Search for Natural Supersymmetry with Low-Momentum and Displaced Tracks

and

Machine Learning-Based Refinement of Simulations

Dissertation
zur Erlangung des Doktorgrades
an der Fakultät für Mathematik, Informatik und
Naturwissenschaften
Fachbereich Physik
der Universität Hamburg

vorgelegt von
Moritz Jonas Wolf

Hamburg
2025

Gutachter der Dissertation: Prof. Dr. Peter Schleper
Prof. Dr. Gregor Kasieczka

Zusammensetzung der Prüfungskommission: Prof. Dr. Daniela Pfannkuche
Prof. Dr. Peter Schleper
Prof. Dr. Gregor Kasieczka
Dr. Isabell Melzer-Pellmann
Dr. Andreas Hinzmann

Vorsitzende der Prüfungskommission: Prof. Dr. Daniela Pfannkuche

Datum der Disputation: 19.05.2025

Vorsitzender des Fach-Promotionsausschusses PHYSIK: Prof. Dr. Wolfgang J. Parak

Leiter des Fachbereichs PHYSIK: Prof. Dr. Markus Drescher

Dekan der Fakultät MIN: Prof. Dr.-Ing. Norbert Ritter

What is fully, completely understood
leaves no trace as memory.

Jiddu Krishnamurti

Abstract

The most successful theory for explaining the matter and forces in our universe is the Standard Model of particle physics (SM). Despite its success, it is known that the SM does not and cannot provide a complete description of all fundamental interactions and observed physics phenomena. Experiments at the CERN Large Hadron Collider (LHC) are at the forefront of worldwide efforts to understand Nature at the smallest scales and highest energies and to pin down the laws and dynamics that lie beyond the SM.

This thesis presents two projects that advance the search for new physics in different ways.
In the first part of this thesis, a machine learning-based morphing approach is introduced which has been developed to facilitate particle physics analyses in a broad context by refining simulations to improve their accuracy.
The second part describes a search for signs of extensions to the SM that are based on the concept of supersymmetry (SUSY). Such manifestations of physics beyond the SM are motivated by the fact that they address numerous theoretical issues of the SM and can provide a dark matter candidate. Specifically, the presented analysis targets so-called *natural* SUSY scenarios with low-mass higgsinos, exhibiting compressed mass spectra.

The refinement tool and its supporting studies demonstrate that classical fast simulation applications can be augmented to achieve an accuracy very near that of more detailed, high-fidelity, but compute-intensive simulation engines. This opens up the prospect for much more efficient and granular studies of new physics models in the future CMS physics program.
The analysis presented in the second part of this thesis establishes sensitivity to some of the last viable phase space regions of natural SUSY, testing for higgsino-like dark matter candidates with masses up to 180 GeV. The data, recorded by the CMS experiment at the LHC, are found to be consistent with the predictions of the SM, and no evidence of new physics is found. Higgsinos are excluded for a range of mass splittings between the lightest chargino and neutralino of 0.3 to 1.2 GeV corresponding to cases which were unconstrained by previous searches.

Zusammenfassung

Das Standardmodell der Teilchenphysik (SM) ist das Modell, welches das gesammelte Wissen über die fundamentalen Bausteine und Kräfte unseres Universums am erfolgreichsten zusammenfasst. Trotzdem ist bekannt, dass es nur als Teil einer noch unbekannten, umfassenderen Theorie verstanden werden kann. Die Experimente am CERN Large Hadron Collider (LHC) nehmen eine weltweit führende Rolle darin ein, die Natur in kleinsten Maßstäben und bei höchsten Energien zu erforschen und so den Gesetzen und Dynamiken jenseits des SM näher zu kommen.

In dieser Arbeit werden zwei Projekte vorgestellt, die auf unterschiedliche Weise die Suche nach neuer Physik voranbringen.
Im ersten Teil der Arbeit wird eine *refinement*-Methode vorgestellt, die auf maschinellem Lernen basiert und entwickelt wurde, um die Genauigkeit von Simulationen zu verbessern. Im übergeordneten Kontext kann diese Methode verwendet werden, um generell Datenanalysen in verschiedenen Gebieten der Teilchenphysik zu verbessern.
Der zweite Teil der Arbeit beschreibt eine konkrete Suche nach Phänomenen, die dafür sprechen, das SM um das Konzept der Supersymmetrie (SUSY) zu erweitern. Auf SUSY basierende Erweiterungen des SM sind in der Lage, einer Vielzahl der theoretischen Unzulänglichkeiten des SM zu begegnen und eine Erklärung für Dunkle Materie zu liefern. Im Speziellen werden sogenannte *natürliche* SUSY-Szenarien untersucht, in denen Higgsinos mit sehr kleinen Massendifferenzen vorhergesagt werden.

