the menu into a single *python* configuration file, that can be used online on the filter farm during data-taking or offline for Monte Carlo studies.

The main user interface of the ConfDB GUI application is shown in Figure A.10.

The HLT software is based on CMSSW, which is the same software framework used for offline reconstruction and analysis. The HLT configuration initialises modules with specific parameters that run the full HLT algorithms. Reconstruction algorithms designed for fast performance are used to identify candidates for muons, electrons, photons, charged and neutral hadrons. The objects have higher purity than that of the L1T, using inputs from the full detector readout. Thanks to the versatility of a software-based system, a wide range of more complicated algorithms can be implemented, such as special iterative tracking or the tagging of heavy-flavoured jets, via the determination of combined secondary vertices (CSV), including the use of deep neural networks (DNN). Processing-heavy algorithms are ran at a later stage, typically preceded by some initial reconstruction and filtering based on calorimeters and muon detectors.

Filtering modules are used to apply requirements on the reconstructed objects, such

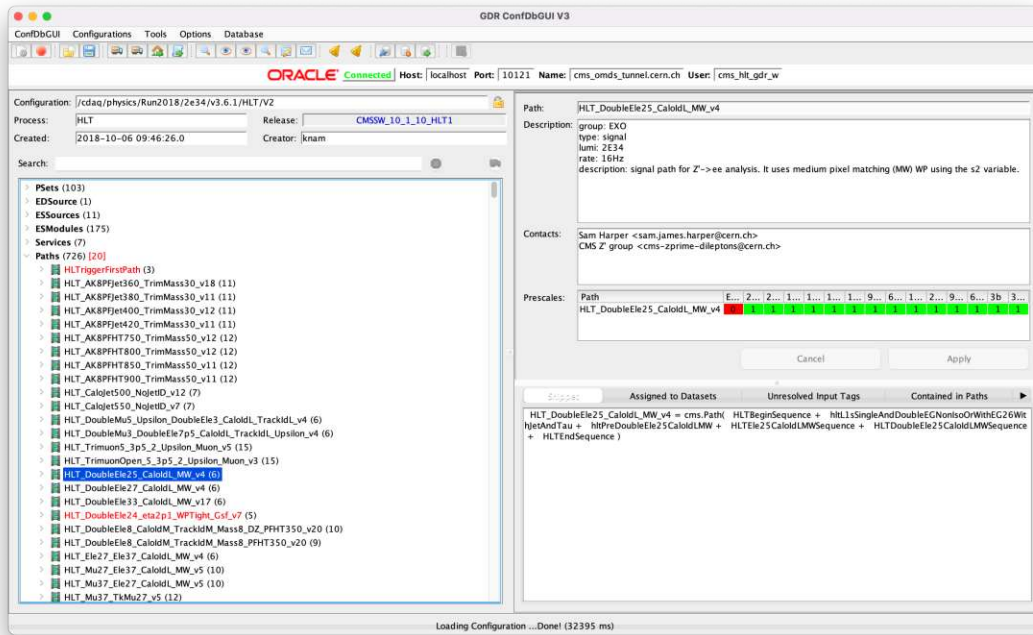


Figure A.10: ConfDB GUI, a graphical user interface for creating and modifying HLT menu configuration files.

as thresholds on energy or momentum or topological requirements. Within a path, the filters are embedded to be ran following early reconstruction algorithms, so that an event is rejected prior to running any further, potentially processing-heavy, algorithms unnecessarily.

Each HLT path is a sequence of specific reconstruction (EDProducer) and filtering (EDFilter) modules. They follow the structure of the Event Data Model (EDM) of CMSSW, that stores the collision data within a C++ type-safe container called `edm:Event`. This includes the *RAW* data (detector data, LIT result, HLT result, HLT objects) together with added *RECO* products and metadata (configuration, and alignment and calibration data), as the events are processed.

A typical HLT path has a standard structure of increasing complexity, as described in Table A.3.

Therefore, each path starts with `HLTBeginSequence`. Each path contains a LIT seed filter as the first module, followed by a prescaler module defining the HLT prescales.

HLT Module	Description
HLTBEGINSequence	special module defining the trigger type, an L1 unpacker sequence and the beam-spot information
hltL1s [L1TSeed]	filter module selecting on given L1T seeds
hltPre [HLTPath]	prescaler module containing HLT prescale values
HLTReco [Object] Sequence	reconstruction sequence
hlt [Object] Filter [Threshold]	filter on reconstructed object
⋮	⋮
HLTEndSequence	returning a True boolean

Table A.3: Typical HLT path structure

Following are series of reconstruction sequences and modules filtering on the reconstructed objects. Specific requirements in the form of cuts can be imposed both in the reconstruction and, naturally, the filtering modules. If all filters pass for an event, the `HLTEndSequence` returns a True boolean, indicating a positive HLT decision.

The typical structure of an HLT path is illustrated in Figure A.11, taking the pure $p_T^{\text{miss}} + H_T^{\text{miss}}$ HLT path `HLT_PFMET120_PFMHT120_IDTight` as an example.

Within the converted python configuration file the path is thus defined by the following snippet of code:

```

HLT_Mu3er1p5_PfJet100er2p5_PFMET80_PFMHT80_IDTight_v1 = cms.
  Path(HLTBeginSequence +
    hltL1Mu3er1p5Jet100er2p5ETMHF400RETMHF50 +
    hltPreMu3er1p5PfJet100er2p5PFMET80PFMHT80IDTight + cms.
    ignore(hltL1sSingleMuOpenObjectMap) +
    hltL1fL1sSingleMuOpenCandidateL1Filtered0 +
    HLTl2muonrecoSequence + cms.ignore(
    hltL2fL1sSingleMuOpenCandidateL1f0L2Filtered0Q) +
    HLTl3muonrecoSequence + cms.ignore(
    hltL1fForIterL3L1fL1sSingleMuOpenCandidateL1Filtered0) +
    hltL3MuFiltered3er1p5 + HLTRecoMETSequence + hltMET50 +
    HLTHBHENoiseCleanerSequence + hltMetClean + hltMETClean40

```

```

+ HLTAK4CaloJetsSequence + hltSingleCaloJet70 + hltMht +
hltMHT50 + HLTAK4PFJetsSequence +
hltPFJetsCorrectedMatchedToCaloJets70 +
hltSinglePFJet100er2p5 + hltPFMHTTightID +
hltPFMHTTightID80 + hltPFMETProducer + hltPFMET80 +
HLTEndSequence)

```

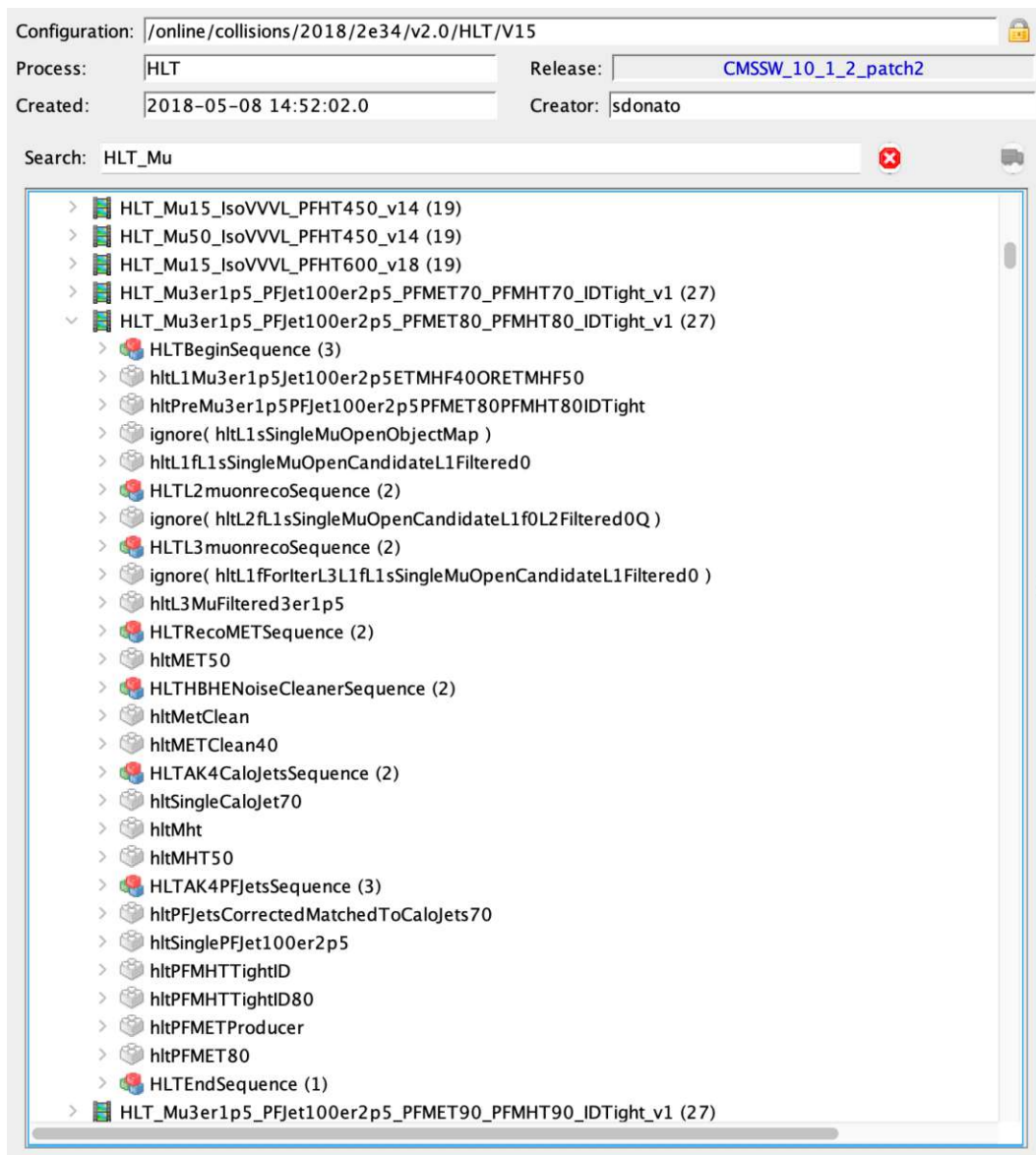


Figure A.11: ConfDB GUI display of the structure of the *soft μ + hard jet + moderate p_T^{miss}* HLT path (`HLT_Mu3er1p5_PFJet100er2p5_PFMET80_PFMHT80_IDTight`)

D Status Quo of Compressed SUSY Searches

There is a significant effort at ATLAS [137] and CMS [136] focusing on compressed SUSY and this section gives a general overview of the most recent results targeting these signatures, starting from a Δm of 1 GeV up to the mass of the top quark $m_t \approx 173$ GeV (top-corridor). In order to illustrate the reach of some analyses in compressed regions, Figures A.12 and A.13 shows a summary of the exclusion limits for models with strong and electroweak SUSY mechanisms, where the compressed regions are near the diagonal. Figures A.12 and A.13 show exclusion limits on stop pair-production (left) as well as chargino-neutralino (χ_1^\pm - χ_2^0) production (right). For stop pair-production, exclusions are made in the plane of the stop \tilde{t}_1 (NLSP) and $\tilde{\chi}_1^0$ neutralino (LSP) masses. The compressed region is interpreted in terms of direct stop decays to a four-body final state, excluding stop masses up to ≈ 700 GeV (for 35.9 fb^{-1}). For chargino-neutralino (χ_1^\pm - χ_2^0) production, exclusions are made in the plane of the chargino $\tilde{\chi}_1^\pm$ (NLSP) and neutralino $\tilde{\chi}_1^0$ (LSP) masses. The compressed region is interpreted in terms of a WZ-mediated model, excluding chargino masses up to ≈ 300 GeV. The strongest limits reach up to $\tilde{t}_1 \approx 1.3$ TeV for both models.

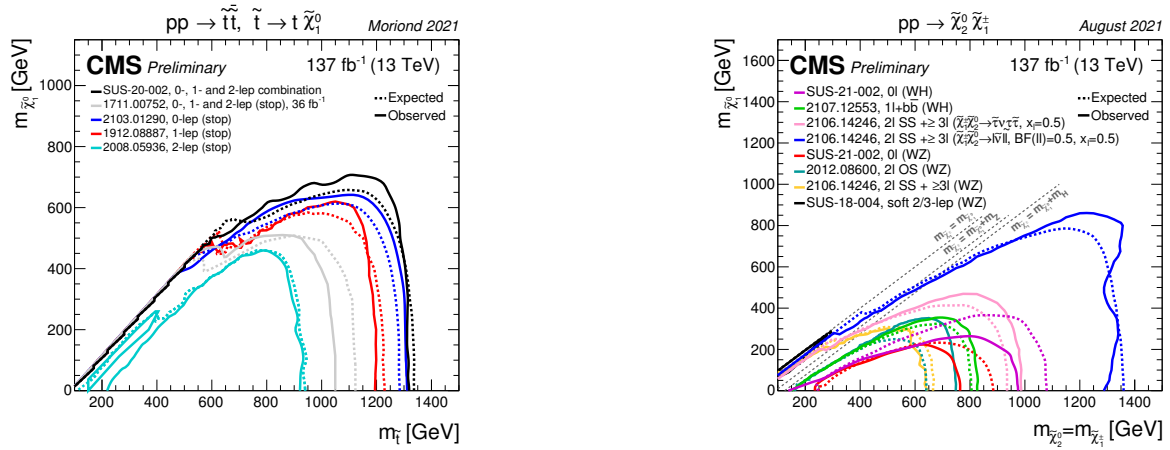


Figure A.12: Summary of limits on CMS SUSY searches [308] for stop pair-production (left) and chargino-neutralino ($\chi_1^\pm - \chi_2^0$) production (right). A 100 % branching ratio is assumed for the chosen decay channel and the decays are assumed to be prompt.

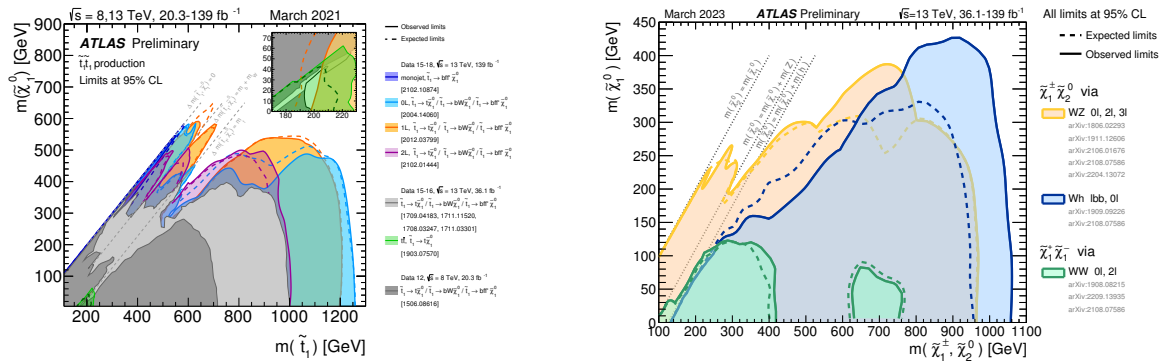


Figure A.13: Summary of limits on ATLAS SUSY searches [309] for stop pair-production (left) and chargino-neutralino ($\chi_1^\pm - \chi_2^0$) production (right). A 100 % branching ratio is assumed for the chosen decay channel and the decays are assumed to be prompt.

Bibliography

- [1] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B716** (2012) 30–61, [arXiv:1207.7235](https://arxiv.org/abs/1207.7235) [hep-ex].
- [2] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1–29, [arXiv:1207.7214](https://arxiv.org/abs/1207.7214) [hep-ex].
- [3] CMS Collaboration, *Search for supersymmetry in events with at least one soft lepton, low jet multiplicity, and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV*, <http://cds.cern.ch/record/2273394?ln=en>.
- [4] CMS Collaboration, *Search for top squarks decaying via four-body or chargino-mediated modes in the single-lepton final state at $\sqrt{s} = 13$ TeV*, <http://cds.cern.ch/record/2308382?ln=en>.
- [5] CMS Collaboration, *Search for top squarks decaying via four-body or chargino-mediated modes in single-lepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Journal of High Energy Physics* **09** no. 9, (2018) 065, 1805.05784. [https://dx.doi.org/10.1007/JHEP09\(2018\)065](https://dx.doi.org/10.1007/JHEP09(2018)065).
- [6] CMS Collaboration, *Search for supersymmetry in events with at least one soft lepton, low jet multiplicity, and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV*, <https://cds.cern.ch/record/2226549>.
- [7] CMS Collaboration, V. Khachatryan et al., *Search for supersymmetry in events with soft leptons, low jet multiplicity, and missing transverse energy in proton proton collisions at $\sqrt{s} = 8$ TeV*, *Phys. Lett. B* **759** (2016) 9, [arXiv:1512.08002](https://arxiv.org/abs/1512.08002) [hep-ex].
- [8] CMS Collaboration, M. Zarucki, *Search for Supersymmetry with a Highly Compressed Mass Spectrum in the Single Soft Lepton Channel with the CMS Experiment at the LHC*, *PoS EPS-HEP2017* (2017) 728.
- [9] CMS Collaboration, M. Zarucki, *Search for SUSY with a Highly Compressed Mass Spectrum in the Soft Single Lepton Channel with the CMS Experiment*, *PoS ALPS2018* (2018) 058. <https://cds.cern.ch/record/2632085>.

- [10] CMS Collaboration, M. Zarucki, *Soft Supersymmetry Searches at ATLAS and CMS*, [PoS LHCP2020 \(2020\) 009](#).
- [11] CMS Collaboration, *Performance of a Soft Muon, Hard Jet and Moderate Missing Energy Trigger in 2018 Data*, <https://cds.cern.ch/record/2710122>.
- [12] D. Griffiths, *Introduction to elementary particles*. 2008.
- [13] M. E. Peskin and D. V. Schroeder, *An Introduction To Quantum Field Theory*. Westview Press, New York, 1995.
- [14] W. Cottingham and D. Greenwood, *An introduction to the standard model of particle Physics*. Cambridge Univ. Press, 2007.
- [15] D. H. Perkins, *Introduction to high energy physics*. Cambridge Univ. Press, 2000.
- [16] J. J. Thomson, *Cathode Rays*, Royal Institution Proceedings no. 15, (1897) 419–432.
- [17] J. J. Thomson, *Cathode Rays*, Philosophical Magazine no. 44, (1897) 293.
- [18] E. Rutherford, *Collision of α particles with light atoms. IV. An anomalous effect in nitrogen*, [Philosophical Magazine 90 no. suppl, \(2010\) 31–37](#), <https://doi.org/10.1080/14786431003659230>.
- [19] J. Chadwick, *Possible Existence of a Neutron*, [Nature 129 \(1932\) 312](#).
- [20] M. Planck, *On the Theory of the Energy Distribution Law of the Normal Spectrum*, Verhandl. Dtsch. phys. Ges. 2 (1900) 237.
- [21] A. Einstein, *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. (German) [On the production and transformation of light from a heuristic viewpoint]*,.
- [22] A. H. Compton, *A Quantum Theory of the Scattering of X-rays by Light Elements*, [Phys. Rev. 21 \(1923\) 483–502](#). <https://link.aps.org/doi/10.1103/PhysRev.21.483>.
- [23] O. Klein, *Quantentheorie und fünfdimensionale Relativitätstheorie*, [Zeitschrift für Physik no. 37, \(1926\)](#).
- [24] W. Gordon, *Der Comptoneffekt nach der Schrödingerschen Theorie*, [Zeitschrift für Physik no. 40, \(1926\)](#).
- [25] P. A. M. Dirac, *The quantum theory of the electron. I.*, <http://www.jstor.org/stable/94981>.
- [26] P. A. M. Dirac, *The quantum theory of the electron. II.*, <http://www.jstor.org/stable/94909>.
- [27] P. A. M. Dirac, *On the annihilation of electrons and protons*,.