Durch die hier eingeführte *refinement*-Methode können klassische, sogenannte *fast simulation*-Anwendungen so erweitert werden, dass sie eine Genauigkeit erreichen, die sehr nah an die der gängigen detaillierteren, aber rechenintensiveren Simulationsanwendungen herankommt, was in dieser Arbeit durch mehrere Studien gezeigt wird. Dies eröffnet die Möglichkeit sehr viel effizienterer detailreicher Suchen nach neuer Physik im zukünftigen CMS-Physikprogramm.
Die im zweiten Teil beschriebene Analyse ist sensitiv in Phasenraumregionen, die einige der letzten möglichen Manifestationen natürlicher SUSY-Szenarien beinhalten. Es zeigt sich, dass die vom CMS-Experiment am LHC aufgenommenen Daten mit den Vorhersagen des SM konsistent sind und somit keine Anzeichen für Physik jenseits des SM vorliegen. In der Analyse können Modelle mit Higgsino-Massen bis zu 180 GeV und Massendifferenzen zwischen 0,3 und 1,2 GeV ausgeschlossen werden. Damit wird der in vorangegangenen Suchen erforschte Bereich signifikant erweitert.

Author's Contributions and Publications

The refinement method introduced in Part I of this thesis has been published as "Fast Perfekt: Regression-based refinement of fast simulation" in *SciPost Phys. Core* [1]. The author played a leading role in the conceptual design of the procedure and its technical implementation. In the published article, the author provided, among other things, the studies compiled in Section 4. Initial concepts of the method were reported as "Refining fast simulation using machine learning" in *EPJ Web Conf.* as proceedings to the 26th International Conference on Computing in High Energy & Nuclear Physics featuring a presentation by the author on the topic [2].

The search for new physics presented in Part II of this thesis is currently being reviewed for publication by the CMS Collaboration. The author served as the main analyzer carrying out the analysis, training the machine learning algorithms, developing the special treatment of low-momentum tracks and the background estimation, as well as making core conceptual contributions to all stages of the analysis.

As part of the CMS Collaboration, the author contributed to the data taking of the CMS experiment during Run 3 of the LHC by taking online data quality monitoring shifts. Furthermore, the author implemented the refinement method in the CMS software framework facilitating its usage as part of the CMS FastSim application.

Contents

1 The Standard Model of Particle Physics and Beyond

Nearly all knowledge about the building blocks of our universe is encapsulated in a remarkable theory called the Standard Model of particle physics (SM). Describing all known fundamental particles and their interactions, the SM has been developed and validated in great detail throughout the past century [3–7]. One of the most recent experimental highlights was the observation of a SM-compatible Higgs boson by the ATLAS and CMS collaborations at the LHC in 2012 [8, 9]. With this discovery, all particles included in the SM have been experimentally confirmed, whereas, as of yet, no particle outside the SM has been observed.

In the Standard Model, particles are described as excitations of quantum fields. Together with the notion of invariance under different symmetries, the theory unfolds. The particle content of the SM is visualized in Fig. 1.1. It can be divided into fermions with half-integer spin and bosons with integer spin. The fermions are further divided into quarks and leptons each arranged in three generations with ascending particle masses. The first generation quarks, *up* and *down*, together with *electrons* (which are first generation leptons) form atoms and make up our everyday matter. The gauge bosons with spin 1 mediate three of the four known fundamental forces: the electromagnetic, weak, and strong force. Gravity is the only observed interaction that is not described in the SM. The Higgs boson is the only SM particle with spin 0 and plays a central role in spontaneous symmetry breaking, responsible for generating the particle masses.

In the following, the formulation of the SM based on its Lagrangian density will be discussed. This exposes the particles along with their interactions and important properties. The chapter will be concluded by a brief review of hints of physics beyond the SM, introducing the concept of supersymmetry as a possible way to extend the theory.