- [28] C. D. Anderson, *Cosmic-Ray Positive and Negative Electrons*, *Phys. Rev.* **44** (1933) 406–416.
<https://link.aps.org/doi/10.1103/PhysRev.44.406>.
- [29] O. Chamberlain, E. Segrè, C. Wiegand, and T. Ypsilantis, *Observation of Antiprotons*, *Phys. Rev.* **100** (1955) 947–950. <https://link.aps.org/doi/10.1103/PhysRev.100.947>.
- [30] B. Cork, G. R. Lambertson, O. Piccioni, and W. A. Wenzel, *Antineutrons Produced from Antiprotons in Charge-Exchange Collisions*, *Phys. Rev.* **104** (1956) 1193–1197.
<https://link.aps.org/doi/10.1103/PhysRev.104.1193>.
- [31] F. J. Dyson, *The Radiation Theories of Tomonaga, Schwinger, and Feynman*, *Phys. Rev.* **75** (1949) 486–502. <https://link.aps.org/doi/10.1103/PhysRev.75.486>.
- [32] M. J. Clerk, *dynamical theory of the electromagnetic field*, *APhil. Trans. R. Soc. no. 155*, (1865) 459–512.
- [33] H. Yukawa, *On the Interaction of Elementary Particles I*, *Proc. Phys. Math. Soc. Jap.* **17** (1935) 48–57.
- [34] S. H. Neddermeyer and C. D. Anderson, *Note on the Nature of Cosmic-Ray Particles*, *Phys. Rev.* **51** (1937) 884–886. <https://link.aps.org/doi/10.1103/PhysRev.51.884>.
- [35] 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.* **52** (1937) 1003–1004.
<https://link.aps.org/doi/10.1103/PhysRev.52.1003>.
- [36] M. Conversi, E. Pancini, and O. Piccioni, *On the Disintegration of Negative Mesons*, *Phys. Rev.* **71** (1947) 209–210. <https://link.aps.org/doi/10.1103/PhysRev.71.209>.
- [37] C. Lattes, H. Muirhead, G. Occhialini, and C. Powell, *Processes Involving Charged Mesons*, *Nature* **159** (1947) 694–697.
- [38] D. H. Perkins, *Nuclear Disintegration by Meson Capture*, *Nature* **159** (1947) 126–127.
- [39] G. Rochester and C. Butler, *Evidence for the Existence of New Unstable Elementary Particles*, *Nature* **160** (1947) 855–857.
- [40] P. F. C.F. Powell and D. Perkins, *Mesons*, *Rep. Prog. Phys* **13** (1950) 350.
- [41] V. D. Hopper and S. Biswas, *Evidence Concerning the Existence of the New Unstable Elementary Neutral Particle*, *Phys. Rev.* **80** (1950) 1099–1100.
<https://link.aps.org/doi/10.1103/PhysRev.80.1099>.
- [42] W. Pauli, *Pauli letter collection: letter to Lise Meitner*,. <https://cds.cern.ch/record/83282>.
Typed copy.

- [43] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, *Detection of the Free Neutrino: a Confirmation*, *Science* **124** no. 3212, (1956) 103–104, <https://www.science.org/doi/pdf/10.1126/science.124.3212.103>.
- [44] J. P. Blewett, *The Proton Synchrotron*, *Reports on Progress in Physics* **19** no. 1, (1956) 37. <https://dx.doi.org/10.1088/0034-4885/19/1/302>.
- [45] J. T. Seeman, *The Stanford Linear Collider*, *Ann. Rev. Nucl. Part. Sci.* **41** (1991) 389–428.
- [46] S. Holmes, R. S. Moore, and V. Shiltsev, *Overview of the Tevatron collider complex: goals, operations and performance*, *Journal of Instrumentation* **6** no. 08, (2011) T08001–T08001. <https://doi.org/10.1088/1748-0221/6/08/T08001>.
- [47] *The Super Proton Synchrotron*,. <https://cds.cern.ch/record/1997188>.
- [48] T. Taylor and D. Treille, *The Large Electron Positron Collider (LEP): Probing the Standard Model*, *Adv. Ser. Direct. High Energy Phys.* **27** (2017) 217–261. <https://cds.cern.ch/record/2312570>.
- [49] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*,. <https://cds.cern.ch/record/782076>.
- [50] L. R. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001. 164 p. <https://cds.cern.ch/record/1129806>. This report is an abridged version of the LHC Design Report (CERN-2004-003).
- [51] Aubert, J. J. et al., *Experimental Observation of a Heavy Particle J* , *Phys. Rev. Lett.* **33** (1974) 1404–1406. <https://link.aps.org/doi/10.1103/PhysRevLett.33.1404>.
- [52] Augustin, J. E. et al., *Discovery of a Narrow Resonance in e^+e^- Annihilation*, *Phys. Rev. Lett.* **33** (1974) 1406–1408. <https://link.aps.org/doi/10.1103/PhysRevLett.33.1406>.
- [53] Perl, M. L. et al., *Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation*, *Phys. Rev. Lett.* **35** (1975) 1489–1492. <https://link.aps.org/doi/10.1103/PhysRevLett.35.1489>.
- [54] R. Brandelik et al., *Evidence for planar events in e^+e^- annihilation at high energies*, *Physics Letters B* **86** no. 2, (1979) 243–249. <https://www.sciencedirect.com/science/article/pii/037026937990830X>.
- [55] Barber, D. P. et al., *Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA*, *Phys. Rev. Lett.* **43** (1979) 830–833. <https://link.aps.org/doi/10.1103/PhysRevLett.43.830>.
- [56] CDF Collaboration, *Observation of Top Quark Production in $\bar{P}-P$ Collisions with the Collider Detector at Fermilab*, *Physical Review Letters* **74** no. 14, (1995) 2626–2631. <https://doi.org/10.1103/PhysRevLett.74.2626>.

- [57] D0 Collaboration, *Observation of the Top Quark*, *Physical Review Letters* **74** no. 14, (1995) 2632–2637. <https://doi.org/10.1103/PhysRevLett.74.2632>.
- [58] UA1 Collaboration, G. Arnison et al., *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $s^{1/2} = 540\text{-GeV}$* , *Phys. Lett.* **122B** (1983) 103–116. [611(1983)].
- [59] UA2 Collaboration, M. Banner et al., *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, *Phys. Lett.* **122B** (1983) 476–485.
- [60] UA1 Collaboration, *Experimental Observation of Lepton Pairs of Invariant Mass Around $95\text{-GeV}/c^2$ at the CERN SPS Collider*, *Phys. Lett.* **126B** (1983) 398–410.
- [61] UA2 Collaboration, *Evidence for $Z^0 \rightarrow e^+ e^-$ at the CERN anti-p p Collider*, *Phys. Lett.* **129B** (1983) 130–140.
- [62] H. Weyl, *Electron and Gravitation. I. (In German)*, *Z. Phys.* **56** (1929) 330–352.
- [63] E. C. G. Stueckelberg, *Interaction forces in electrodynamics and in the field theory of nuclear forces*, *Helv. Phys. Acta* **11** (1938) 299–328.
- [64] E. Wigner, *Invariance in physical theory.*, *Proceedings of the American Philosophical Society* **93** 7 (1949) 52–6.
- [65] T. Nakano and K. Nishijima, *Charge Independence for V-particles*, *Prog. Theor. Phys.* **10** (1953) 581–582.
- [66] K. Nishijima, *Charge Independence Theory of V Particles*, *Prog. Theor. Phys.* **13** no. 3, (1955) 285–304.
- [67] M. Gell-Mann, *The interpretation of the new particles as displaced charge multiplets*, *Nuovo Cim.* **4** no. S2, (1956) 848–866.
- [68] M. Gell-Mann, *The Eightfold Way: A Theory of strong interaction symmetry.*,
- [69] Y. Ne'eman, *Derivation of strong interactions from a gauge invariance*, *Nucl. Phys.* **26** (1961) 222–229.
- [70] Barnes, V. E. et al., *Observation of a Hyperon with Strangeness Minus Three*, *Phys. Rev. Lett.* **12** (1964) 204–206. <https://link.aps.org/doi/10.1103/PhysRevLett.12.204>.
- [71] G. Zweig, *An $SU(3)$ model for strong interaction symmetry and its breaking. Version 2.*,
- [72] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, *Phys. Lett.* **8** (1964) 214–215.

- [73] G. Zweig, *Origins of the Quark Model*, in *4th International Conference on Baryon Resonances*. 7, 1980.
- [74] Bloom, E. D. et al., *High-Energy Inelastic e-p Scattering at 6° and 10°*, *Phys. Rev. Lett.* **23** (1969) 930–934. <https://link.aps.org/doi/10.1103/PhysRevLett.23.930>.
- [75] Breidenbach, M. et al., *Observed Behavior of Highly Inelastic Electron-Proton Scattering*, *Phys. Rev. Lett.* **23** (1969) 935–939. <https://link.aps.org/doi/10.1103/PhysRevLett.23.935>.
- [76] S. Weinberg, *Recent progress in gauge theories of the weak, electromagnetic and strong interactions*, *Rev. Mod. Phys.* **46** (1974) 255–277. [*J. Phys. Colloq.*34,no.C1,45(1973)].
- [77] S. L. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579–588.
- [78] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264–1266.
- [79] Gargamelle Neutrino Collaboration, *Observation of Neutrino Like Interactions without Muon or Electron in the Gargamelle Neutrino Experiment*, *Nucl. Phys.* **B73** (1974) 1–22.
- [80] S. Gilardoni and D. Manglunki, eds., *Fifty years of the CERN Proton Synchrotron: Volume 1*. CERN Yellow Reports: Monographs. 2011.
- [81] J.-P. Burnet et al., *Fifty years of the CERN Proton Synchrotron : Volume 2*, [arXiv:1309.6923](https://arxiv.org/abs/1309.6923) [[physics.acc-ph](https://arxiv.org/abs/1309.6923)].
- [82] M. Chaichian and K. Nishijima, *An Essay on Color Confinement*, [arXiv:hep-th/9909158](https://arxiv.org/abs/hep-th/9909158) [[hep-th](https://arxiv.org/abs/hep-th/9909158)].
- [83] W. Pauli, *Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren*, *Z. Physik* **31** (1925) 765–783.
- [84] *Standard Model of Elementary Particles*,
https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.
Accessed: 27/03/2023.
- [85] E. Noether, *Invariante Variationsprobleme*, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1918** (1918) 235–257.
<http://eudml.org/doc/59024>.
- [86] B. Hall, *Lie Groups, Lie Algebras and Representations*. Springer Science & Business Media, 2003.
- [87] S. Coleman and J. Mandula, *All Possible Symmetries of the S Matrix*, *Phys. Rev.* **159** (1967) 1251–1256. <https://link.aps.org/doi/10.1103/PhysRev.159.1251>.
- [88] C. N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, *Phys. Rev.* **96** (1954) 191–195. <https://link.aps.org/doi/10.1103/PhysRev.96.191>.