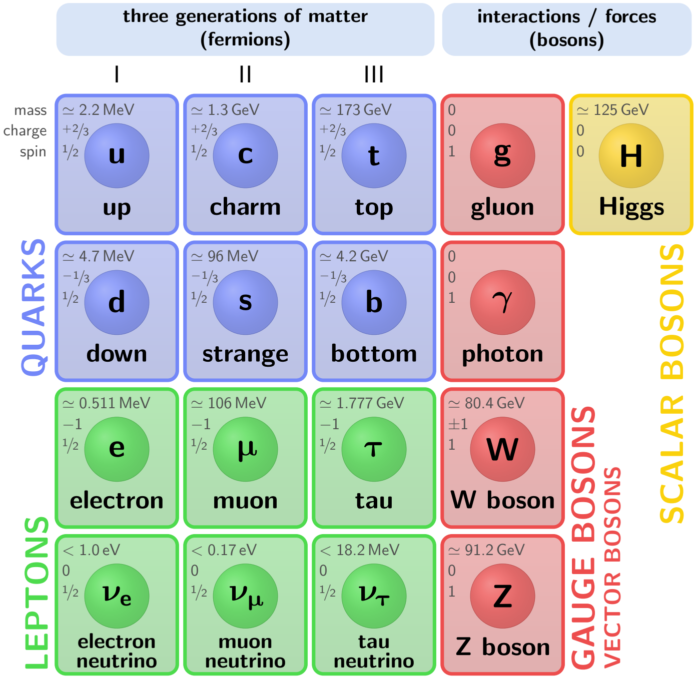

Figure 1.1: Particles of the SM with their mass, charge, and spin [10].

1.1 The Lagrangian of the Standard Model

The Standard Model is formulated as a quantum field theory based on the gauge symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. Here, $SU(3)_C$ refers to the group associated to the strong force which acts on the color charge C and the group $SU(2)_L \times U(1)_Y$ describes the electroweak part of the SM. The subscript L indicates that the weak interaction only acts on left-handed particles, and Y is the weak hypercharge. The SM Lagrangian density is invariant under local $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge transformations as well as under global Poincaré transformations, which include translations, rotations, and Lorentz-boosts.

The full SM Lagrangian can be decomposed into multiple parts, each describing a distinct part of the theory:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EWK}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}} \,. \tag{1.1}$$

Fermions can be represented as 4-component Dirac spinors ψ (combining two spin states of a particle and of its antiparticle) that are solutions of the Dirac equation[1]

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \,. \tag{1.2}$$

---

[1] As throughout this thesis, natural units are used: $c = \hbar = 1$.

The 4×4 gamma matrices can be written with the three Pauli matrices σ^i as:

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \quad i = 1, 2, 3.$$

With this formalism and the Dirac adjoint $\bar{\psi}$ defined as $\bar{\psi} = \psi^\dagger \gamma^0$, the kinetic term for fermions reads:

$$\mathcal{L}_\text{fermion} = i\bar{\psi}\gamma^\mu \partial_\mu \psi \,. \tag{1.3}$$

Fermion spinors can be decomposed according to *chirality* into left-handed and right-handed components. The chiral projection operators are written with $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ as

$$P_L = \frac{1}{2}(1 - \gamma_5) \quad \text{and} \quad P_R = \frac{1}{2}(1 + \gamma_5)\,.$$

Left-handed particles carry the charge related to $SU(2)$ (weak isospin I_3) and are arranged in isospin doublets with weak isospin values of $I_3 = \pm\frac{1}{2}$:

$$\psi_L = \begin{pmatrix} \nu_\text{e} \\ \text{e} \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \begin{pmatrix} \text{u} \\ \text{d}' \end{pmatrix}_L, \begin{pmatrix} \text{c} \\ \text{s}' \end{pmatrix}_L, \begin{pmatrix} \text{t} \\ \text{b}' \end{pmatrix}_L .$$

The flavor eigenstates d$'$, s$'$, b$'$ and ν_e, ν_μ, ν_τ mix to form the mass eigenstates d, s, b and ν_1, ν_2, ν_3.

Right-handed particles carry no weak isospin and are arranged as singlets:

$$\psi_R = \text{e}_R,\ \mu_R,\ \tau_R,\ \text{u}_R,\ \text{d}_R,\ \text{c}_R,\ \text{s}_R,\ \text{b}_R,\ \text{t}_R\,.$$

Many extensions of the SM also include right-handed neutrinos which allows for neutrino masses to be non-zero.