- [89] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, *Phys. Rev. Lett.* **30** (1973) 1343–1346.
- [90] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346–1349. <https://link.aps.org/doi/10.1103/PhysRevLett.30.1346>.
- [91] M. Goldhaber, L. Grodzins, and A. W. Sunyar, *Helicity of Neutrinos*, *Phys. Rev.* **109** (1958) 1015–1017.
- [92] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, *Phys. Rev.* **105** (1957) 1413–1414.
- [93] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652–657.
- [94] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531–533. <https://link.aps.org/doi/10.1103/PhysRevLett.10.531>.
- [95] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508–509.
- [96] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321–323.
- [97] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585–587.
- [98] Y. Nambu, *Quasi-Particles and Gauge Invariance in the Theory of Superconductivity*, *Phys. Rev.* **117** (1960) 648–663. <https://link.aps.org/doi/10.1103/PhysRev.117.648>.
- [99] J. Goldstone, *Field Theories with Superconductor Solutions*, *Nuovo Cim.* **19** (1961) 154–164.
- [100] H. Greaves and T. Thomas, *On the CPT theorem*, *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **45** (2014) 46–65. <https://doi.org/10.1016%2Fj.shpsb.2013.10.001>.
- [101] CMS Collaboration, *Summaries of CMS cross section measurements*, Public CMS Wiki (2021). <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined>.
- [102] B. Pontecorvo, *Mesonium and anti-mesonium*, *Sov. Phys. JETP* **6** (1957) 429.
- [103] B. Pontecorvo, *Inverse β processes and nonconservation of lepton charge*, *Zhur. Eksptl'. i Teoret. Fiz.* **34** (1958). <https://www.osti.gov/biblio/4349231>.
- [104] B. Pontecorvo, *Neutrino Experiments and the Problem of Conservation of Leptonic Charge*, *Zh. Eksp. Teor. Fiz.* **53** (1967) 1717–1725.

- [105] R. Davis, *Solar Neutrinos. II. Experimental*, *Phys. Rev. Lett.* **12** (1964) 303–305.
<https://link.aps.org/doi/10.1103/PhysRevLett.12.303>.
- [106] Super-Kamiokande Collaboration, Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562–1567, [arXiv:hep-ex/9807003](https://arxiv.org/abs/hep-ex/9807003) [hep-ex].
- [107] SNO Collaboration, Q. R. Ahmad et al., *Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by 8B solar neutrinos at the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **87** (2001) 071301, [arXiv:nucl-ex/0106015](https://arxiv.org/abs/nucl-ex/0106015) [nucl-ex].
- [108] KATRIN Collaboration, *Direct neutrino-mass measurement with sub-electronvolt sensitivity*, *Nature Physics* **18** no. 2, (2022) 160–166. <https://doi.org/10.1038/s41567-021-01463-1>.
- [109] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the Unified Model of Elementary Particles*, *Progress of Theoretical Physics* **28** no. 5, (1962) 870–880.
<https://doi.org/10.1143/PTP.28.870>.
- [110] NEXT Collaboration, *NEXT, a HPGXe TPC for neutrinoless double beta decay searches*, [arXiv:0907.4054](https://arxiv.org/abs/0907.4054) [hep-ex].
- [111] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, *Helvetica Physica Acta* **6** (1933) 110–127.
- [112] V. C. Rubin, J. Ford, W. K., and N. Thonnard, *Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 ($R=4kpc$) to UGC 2885 ($R=122kpc$).*, *ApJ* **238** (1980) 471–487.
- [113] D. Maoz and H.-W. Rix, *Early-Type Galaxies, Dark Halos, and Gravitational Lensing Statistics*, *ApJ* **416** (1993) 425.
- [114] D. Clowe, M. Bradac, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky, *A direct empirical proof of the existence of dark matter*, *Astrophys. J.* **648** (2006) L109–L113, [arXiv:astro-ph/0608407](https://arxiv.org/abs/astro-ph/0608407) [astro-ph].
- [115] Planck Collaboration, *Planck 2018 results*, *Astronomy & Astrophysics* **641** (2020) A6.
<https://doi.org/10.1051/2F0004-6361/2F201833910>.
- [116] L. Bergström, *Dark matter candidates*, *New Journal of Physics* **11** no. 10, (2009) 105006.
<https://doi.org/10.1088/2F1367-2630/2F11/2F10/2F105006>.
- [117] J. L. Feng, *Dark Matter Candidates from Particle Physics and Methods of Detection*, *Annual Review of Astronomy and Astrophysics* **48** no. 1, (2010) 495–545,
<https://doi.org/10.1146/annurev-astro-082708-101659>.

- [118] T. M. Undagoitia and L. Rauch, *Dark matter direct-detection experiments*, *Journal of Physics G: Nuclear and Particle Physics* **43** no. 1, (2015) 013001.
<https://doi.org/10.1088%2F0954-3899%2F43%2F1%2F013001>.
- [119] LIGO Scientific Collaboration and Virgo Collaboration, *Observation of Gravitational Waves from a Binary Black Hole Merger*, *Phys. Rev. Lett.* **116** (2016) 061102.
<https://link.aps.org/doi/10.1103/PhysRevLett.116.061102>.
- [120] S. P. Martin, *A Supersymmetry Primer*, *Adv. Ser. Direct. High Energy Phys.* **18** (1998) 1, [arXiv:hep-ph/9709356](https://arxiv.org/abs/hep-ph/9709356) [hep-ph].
- [121] P. Fayet, *Supersymmetry and Weak, Electromagnetic and Strong Interactions*, *Physics Letters B* **64** (1976) 159.
- [122] P. Fayet, *Spontaneously Broken Supersymmetric Theories of Weak, Electromagnetic and Strong Interactions*, *Physics Letters B* **69** (1977) 489.
- [123] R. Barbieri and G. Giudice, *Upper Bounds on Supersymmetric Particle Masses*, *Nuclear Physics B* **306** (1987) 63–76. <https://cds.cern.ch/record/180560>.
- [124] R. Barbieri and D. Pappadopulo, *S-particles at their naturalness limits*, *Journal of High Energy Physics* **2009** no. 10, (2009) 061–061.
<https://doi.org/10.1088%2F1126-6708%2F2009%2F10%2F061>.
- [125] R. Arnowitt and P. Nath, *Developments in Supergravity Unified Models*,
https://doi.org/10.1142%2F9789814307505_0003.
- [126] C. Kolda, *Gauge-mediated supersymmetry breaking: Introduction, review and update*, *Nuclear Physics B - Proceedings Supplements* **62** no. 1-3, (1998) 266–275.
<https://doi.org/10.1016%2Fs0920-5632%2897%2900667-1>.
- [127] G. Giudice and R. Rattazzi, *Theories with gauge-mediated supersymmetry breaking*, *Physics Reports* **322** no. 6, (1999) 419–499.
<https://doi.org/10.1016%2Fs0370-1573%2899%2900042-3>.
- [128] J. R. Ellis and S. Rudaz, *Search for Supersymmetry in Toponium Decays*, *Phys. Lett. B* **128** (1983) 248–252.
- [129] A. Delgado, G. F. Giudice, G. Isidori, M. Pierini, and A. Strumia, *The light stop window*, *Eur. Phys. J. C* **73** no. 3, (2013) 2370, [arXiv:1212.6847](https://arxiv.org/abs/1212.6847) [hep-ph].
- [130] C. Balázs, M. Carena, and C. E. M. Wagner, *Dark Matter, Light Stops and Electroweak Baryogenesis*, *Physical Review D* **70** (2004) 015007, [arXiv:hep-ph/0403224](https://arxiv.org/abs/hep-ph/0403224) [hep-ph].

- [131] J. Edsjö and P. Gondolo, *Neutralino Relic Density including Coannihilations*, *Physical Review D* **56** (1997) 1879–1894, [arXiv:hep-ph/9704361](https://arxiv.org/abs/hep-ph/9704361).
- [132] K. Griest and D. Seckel, *Three exceptions in the calculation of relic abundances*, *Phys. Rev. D* **43** (1991) 3191–3203. <https://link.aps.org/doi/10.1103/PhysRevD.43.3191>.
- [133] E. Mobs, *The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019*, <https://cds.cern.ch/record/2684277>. General Photo.
- [134] K. Hanke, *Past and Present Operation of the CERN PS Booster*, *International Journal of Modern Physics A* **28** no. 13, (2013) 1330019, <https://doi.org/10.1142/S0217751X13300196>.
- [135] M. Chanel, *LEIR: The low energy ion ring at CERN*, *Nucl. Instrum. Meth. A* **532** (2004) 137–143.
- [136] CMS Collaboration, *The CMS Experiment at the CERN LHC*, *Journal of Instrumentation* **3** no. 08, (2008) S08004.
- [137] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *Journal of Instrumentation* **3** no. 08, (2008) S08003.
- [138] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002.
- [139] LHCb Collaboration, *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005.
- [140] TOTEM Collaboration, *Performance of the TOTEM Detectors at the LHC*, *International Journal of Modern Physics A* **28** no. 31, (2013) 1330046. <https://doi.org/10.1142/S0217751X13300469>.
- [141] E. Berti, *Highlight: Forward Physics (LHCf+ FASER)*, [arXiv:2109.04974](https://arxiv.org/abs/2109.04974) [hep-ex].
- [142] J. Pinfold, *MoEDAL-MAPP, an LHC Dedicated Detector Search Facility*, [arXiv:2209.03988](https://arxiv.org/abs/2209.03988) [hep-ph].
- [143] LHC Collaboration, *SND@LHC: The Scattering and Neutrino Detector at the LHC*, [arXiv:2210.02784](https://arxiv.org/abs/2210.02784) [hep-ex].
- [144] CMS-TOTEM Collaboration, *CMS-TOTEM Precision Proton Spectrometer*, <https://cds.cern.ch/record/1753795>.
- [145] CMS Collaboration, *Technical proposal for the upgrade of the CMS detector through 2020*, <https://cds.cern.ch/record/1355706>.
- [146] CMS Collaboration, *CMS Technical Design Report for the Pixel Detector Upgrade*, <https://cds.cern.ch/record/1481838>.
- [147] CMS Collaboration, *Public CMS Luminosity Information*, Public CMS Wiki (2022). <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.