To make the Lagrangian invariant under local $SU(3)$ transformations connected to the theory of **quantum chromodynamics** (QCD), the derivative in Eq. 1.3 is replaced by the *covariant derivative*:

$$D_\mu = \partial_\mu + ig_s T^a G^a_\mu \,, \tag{1.4}$$

where g_s is the strong coupling constant and T^a are the generators of the $SU(3)$ group, explicitly represented by $T^a = \frac{\lambda^a}{2}$ with the Gell-Mann matrices λ^a. The 8 vector fields G^a correspond to 8 gluons that themselves carry color charge and therefore couple to themselves and to quarks. Defining the field strength tensor with structure constant $f^{abc} \neq 0$ as

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f^{abc} G^b_\mu G^c_\nu \,, \tag{1.5}$$

the QCD Lagrangian becomes:

$$\mathcal{L}_{\text{QCD}} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - g_s\bar{\psi}\left(\gamma^{\mu}T^{a}G^{a}_{\mu}\right)\psi - \frac{1}{4}G^{a}_{\mu\nu}G^{a\mu\nu}\,. \tag{1.6}$$

In this way, the kinetic term for fermions (first term) is complemented by a term describing the interaction between quarks and gluons (second term) and a third expression including the kinetic term for gluons as well as their self-interaction in the form of three- and four-gluon vertices.

To describe the **electroweak** part of the SM, local invariance under $SU(2)_L \times U(1)_Y$ is required. The corresponding generators for $SU(2)$ are defined with the Pauli matrices σ^a to be $T^a = \frac{\sigma^a}{2}$. Three new vector fields W^a_μ are introduced with field strength tensor $W^a_{\mu\nu}$ and coupling constant g analogously to Eq. 1.5:

$$W^{a}_{\mu\nu} = \partial_{\mu}W^{a}_{\nu} - \partial_{\nu}W^{a}_{\mu} - g\epsilon^{abc}W^{b}_{\mu}W^{c}_{\nu}\,, \tag{1.7}$$

where ϵ^{abc} is the structure constant. On the other hand, invariance under $U(1)$ leads to one new gauge field B_μ with field strength tensor

$$B_{\mu\nu} = \partial_{\mu}B_{\nu} - \partial_{\nu}B_{\mu}\,. \tag{1.8}$$

With this, the Lagrangian reads

$$\mathcal{L}_{\text{EWK}} = i\bar{\psi}_L\gamma^{\mu}D_{\mu}\psi_L + i\bar{\psi}_R\gamma^{\mu}D_{\mu}\psi_R - \frac{1}{4}W^{a}_{\mu\nu}W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}\,. \tag{1.9}$$

It differentiates between left- and right-handed fermions and the respective covariant derivatives are:

$$D_{\mu,L}\psi_L = \left(\partial_{\mu} + \frac{ig}{2}\sigma^{a}W^{a}_{\mu} + \frac{ig'}{2}YB_{\mu}\right)\psi_L \tag{1.10}$$

$$D_{\mu,R}\psi_R = \left(\partial_{\mu} + \frac{ig'}{2}YB_{\mu}\right)\psi_R. \tag{1.11}$$

In order to include mass terms of the form $-m\bar{\psi}\psi$ in the SM Lagrangian, the **Higgs** mechanism postulates a complex $SU(2)$ doublet with spin 0:

$$\phi = \begin{pmatrix}\phi^{+}\\ \phi^{0}\end{pmatrix} = \begin{pmatrix}\phi_1 + i\phi_2\\ \phi_3 + i\phi_4\end{pmatrix}$$

with corresponding Lagrangian

$$\mathcal{L}_{\text{Higgs}} = \left|D_{\mu}\phi\right|^{2} - V(\phi). \tag{1.12}$$

The potential is assumed to have the form

$$V(\phi) = \mu^2 |\phi|^2 + \lambda |\phi|^4 \ , \tag{1.13}$$

where μ and λ are free parameters. However, for the vacuum to be stable $\lambda > 0$ has to be fulfilled otherwise the potential would not be bounded from below. Also, only $\mu^2 < 0$ leads to the desired effect of spontaneous symmetry breaking where the minimum of the potential lies at $|\phi| = v = \sqrt{\frac{-\mu^2}{\lambda}} \neq 0$ with the vacuum expectation value v.