- [148] G. Apollinari, I. Béjar Alonso, O. Brüning, M. Lamont, and L. Rossi, *High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2015. <https://cds.cern.ch/record/2116337>.
- [149] FCC Collaboration, *HE-LHC: The High-Energy Large Hadron Collider*, *The European Physical Journal Special Topics* **228** no. 5, (2019) 1109–1382. <https://doi.org/10.1140/epjst/e2019-900088-6>.
- [150] Benedikt, Michael et al., *Future Circular Collider*,. <https://cds.cern.ch/record/2651299>. Published in Eur. Phys. J. ST.
- [151] Philip Bambade et al., *The International Linear Collider: A Global Project*, [arXiv:1903.01629](https://arxiv.org/abs/1903.01629) [hep-ex].
- [152] CERN, *CERN Yellow Reports: Monographs, Vol 4 (2018): The Compact Linear Collider (CLIC) – Project Implementation Plan*,. <https://e-publishing.cern.ch/index.php/CYRM/issue/view/68>.
- [153] The CEPC Study Group, *CEPC Conceptual Design Report: Volume I - Accelerator*, [arXiv:1809.00285](https://arxiv.org/abs/1809.00285) [physics.acc-ph].
- [154] CMS Collaboration and Mc Cauley, Thomas, *Collisions recorded by the CMS detector on 14 Oct 2016 during the high pile-up fill*,. <https://cds.cern.ch/record/2231915>. CMS Collection.
- [155] CMS Collaboration, *Technical proposal*. LHC Tech. Proposal. CERN, Geneva, 1994. <http://cds.cern.ch/record/290969>. Cover title : CMS, the Compact Muon Solenoid : technical proposal.
- [156] CMS Collaboration, *CMS physics: Technical design report*, CERN-LHCC-2006-001 (2006).
- [157] I. Neutelings and A. Tsagaropolulos, *Graphics with TikZ in LaTeX*,. https://tikz.net/axis3d_cms/.
- [158] CMS Collaboration, *Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV*, *JINST* **7** (2012) P10002, [arXiv:1206.4071](https://arxiv.org/abs/1206.4071) [physics.ins-det].
- [159] CMS Collaboration, *The CMS magnet project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. <https://cds.cern.ch/record/331056>.
- [160] CMS Collaboration, *Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays*, *JINST* **5** no. [arXiv:0910.5530](https://arxiv.org/abs/0910.5530). CMS-CFT-09-015, (2009) T03021. 35 p. <https://cds.cern.ch/record/1215500>.
- [161] CMS Collaboration, *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. <https://cds.cern.ch/record/368412>.

- [162] CMS Collaboration, *The CMS tracker: addendum to the Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 2000. <https://cds.cern.ch/record/490194>.
- [163] CMS Collaboration, *Precision measurement of the structure of the CMS inner tracking system using nuclear interactions*, JINST **13** no. arXiv:1807.03289. CMS-TRK-17-001-003, (2018) P10034. 41 p. <https://cds.cern.ch/record/2629890>. Submitted to JINST.
- [164] CMS Collaboration, *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical Design Report CMS. CERN, Geneva, 2002. <https://cds.cern.ch/record/578006>.
- [165] CMS Collaboration, *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. <http://cds.cern.ch/record/349375>.
- [166] P. Lecoq et al., *Lead tungstate (PbWO₄) scintillators for LHC EM calorimetry*, Nucl. Instrum. Meth. A **365** (1995) 291–298.
- [167] CMS Collaboration, *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. <http://cds.cern.ch/record/357153>.
- [168] CMS Collaboration, *CMS Technical Design Report for the Phase I Upgrade of the Hadron Calorimeter*,. <https://cds.cern.ch/record/1481837>.
- [169] CMS Collaboration, *The CMS Outer Hadron Calorimeter*,. <https://cds.cern.ch/record/973131>.
- [170] CMS Collaboration, *Design, Performance and Calibration of the CMS Forward Calorimeter Wedges*,. <https://cds.cern.ch/record/951395>.
- [171] CMS Collaboration, *The CMS muon project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. <https://cds.cern.ch/record/343814>.
- [172] CMS Collaboration, *CMS endcap RPC gas gap production for upgrade*, Journal of Instrumentation **7** (2012) P11013. <https://cds.cern.ch/record/1428495>.
- [173] CMS Collaboration, *Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV*, JINST **13** no. CMS-MUO-16-001. CMS-MUO-16-001-003. 06, (2018) P06015. 53 p. <https://cds.cern.ch/record/2313130>.
- [174] CMS Collaboration, *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*. Technical Design Report CMS. CERN, 2000. <https://cds.cern.ch/record/706847>.
- [175] CMS Collaboration, *The CMS trigger system*, JINST **12** no. 01, (2017) P01020, arXiv:1609.02366 [physics.ins-det].

- [176] L. Tuura, G. Eulisse, and A. Meyer, *CMS data quality monitoring web service*, *J. Phys.: Conf. Ser.* **219** (2010) 072055. 6 p. <https://cds.cern.ch/record/1353200>.
- [177] Morovic, Srečko et al., *The New CMS DAQ System for Run 2 of the LHC*,
<https://cds.cern.ch/record/1711011>.
- [178] *Worldwide LHC Computing Grid*, <https://cds.cern.ch/record/1997398>.
<https://wlcg-public.web.cern.ch/>.
- [179] CMS Collaboration, *Performance of the CMS Level-1 Trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV*, <https://cds.cern.ch/record/2702897>.
- [180] CMS Collaboration, *CMS Technical Design Report for the Level-1 Trigger Upgrade*,
<https://cds.cern.ch/record/1556311>.
- [181] CMS Collaboration, A. Bocci, V. Innocente, M. Kortelainen, F. Pantaleo, and M. Rovere, *Heterogeneous Reconstruction of Tracks and Primary Vertices With the CMS Pixel Tracker*, *Frontiers in Big Data* **3** (2020).
<https://www.frontiersin.org/articles/10.3389/fdata.2020.601728>.
- [182] CMS Collaboration, D. Guest, J. Collado, P. Baldi, S.-C. Hsu, G. Urban, and D. Whiteson, *Jet flavor classification in high-energy physics with deep neural networks*, *Phys. Rev. D* **94** (2016) 112002. <https://link.aps.org/doi/10.1103/PhysRevD.94.112002>.
- [183] CMS Collaboration, E. Bols, J. Kieseler, M. Verzetti, M. Stoye, and A. Stakia, *Jet flavour classification using DeepJet*, *Journal of Instrumentation* **15** no. 12, (2020) P12012.
<https://dx.doi.org/10.1088/1748-0221/15/12/P12012>.
- [184] CMS Collaboration, H. Qu and L. Gouskos, *Jet tagging via particle clouds*, *Physical Review D* **101** no. 5, (2020). <https://doi.org/10.1103/PhysRevD.101.056019>.
- [185] CMS Collaboration, *HLT Dimuon Invariant Mass Distributions In 2017 and 2018*,
<https://cds.cern.ch/record/2644377>.
- [186] CMS Collaboration, *Performance of various triggers used in the Supersymmetry group in early 2018 data*, <https://cds.cern.ch/record/2631527>.
- [187] CMS Collaboration, *Early 2018 High-Level Trigger rates*,
<https://cds.cern.ch/record/2644379>.
- [188] CMS Collaboration, *Summary of performance plots for all Run 2 Data*,
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/HighLevelTriggerRunIIResults>.
- [189] CMS Collaboration, *Performance of JetMET high level trigger algorithms in the CMS experiment using proton-proton collisions data at $\sqrt{s} = 13$ TeV during Run-2*,
<http://cds.cern.ch/record/2842377>.

- [190] CMS Collaboration, *Particle-flow reconstruction and global event description with the CMS detector*, *JINST* **12** no. 10, (2017) P10003, [arXiv:1706.04965](https://arxiv.org/abs/1706.04965) [physics.ins-det].
- [191] ALEPH Collaboration, D. Buskulic et al., *Performance of the ALEPH detector at LEP*, *Nucl. Instrum. Meth.* **A360** (1995) 481–506.
- [192] CMS Collaboration, *ECAL Clustering*,
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideEcalRecoClustering>.
Accessed: 07/04/2020.
- [193] CMS Collaboration, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*, *JINST* **9** no. 10, (2014) P10009, [arXiv:1405.6569](https://arxiv.org/abs/1405.6569) [physics.ins-det].
- [194] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, *Nucl. Instrum. Meth.* **A262** (1987) 444–450.
- [195] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, *Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC*, *Journal of Physics G: Nuclear and Particle Physics* **31** no. 9, (2005) N9. <https://dx.doi.org/10.1088/0954-3899/31/9/N01>.
- [196] K. Rose, *Deterministic Annealing for Clustering, Compression, Classification, Regression, and Related Optimization Problems*,.
- [197] R. Frühwirth, W. Waltenberger, and P. Vanlaer, *Adaptive vertex fitting*, *J. Phys.* **G34** (2007) N343.
- [198] CMS Collaboration, *Pileup mitigation at CMS in 13 TeV data*, *Journal of Instrumentation* **15** no. 09, (2020) P09018–P09018.
<https://doi.org/10.1088/1748-0221/15/9/P09018>.
- [199] D. Bertolini, P. Harris, M. Low, and N. Tran, *Pileup per particle identification*, *Journal of High Energy Physics* **2014** no. 10, (2014). [https://doi.org/10.1007/JHEP10\(2014\)29059](https://doi.org/10.1007/JHEP10(2014)29059).
- [200] CMS Collaboration, *Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ TeV*, *JINST* **13** no. 10, (2018) P10005, [arXiv:1809.02816](https://arxiv.org/abs/1809.02816) [hep-ex].
- [201] CMS Collaboration, *Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS*, *JINST* **11** no. 01, (2016) P01019, [arXiv:1510.07488](https://arxiv.org/abs/1510.07488) [physics.ins-det].
- [202] CMS Collaboration, *Performance of tau-lepton reconstruction and identification in CMS*, *JINST* **7** (2012) P01001, [arXiv:1109.6034](https://arxiv.org/abs/1109.6034) [physics.ins-det].
- [203] CMS Collaboration, *Identification of hadronic tau lepton decays using a deep neural network*, *Journal of Instrumentation* **17** no. 07, (2022) P07023.
<https://doi.org/10.1088/1748-0221/17/7/P07023>.

- [204] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti- $k(t)$ jet clustering algorithm*, *JHEP* **04** (2008) 063, [arXiv:0802.1189 \[hep-ph\]](#).
- [205] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896, [arXiv:1111.6097 \[hep-ph\]](#).
- [206] CMS Collaboration, *Jet algorithms performance in 13 TeV data*, <https://cds.cern.ch/record/2256875>.
- [207] CMS Collaboration, *Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV*, *JINST* **12** no. 02, (2017) P02014, [arXiv:1607.03663 \[hep-ex\]](#).
- [208] CMS Collaboration, *Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV*, *Journal of Instrumentation* **13** no. 05, (2018) P05011, 1712.07158. <https://doi.org/10.1088/1748-0221/13/05/P05011>.
- [209] CMS Collaboration, *Identification of b quark jets at the CMS Experiment in the LHC Run 2*, <http://cds.cern.ch/record/2138504>.
- [210] CMS Collaboration, *Identification of b-quark jets with the CMS experiment*, *JINST* **8** (2013) P04013, [arXiv:1211.4462 \[hep-ex\]](#).
- [211] CMS Collaboration, *Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV*, *Journal of Instrumentation* **10** no. 06, (2015) P06005, [arXiv:1502.02701 \[physics.ins-det\]](#).
- [212] A. Buckley et al., *General-purpose event generators for LHC physics*, *Phys. Rept.* **504** (2011) 145–233, [arXiv:1101.2599 \[hep-ph\]](#).
- [213] F. Siegert, *Monte-Carlo event generation for the LHC*. PhD thesis, 2010. <https://cds.cern.ch/record/1600005>.
- [214] Collins, John C. and Soper, Davison E., *The Theorems of Perturbative QCD*, *Ann. Rev. Nucl. Part. Sci.* **37** (1987) 383–409.
- [215] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, *Eur. Phys. J. C* **63** (2009) 189, [arXiv:0901.0002 \[hep-ph\]](#).
- [216] NNPDF Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040, [arXiv:1410.8849 \[hep-ph\]](#).
- [217] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [arXiv:1405.0301 \[hep-ph\]](#).

- [218] S. Alioli, P. Nason, C. Oleari, and E. Re, *NLO single-top production matched with shower in POWHEG: s- and t-channel contributions*, *JHEP* **09** (2009) 133, [arXiv:0907.4076](https://arxiv.org/abs/0907.4076) [hep-ph].
[Erratum: [doi: 10.1007/JHEP02\(2010\)011](https://doi.org/10.1007/JHEP02(2010)011)].
- [219] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, *Eur. Phys. J. C* **71** (2011) 1547, [arXiv:1009.2450](https://arxiv.org/abs/1009.2450) [hep-ph].
- [220] J. Winter, S. Hoeche, H. Hoeth, F. Krauss, M. Schonherr, K. Zapp, S. Schumann, and F. Siegert, *Systematic improvement of QCD parton showers*, *Frascati Phys. Ser.* **54** (2012) 336–355, [arXiv:1204.5668](https://arxiv.org/abs/1204.5668) [hep-ph].
- [221] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026, [arXiv:hep-ph/0603175](https://arxiv.org/abs/hep-ph/0603175) [hep-ph].
- [222] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852, [arXiv:0710.3820](https://arxiv.org/abs/0710.3820) [hep-ph].
- [223] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A simulation toolkit*, *Nucl. Instrum. Meth.* **A506** (2003) 250–303.
- [224] CMS Collaboration, S. Abdullin, P. Azzi, F. Beaudette, P. Janot, and A. Perrotta, *The fast simulation of the CMS Detector at LHC*, *J. Phys. Conf. Ser.* **331** (2011) 032049.
- [225] J. Alwall, P. Schuster, and N. Toro, *Simplified Models for a First Characterization of New Physics at the LHC*, *Physical Review D* **79** (2009) 075020, [arXiv:0810.3921](https://arxiv.org/abs/0810.3921) [hep-ph].
- [226] LHC New Physics Working Group Collaboration, *Simplified Models for LHC New Physics Searches*, *Journal of Physics G* **39** (2012) 105005, [arXiv:1105.2838](https://arxiv.org/abs/1105.2838) [hep-ph].
- [227] CMS Collaboration, *Search for top squarks decaying via four-body or chargino-mediated modes in single-lepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* (2018), [arXiv:1805.05784](https://arxiv.org/abs/1805.05784) [hep-ex].
- [228] W. Beenakker, R. Höpker, M. Spira, and P. M. Zerwas, *Squark and gluino production at hadron colliders*, *Nucl. Phys. B* **492** (1997) 51, [arXiv:hep-ph/9610490](https://arxiv.org/abs/hep-ph/9610490) [hep-ph].
- [229] W. Beenakker, M. Krämer, T. Plehn, M. Spira, and P. Zerwas, *Stop production at hadron colliders*, *Nuclear Physics B* **515** no. 1-2, (1998) 3–14.
<https://doi.org/10.1016%2Fs0550-3213%2898%2900014-5>.
- [230] W. Beenakker, C. Borschensky, M. Krämer, A. Kulesza, and E. Laenen, *NNLL-fast: predictions for coloured supersymmetric particle production at the LHC with threshold and Coulomb resummation*, *Journal of High Energy Physics* **2016** no. 12, (2016).
<https://doi.org/10.1007%2Fjhep12%282016%29133>.

- [231] LHC SUSY Cross Section Working Group, *SUSY Cross Sections for 13, 14, 33 and 100 TeV pp Collisions*,.
- [232] M. Kramer, A. Kulesza, R. van der Leeuw, M. Mangano, S. Padhi, T. Plehn, and X. Portell, *Supersymmetry production cross sections in pp collisions at $\sqrt{s} = 7$ TeV*, [arXiv:1206.2892 \[hep-ph\]](#).
- [233] W. Beenakker, R. Höpker, and M. Spira, *PROSPINO: A Program for the production of supersymmetric particles in next-to-leading order QCD*, [arXiv:hep-ph/9611232 \[hep-ph\]](#).
- [234] A. Kulesza and L. Motyka, *Threshold resummation for squark-antisquark and gluino-pair production at the LHC*, *Phys. Rev. Lett.* **102** (2009) 111802, [arXiv:hep-ph/0807.2405 \[hep-ph\]](#).
- [235] A. Kulesza and L. Motyka, *Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC*, *Phys. Rev. D* **80** (2009) 095004, [arXiv:hep-ph/0905.4749 \[hep-ph\]](#).
- [236] W. Beenakker, S. Brensing, M. Krämer, A. Kulesza, E. Laenen, and I. Niessen, *Soft-gluon resummation for squark and gluino hadroproduction*, *JHEP* **12** (2009) 41, [arXiv:hep-ph/0909.4418 \[hep-ph\]](#).
- [237] W. Beenakker, S. Brensing, M. Krämer, A. Kulesza, E. Laenen, L. Motyka, and I. Niessen, *Squark and gluino production*, *Int. J. Mod. Phys. A* **26** (2011) 2637, [arXiv:hep-ph/1105.1110 \[hep-ph\]](#).
- [238] W. Beenakker, M. Klasen, M. Krämer, T. Plehn, M. Spira, and P. M. Zerwas, *Production of Charginos, Neutralinos, and Stopped Squarks at Hadron Colliders*, *Phys. Rev. Lett.* **83** (1999) 3780–3783. <https://link.aps.org/doi/10.1103/PhysRevLett.83.3780>.
- [239] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, *Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV*, *JHEP* **10** (2012) 081, [arXiv:1207.2159 \[hep-ph\]](#).
- [240] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, *Precision predictions for electroweak superpartner production at hadron colliders with RESUMMINO*, *Eur. Phys. J. C* **73** (2013) 2480, [arXiv:1304.0790 \[hep-ph\]](#).
- [241] ATLAS Collaboration, *Search for top squark pair production in final states with one isolated lepton, jets, and missing transverse momentum in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector*, *Journal of High Energy Physics* **2014** no. 11, (2014). [https://doi.org/10.1007/JHEP11\(2014\)29118](https://doi.org/10.1007/JHEP11(2014)29118).

- [242] CMS Collaboration, *Search for direct production of supersymmetric partners of the top quark in the all-jets final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Journal of High Energy Physics* **10** no. 10, (2017) 005, 1707.03316. [https://dx.doi.org/10.1007/JHEP10\(2017\)005](https://dx.doi.org/10.1007/JHEP10(2017)005).
- [243] CMS Collaboration, *Search for top squark production in fully hadronic final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Physical Review D* **104** no. 5, (2021). <https://doi.org/10.1103/PhysRevD.104.052001>.
- [244] CMS Collaboration, *Search for new physics in events with two soft oppositely charged leptons and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Physics Letters B* **782** (2018) 440 – 467, 1801.01846. <https://www.sciencedirect.com/science/article/pii/S037026931830426X>.
- [245] CMS Collaboration, *Search for supersymmetry in final states with two or three soft leptons and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Journal of High Energy Physics* no. 91, (2022), 2111.06296. [https://doi.org/10.1007/JHEP04\(2022\)091](https://doi.org/10.1007/JHEP04(2022)091).
- [246] R. Gröber, M. M. Mühlleitner, E. Popenza, and A. Wlotzka, *Light Stop Decays: Implications for LHC Searches*, *The European Physical Journal C* **75** no. 9, (2015) 420, 1408.4662. <https://doi.org/10.1140/epjc/s10052-015-3626-z>.
- [247] CMS Collaboration, *Search for the pair production of third-generation squarks with two-body decays to a bottom or charm quark and a neutralino in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Physics Letters B* **778** (2018) 263 – 291, 1707.07274. <https://www.sciencedirect.com/science/article/pii/S0370269318300200>.
- [248] CMS Collaboration, *Displays of candidate events with soft leptons and missing energy, as examined in Higgsino searches*,. <https://cds.cern.ch/record/2759686>. CMS Collection.
- [249] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, PoS ACAT (2007) 040, [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039).
- [250] L. Rokach and O. Maimon, *Data mining with decision trees: theory and applications*. World Scientific Pub Co Inc., 2008.
- [251] CMS Collaboration, *Search for top squarks in the four-body decay mode with single lepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, [arXiv:2301.08096](https://arxiv.org/abs/2301.08096). <https://cds.cern.ch/record/2846517>.
- [252] N. K. Rad, *Search for supersymmetry partners of the top quark in models with compressed mass spectra with the CMS detector*. PhD thesis, Technische Universität Wien, Wien, 2018.
- [253] Brun, R. and Rademakers, F., *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth. A* **389** (1997) 81–86.

- [254] CMS Collaboration, C. D. Jones et al., *The New CMS Event Data Model and Framework*, CHEP Proc. (2006) .
- [255] CMS Collaboration, P. Elmer, B. Hegner, and L. Sexton-Kennedy, *Experience with the CMS event data model*, *Journal of Physics: Conference Series* **219** no. 3, (2010) 032022.
<https://dx.doi.org/10.1088/1742-6596/219/3/032022>.
- [256] CMS Collaboration, G. Petrucciani and A. Rizzi and C. Vuosalo, *Mini-AOD: A New Analysis Data Format for CMS*, *Journal of Physics: Conference Series* **664** no. 7, (2015) 072052.
<https://doi.org/10.1088/1742-6596/664/7/072052>.
- [257] CMS Collaboration, M. Peruzzi, G. Petrucciani, A. Rizzi, and for the CMS Collaboration, *The NanoAOD event data format in CMS*, *Journal of Physics: Conference Series* **1525** no. 1, (2020) 012038. <https://dx.doi.org/10.1088/1742-6596/1525/1/012038>.
- [258] CMS Collaboration, *Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector*, *Journal of Instrumentation* **14** no. 07, (2019) P07004–P07004.
<https://doi.org/10.1088/1748-0221/14/07/P07004>.
- [259] J. Alwall et al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, *Eur. Phys. J. C* **53** (2008) 473–500, arXiv:0706.2569 [hep-ph].
- [260] M. Perelstein and A. Weiler, *Polarized tops from stop decays at the LHC*, *Journal of High Energy Physics* **2009** no. 03, (2009) 141–141.
<https://doi.org/10.1088/1126-6708/2009/03/141>.
- [261] I. Mikulec, *MC Production of 4-body Stop Decay*,. Private Communication.
- [262] CMS Collaboration, *Measuring Electron Efficiencies at CMS with Early Data*,.
<https://cds.cern.ch/record/1194482>.
- [263] I. Mikulec, *Lepton efficiencies for identification, isolation and impact parameter requirements using the tag-and-probe method*,. Private Communication.
- [264] I. Mikulec, *Triggers for soft lepton analysis*,. Private Communication.
- [265] Particle Data Group Collaboration, R. L. Workman et al., *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.
- [266] CMS Collaboration, *Measurements of the W boson rapidity, helicity, double-differential cross sections, and charge asymmetry in pp collisions at $\sqrt{s} = 13$ TeV*, *Physical Review D* **102** no. 9, (2020). <https://doi.org/10.1103/PhysRevD.102.092012>.