Expanding the Higgs field around this minimum using the unitary gauge,

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} , \tag{1.14}$$

and inserting it into the first part of the Lagrangian (Eq. 1.12), the physical gauge bosons and their masses are exposed:

$$m_{W^\pm} = \frac{v}{2} g \, , \quad m_Z = \frac{v}{2} \sqrt{g^2 + g'^2} \, , \quad m_A = 0.$$

Here, the $W^\pm$ bosons are linear combinations of $W_{1,2}$, $W^\pm = (W_1 \mp i W_2)/\sqrt{2}$, and the physical Z boson and photon A are formed by rotating the W_3 and B fields by the Weinberg angle θ_W:

$$\begin{pmatrix} A \\ Z \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix} .$$

Given this, the potential can now be written as

$$V(\phi) = -\mu^2 H^2 + \lambda v H^3 + \frac{1}{4} \lambda H^4 \tag{1.15}$$

and includes the mass term of the physical Higgs boson with $m_{\mathrm{H}} = \sqrt{2\mu^2}$ and its self-interaction in the form of three- and four-vertex couplings.

Finally, also fermions acquire mass through spontaneous symmetry breaking by a **Yukawa** interaction between their fields and the Higgs field:

$$\mathcal{L}_{\text{Yukawa}} = -\frac{c_f v}{\sqrt{2}} \bar{\psi}_L \psi_R - \frac{c_f}{\sqrt{2}} \bar{\psi}_L H \psi_R . \tag{1.16}$$

One can see that the coupling strength of a fermion to the Higgs field is proportional to its mass $m_f = c_f v / \sqrt{2}$.

1.2 Indications of Physics Beyond the Standard Model

Despite the SM's success in explaining and predicting a vast and comprehensive set of experimental outcomes, including the measurement of cross-sections over many orders of magnitude, it can only be understood as an approximation to a yet unknown more fundamental theory or even a *theory of everything*. Reasons in the form of experimental evidence include the following:

- The existence of **gravity** is not explained by the SM. The Planck scale $\Lambda_{\text{Planck}} \approx 10^{28}\,\text{eV}$ (many of orders of magnitude larger than what is probed with current high-energy physics experiments) marks the energy scale at which a theory of quantum gravity becomes necessary.
- In the Big Bang, unequal amounts of matter and antimatter were produced. This **baryon asymmetry** can not be explained within the SM, which predicts that any asymmetry is constrained by the amount of CP violation.
- The accelerated expansion of the universe is assumed to be driven by a form of energy called **dark energy**. Although it is measured to constitute the largest share of the universe's energy density (roughly 68% [11]), the SM provides no explanation of dark energy.
- Moreover, there is a host of observations that imply the existence of an additional form of matter that does not interact electromagnetically or via the strong interaction, referred to as **dark matter**. Measurements of the cosmic microwave background by the Planck collaboration, interpreted within the standard model of cosmology, indicate that dark matter makes up approximately 27% of the energy density of the universe, compared to only 5% attributed to ordinary matter [11]. Its effects can also be observed in multiple ways including gravitational lensing, structure formation, and rotational curves of galaxies.

Additionally to such experimental hints for physics beyond the SM, also conceptual flaws have been pointed out. These include the large number of free parameters (mainly the masses of the particles) and the **hierarchy problem**. The latter refers to the fact that the mass of the Higgs boson receives radiative corrections proportional to the energy scale of the theory in which the SM is to be embedded. Assuming this scale to be $\Lambda_{\text{Planck}} \approx 10^{28}\,\text{eV}$, it appears highly unnatural that the corrections cancel to the level necessary to explain the measured Higgs boson mass of approximately 125 GeV.

1.3 Supersymmetry

Supersymmetry (SUSY) is a compelling concept that extends the SM and addresses several of the above mentioned puzzles [12]. It postulates a symmetry between fermions and boson and thus predicts the existence of yet-to-be-observed particles. The minimal supersymmetric Standard Model (MSSM) is the realization of SUSY which introduces the least number of new particles. Its particle content can be seen in Fig. 1.2. The SM Higgs doublet is extended to contain two such doublets, leading to four additional Higgs bosons. The superpartners of the SM fermions are scalar bosons with spin 0 and are denoted with a prepending "s", e.g., *selectron* referring to the superpartner of the electron. For the names of the superpartners of the SM bosons, the suffix "-ino" is appended to those of their SM counterparts. The superpartners of the SM gauge bosons are called *gluino*, *bino*, and *wino*; whereas those of the Higgs bosons are named *higgsino.*

Together with the neutral higgsinos $\widetilde{H}_u^0$ and $\widetilde{H}_d^0$, the bino $\widetilde{B}$ and neutral wino $\widetilde{W}^0$ mix to form mass eigenstates $\widetilde{\chi}_i^0$ called neutralinos with $i \in \{1,2,3,4\}$ ascending in mass. Using these states as basis vectors, the mixing is characterized by the mass matrix

$$M_N = \begin{pmatrix} M_1 & 0 & -g'v_d/2 & g'v_u/2 \\ 0 & M_2 & gv_d/2 & -gv_u/2 \\ -g'v_d/2 & gv_d/2 & 0 & -\mu \\ g'v_u/2 & -gv_u/2 & -\mu & 0 \end{pmatrix} , \tag{1.17}$$

where M_1, M_2 and μ are the bino, wino and higgsino mass parameters, respectively, and $v_{u,d}$ are the vacuum expectation values of the Higgs doublets.