- [267] CMS Collaboration, *Search for new physics in same-sign dilepton events in proton–proton collisions at $s = 13$ TeV*, *The European Physical Journal C* **76** no. 8, (2016).
<https://doi.org/10.1140/epjc%2Fs10052-016-4261-z>.
- [268] J. W. Andrejkovič, *Data-Driven Background Modeling for Precision Studies of the Higgs Boson and Searches for New Physics with the CMS Experiment*. PhD thesis, Technische Universität Wien, Wien, 2022.
- [269] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) 1554, [arXiv:1007.1727](https://arxiv.org/abs/1007.1727) [physics.data-an]. [Erratum: doi: 10.1140/epjc/s10052-013-2501-z].
- [270] T. Junk, *Confidence level computation for combining searches with small statistics*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **434** no. 2-3, (1999) 435–443.
<https://doi.org/10.1016%2Fs0168-9002%2899%2900498-2>.
- [271] A. L. Read, *Presentation of search results: the CLs technique*, *Journal of Physics G: Nuclear and Particle Physics* **28** no. 10, (2002) 2693.
<https://dx.doi.org/10.1088/0954-3899/28/10/313>.
- [272] ATLAS, CMS, LHC Higgs Combination Group Collaboration, *Procedure for the LHC Higgs boson search combination in Summer 2011*, <https://cds.cern.ch/record/1379837>.
- [273] G. F. Giudice, T. Han, K. Wang, and L.-T. Wang, *Nearly Degenerate Gauginos and Dark Matter at the LHC*, *Physical Review D* **81** no. 11, (2010) 115011, 1004.4902 [hep-ph].
<http://dx.doi.org/10.1103/PhysRevD.81.115011>.
- [274] A. Mustafayev and X. Tata, *Supersymmetry, Naturalness, and Light Higgsinos*, *Indian Journal of Physics* **88** (2014) 991–1004, [arXiv:1404.1386](https://arxiv.org/abs/1404.1386) [hep-ph].
<http://dx.doi.org/10.1007/s12648-014-0504-8>.
- [275] H. Baer, V. Barger, and P. Huang, *Hidden SUSY at the LHC: the light higgsino-world scenario and the role of a lepton collider*, *Journal of High Energy Physics* **2011** no. 11, (2011) 31.
[https://doi.org/10.1007/JHEP11\(2011\)031](https://doi.org/10.1007/JHEP11(2011)031).
- [276] MSSM Working Group Collaboration, *The Minimal Supersymmetric Standard Model: Group Summary Report*, in *GDR (Groupement de Recherche) - Supersymétrie*. 1998. [hep-ph/9901246](https://cds.cern.ch/record/376049).
<https://cds.cern.ch/record/376049>. 1998.
- [277] CMS Collaboration, *Muon Reconstruction and Identification Performance with Run–2 Data*,
<https://cds.cern.ch/record/2727091>.

- [278] CMS Collaboration, *Level-1 muon trigger performance in 2017 data and comparison with the legacy muon trigger system*, <https://cds.cern.ch/record/2286327>.
- [279] Geoffrey N Smith and Charles N Mueller and Andrew S Wightman and on behalf of CMS, *Tools for Trigger Rate Monitoring at CMS*, *Journal of Physics: Conference Series* **1085** no. 4, (2018) 042032. <https://dx.doi.org/10.1088/1742-6596/1085/4/042032>.
- [280] P. Katsoulis, *Development of triggering systems and searches for signatures from theories beyond the standard model with the CMS detector at the LHC*, <https://olympias.lib.uoi.gr/jspui/handle/123456789/32271>.
- [281] C. J. Alimena, J. Alison, C. Botta, **C. Fountas**, Z. Gecse, S. Ghosh, J. Heikkilä, C. Herwig, **G. Karathanasis**, **P. Katsoulis**, P. E. Meiring, I. Papavergou, M. Peruzzi, P. Sphicas, W. Terrill, K. Vellidis, and E. Vourliotis, *Single-Muon Trigger Performance in Low MET Region*,. Private Communication.
- [282] CMS Collaboration, Dinyar Rabady et al., *A 40 MHz Level-1 trigger scouting system for the CMS Phase-2 upgrade*, *Nucl. Instrum. Meth. A* **1047** (2023) 167805.
- [283] CMS Collaboration, *40 MHz Scouting with Deep Learning in CMS*, <https://cds.cern.ch/record/2843741>.
- [284] CMS Collaboration, J. Pata, J. Duarte, F. Mokhtar, E. Wulff, J. Yoo, J.-R. Vlimant, M. Pierini, M. Girone, and on behalf of the CMS Collaboration, *Machine Learning for Particle Flow Reconstruction at CMS*, *Journal of Physics: Conference Series* **2438** no. 1, (2023) 012100. <https://dx.doi.org/10.1088/1742-6596/2438/1/012100>.
- [285] R. T. d'Agnolo, G. Grosso, M. Pierini, A. Wulzer, and M. Zanetti, *Learning New Physics from an Imperfect Machine*, *The European Physical Journal C* **82** no. 3, (2022) 275. <https://doi.org/10.1140/epjc/s10052-022-10226-y>.
- [286] CMS Collaboration, A. A. Pol, V. Azzolini, G. Cerminara, F. De Guio, G. Franzoni, M. Pierini, F. Široký, and J.-R. Vlimant, *Anomaly detection using Deep Autoencoders for the assessment of the quality of the data acquired by the CMS experiment*, <https://cds.cern.ch/record/2650715>.
- [287] Ekaterina Govorkova et al., *Autoencoders on field-programmable gate arrays for real-time, unsupervised new physics detection at 40 MHz at the Large Hadron Collider*, *Nature Machine Intelligence* **4** no. 2, (2022) 154–161. <https://doi.org/10.1038/s42256-022-00441-3>.
- [288] CMS Collaboration, *Recording and reconstructing 10 billion unbiased b hadron decays in CMS*, <https://cds.cern.ch/record/2704495>.

- [289] Bainbridge, Robert, *Recording and reconstructing 10 billion unbiased b hadron decays in CMS*, *EPJ Web Conf.* **245** (2020) 01025. <https://doi.org/10.1051/epjconf/202024501025>.
- [290] ATLAS Collaboration, *Identification of very-low transverse momentum muons in the ATLAS experiment*,. <https://cds.cern.ch/record/2710574>.
- [291] ATLAS Collaboration, *Searches for electroweak production of supersymmetric particles with compressed mass spectra in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, *Physical Review D* **101** (2020) 052005, 1911.12606. <https://link.aps.org/doi/10.1103/PhysRevD.101.052005>.
- [292] CMS Collaboration, *Search for supersymmetry with a compressed mass spectrum in the vector boson fusion topology with l -lepton and 0 -lepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Journal of High Energy Physics* **08** no. 8, (2019) 150, 1905.13059. [https://doi.org/10.1007/JHEP08\(2019\)150](https://doi.org/10.1007/JHEP08(2019)150).
- [293] CMS Collaboration, *Search for Supersymmetry with a Compressed Mass Spectrum in Events with a Soft τ Lepton, a Highly Energetic Jet, and Large Missing Transverse Momentum in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV*, *Physical Review Letters* **124** (2020) 041803, 1910.01185. <https://link.aps.org/doi/10.1103/PhysRevLett.124.041803>.
- [294] ATLAS Collaboration, *Search for chargino-neutralino pair production in final states with three leptons and missing transverse momentum in $\sqrt{s} = 13$ TeV p - p collisions with the ATLAS detector*,. <https://cds.cern.ch/record/2719521>.
- [295] ATLAS Collaboration, *Search for a scalar partner of the top quark in the all-hadronic $t\bar{t}$ plus missing transverse momentum final state at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *The European Physical Journal C* **80** no. 8, (2020) 737, 2004.14060. <https://dx.doi.org/10.1140/epjc/s10052-020-8102-8>.
- [296] ATLAS Collaboration, *Search for new phenomena with top quark pairs in final states with one lepton, jets, and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*,. <https://cds.cern.ch/record/2711489>.
- [297] CMS Collaboration, *Search for direct top squark pair production in events with one lepton, jets, and missing transverse momentum at 13 TeV with the CMS experiment*, *Journal of High Energy Physics* **05** no. 5, (2020) 032, 1912.08887. [https://dx.doi.org/10.1007/JHEP05\(2020\)032](https://dx.doi.org/10.1007/JHEP05(2020)032).
- [298] ATLAS Collaboration, *Soft b -hadron tagging for compressed SUSY scenarios*,. <https://cds.cern.ch/record/2682131>.

- [299] CMS Collaboration, *Measurement of $B\bar{B}$ angular correlations based on secondary vertex reconstruction at $\sqrt{s} = 7$ TeV*, *Journal of High Energy Physics* **03** no. 3, (2011) 136, 1102.3194. [https://dx.doi.org/10.1007/JHEP03\(2011\)136](https://dx.doi.org/10.1007/JHEP03(2011)136).
- [300] A. Ismail, R. Schwienhorst, J. S. Virzi, and D. G. Walker, *Deconstructed transverse mass variables*, *Physical Review D* **91** no. 7, (2015). <https://doi.org/10.1103/PhysRevD.91.074002>.
- [301] P. Jackson and C. Rogan, *Recursive Jigsaw Reconstruction: HEP event analysis in the presence of kinematic and combinatoric ambiguities*, *Physical Review D* **96** (2017) 112007, 1705.10733. <https://link.aps.org/doi/10.1103/PhysRevD.96.112007>.
- [302] P. Jackson, C. Rogan, and M. Santoni, *Sparticles in Motion: Analyzing compressed SUSY scenarios with a new method of event reconstruction*, *Physical Review D* **95** no. 3, (2017) 035031, [arXiv:1607.08307](https://arxiv.org/abs/1607.08307) [hep-ph].
- [303] CMS Collaboration, *Performance of quark/gluon discrimination in 8 TeV pp data*,. <https://cds.cern.ch/record/1599732>.
- [304] ATLAS Collaboration, *Discrimination of Light Quark and Gluon Jets in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*,. <https://cds.cern.ch/record/2200202>.
- [305] CMS Collaboration, *Performance of quark/gluon discrimination in 13 TeV data*,. <https://cds.cern.ch/record/2234117>.
- [306] Juliette Alimena et. al, *Searching for long-lived particles beyond the Standard Model at the Large Hadron Collider*, *Journal of Physics G: Nuclear and Particle Physics* **47** no. 9, (2020) 090501. <https://doi.org/10.1088/1361-6471/2020/9/090501>.
- [307] A. De Roeck, *Searching for long-lived particles at the Large Hadron Collider and beyond*, *Philos. Trans. R. Soc. Lond. A* **377** no. 2161, (2019) 20190047. <https://cds.cern.ch/record/2827283>.
- [308] CMS Collaboration, *Run 2 Summary Plots (13 TeV)*,. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS>.
- [309] ATLAS Collaboration, *SUSY March 2023 Summary Plot Update*,. <https://cds.cern.ch/record/2852738>. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2023-005>.
- [310] CMS Collaboration, *Searches for light higgsino-like charginos and neutralinos at the HL-LHC with the Phase-2 CMS detector*,. <https://cds.cern.ch/record/2648538>.
- [311] CMS Collaboration, *Combined search for supersymmetry with photons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Physics Letters B* **801** (2020) 135183. <https://doi.org/10.1016/j.physletb.2019.135183>.

- [312] J. Fan, M. Reece, and J. T. Ruderman, *Stealth supersymmetry*, *Journal of High Energy Physics* **2011** no. 11, (2011). <https://doi.org/10.1007%2Fjhep11%282011%29012>.
- [313] CMS Collaboration, *Search for top squarks in final states with two top quarks and several light-flavor jets in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Physical Review D* **104** no. 3, (2021). <https://doi.org/10.1103%2Fphysrevd.104.032006>.
- [314] G. Giudice and A. Romanino, *Split supersymmetry*, *Nuclear Physics B* **699** no. 1, (2004) 65–89. <https://www.sciencedirect.com/science/article/pii/S0550321304005759>.
- [315] Wolfgang Waltenberger et al., *Reinterpretation of LHC Results for New Physics: Status and recommendations after Run 2*, *SciPost Physics* **9** no. 2, (2020). <https://doi.org/10.21468%2Fscipostphys.9.2.022>.
- [316] Federico Ambrogio et al., *SModelS v1.2: Long-lived particles, combination of signal regions, and other novelties*, *Computer Physics Communications* **251** (2020) 106848. <https://www.sciencedirect.com/science/article/pii/S0010465519302255>.
- [317] S. Weinberg, *Effective gauge theories*, *Physics Letters B* **91** no. 1, (1980) 51–55. <https://www.sciencedirect.com/science/article/pii/0370269380906607>.
- [318] I. Brivio, S. Bruggisser, F. Maltoni, R. Moutafis, T. Plehn, E. Vryonidou, S. Westhoff, and C. Zhang, *O new physics, where art thou? A global search in the top sector*, *Journal of High Energy Physics* **2020** no. 2, (2020). <https://doi.org/10.1007%2Fjhep02%282020%29131>.
- [319] CMS Collaboration, CMS μ GT Trigger Group (HEPHY), *The Level-1 Global Trigger for the CMS Experiment at LHC*,. <https://globaltrigger.web.cern.ch/>.
- [320] CMS Collaboration, *Scales for inputs to μ GT (ϕ , η , p_T , E_T , and others)*,. CMS-DN-2014-029.
- [321] G Bauer et al., *High level trigger configuration and handling of trigger tables in the CMS filter farm*, *Journal of Physics: Conference Series* **119** no. 2, (2008) 022011. <https://dx.doi.org/10.1088/1742-6596/119/2/022011>.

Mateusz Zarucki

Curriculum Vitae

☎ (+43) 660-465-9440

✉ zaruckim@proton.me

📄 LinkedIn: mateusz-zarucki-270662147

Young physicist motivated to play an active role in shaping the future of experimental particle physics.

Work Experience

- 2021–Ongoing **Research Fellow**, University of Notre Dame (US), CERN, Geneva, Switzerland.
Research Fellow position at the University of Notre Dame (US), working as a member of the CMS Collaboration at the European Organisation for Nuclear Research (CERN).
- Convener (L2) of the Field Operations Group (FOG) of the Trigger Studies Group (TSG), responsible for leading a team of up to ≈ 20 members.
 - CMS data-taking operations of the High Level Trigger (HLT) system during the LHC *Long Shutdown 2* (LS2) and successful startup of LHC *Run 3*, including commissioning of heterogenous HLT reconstruction on GPUs.
 - Responsible for training of Detector on Call (DOC) experts and regular organisation of shifts.
 - Software development of the HLT Online Test Stand (Hilton) and Rate Monitoring (RateMon) in the context of AutoDevOps: CI/CD, containerisation and cloud deployments (OpenShift/OpenStack/PaaS).
 - Awarded the 2021 CMS Award "for his exemplary leadership of the TSG Field Operations Group and outstanding support of online operations."
- 2015–2020 **PhD Researcher**, Institute of High Energy Physics (HEPHY), Vienna, Austria.
PhD position at the Institute of High Energy Physics (HEPHY) Vienna at the Technical University (TU) Vienna, as part of the Doktoratskolleg Particles and Interactions (DKPI) doctoral programme. Member of the CMS Collaboration at the European Organisation for Nuclear Research (CERN).
- Analysis of data collected by the Compact Muon Solenoid (CMS) detector at the Large Hadron Collider (LHC) in the context of Supersymmetry (SUSY) searches.
 - Level-1 (L1) trigger menu coordinator role (L3) responsible for delivering the full set of L1 triggers that pre-select the type of data, encapsulating the physics programme of CMS.
 - Data acquisition (DAQ) shifter responsible for controlling the DAQ system in the CMS control room and monitoring the collection of data.
 - Data quality monitoring (DQM) software developer for the L1 Global Trigger Upgrade.
- Jan.–Apr. 2015 **IAEA Internship Programme**, International Atomic Energy Agency, Vienna, Austria.
Internship at the NAPC Physics Section, with a specific focus on nuclear fusion.
- Jul.–Sep. 2014 **IIASS Summer Internship Programme**, International Institute of Advanced Academic Studies "E. R. Caianiello"/University of Salerno, Vietri sul Mare/Fisciano, Italy.
- Controlling and maintaining the university observatory for collecting astronomical image data for microlensing surveys, including extrasolar planet searches.
 - Completed project on "Parallax in Microlensing Searches towards the Galactic Bulge", which modelled possible observable data from the NASA Spitzer Space Telescope.

Education

2015–2020, 2022–2023 **Vienna University of Technology (TU Wien)**

Doctorate in Natural Sciences (Dr. rer. nat.)

Thesis: *Data-driven background modelling and trigger algorithms for compressed supersymmetry searches with the CMS experiment at the LHC*

Description: Analysis on the search for Supersymmetric (SUSY) particles in models with highly compressed mass spectra in the single-lepton channel, performed on a sample of 35.9 fb^{-1} proton-proton data from *Run 2* of the LHC, recorded with the CMS detector.

The thesis focuses on several data-driven background modelling methods of nonprompt leptons using Monte Carlo simulation together with normalisation to data. Another focus of the thesis are improvements of the trigger strategy to extend the signal acceptance, including electroweakino signals. Dedicated trigger algorithms were developed, exploiting the typical ISR-boosted signature, in order to achieve lower missing momentum thresholds. The algorithms were used online during year 2018 of LHC *Run 2* and are currently collecting data during the ongoing LHC *Run 3*.

2010–2014 **Imperial College London**

Master in Science (MSci) & Associateship of the Royal College of Science (ARCS)
Physics with a Year in Europe | Second Class Honours (Upper Division)

[Master in Science \(MSci.\)](#)

Thesis: *Calibration of the Hamamatsu R11410-10 PMTs for the NEXT-100 TPC Detector*

Description: The thesis focuses on the energy calibration of 55 photomultiplier devices (PMTs) and an evaluation of the contribution to the total energy resolution of the detector.

2012–2013 **University of Valencia (Universitat de València)**

3rd academic year in the form of an exchange year in Valencia, Spain, under the Physics with a Year in Europe and Erasmus programmes.

Formed part of the NEXT research group at IFIC (CSIC).

1996–2009 **Vienna International School**

International Baccalaureate Diploma awarded together with the Austrian Matura.

Publications

- 2020 *Soft Supersymmetry Searches at ATLAS and CMS (PoS LHCP2020 009)*
- 2020 *Performance of a Soft Muon, Hard Jet and Moderate Missing Energy Trigger in 2018 Data (CMS DP-2020 004)*
- 2018 *Search for top squarks decaying via four-body or chargino-mediated modes in single-lepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV (JHEP 1809 065)*
- 2018 *Search for SUSY with a Highly Compressed Mass Spectrum in the Soft Single Lepton Channel with the CMS Experiment (PoS ALPS2018 058)*
- 2017 *Search for Supersymmetry with a Highly Compressed Mass Spectrum in the Single Soft Lepton Channel with the CMS Experiment at the LHC (PoS EPS - HEP2017 728)*
- 2015 *Pathway to the Galactic Distribution of Planets: Combined Spitzer and Ground-Based Microlens Parallax Measurements of 21 Single-Lens Events (ApJ, 804, 20)*

Skills

- Fluent in Python and C++ object-oriented programming (OOP).
- Proficient with CERN's ROOT data-analysis framework and CMS Software (CMSSW).
- Proficient in use of the GitHub/GitLab software development platform.
- Experienced with AutoDevOps: CI/CD software development, containerisation and cloud deployments to OpenShift/OpenStack/PaaS.
- Experienced with Toolkit for Multivariate Data Analysis (TMVA) for Boosted Decision Trees (BDT) analysis, including training, validation and application of classifiers.

Languages

- **Polish:** Mother Tongue
- **English:** Fluent (C2 according to CEFR)
- **German:** Fluent (C1 according to CEFR)
- **Spanish:** Fluent (C1 according to CEFR)
- **Italian:** Beginner (A2 according to CEFR)