Similarly, the charged winos $\widetilde{W}^\pm$ and the charged higgsinos $\widetilde{H}_u^+$ and $\widetilde{H}_d^-$ mix to chargino states $\widetilde{\chi}_i^\pm$ $(i \in \{1,2\})$ according to the mass matrix

$$M_C = \begin{pmatrix} M_2 & gv_u/\sqrt{2} \\ gv_d/\sqrt{2} & \mu \end{pmatrix} . \tag{1.18}$$

In the context of the MSSM, a new quantum number called R-parity is introduced as:

$$P_R = (-1)^{3B+L+2s}. \tag{1.19}$$

Here, B is the baryon number, L the lepton number, and s the spin of a particle. Particles of the SM have R-parity +1 and their superpartners $P_R = -1$. If R-parity is conserved, SUSY particles can only be produced in even numbers. Conservation of R-parity is motivated for example by the stability of the proton. Furthermore, it leads to the fact that the lightest supersymmetric particle (LSP) is stable which makes it a suitable dark matter candidate.

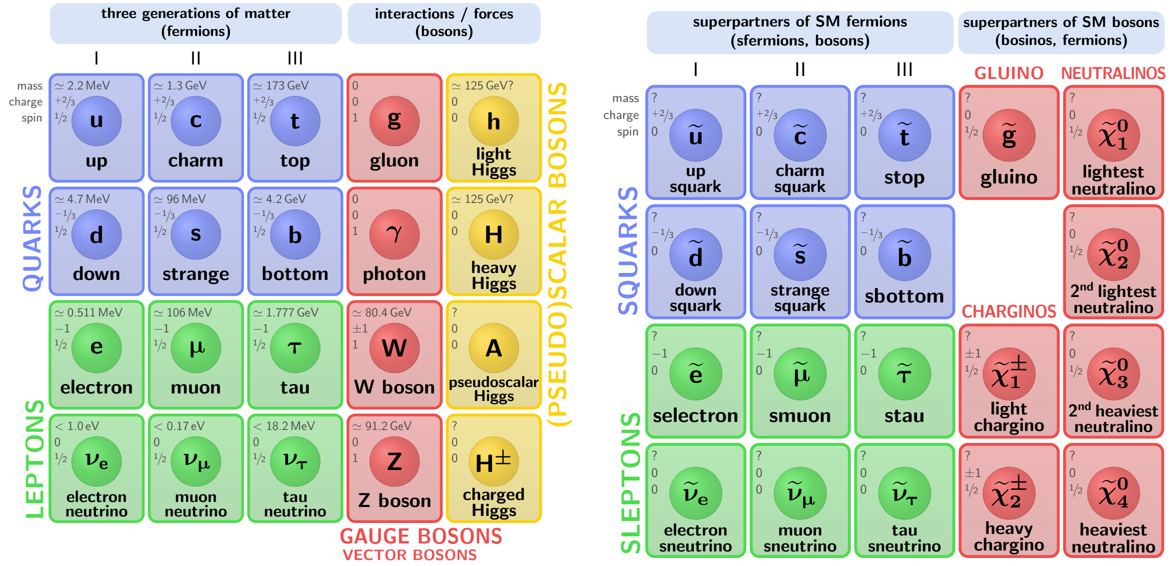

Figure 1.2: Particle content of the minimal supersymmetric Standard Model divided into particles with R-parity +1 (left) and -1 (right) [10].

The hierarchy problem of the SM receives an elegant solution in the MSSM. The seemingly accidental cancellations in the radiative corrections to the Higgs boson mass are explained by the fact that fermions and bosons contribute with opposite sign.

It is noted that an unbroken version of supersymmetry would mean that the superpartners of the SM particles should have the same mass as their counterparts. Apparent from the absence of evidence for such particles, SUSY must therefore be a broken symmetry. If the superpartner masses are too large, a new fine-tuning problem arises. Still, the corresponding *little* hierarchy problem is orders of magnitude smaller than the original hierarchy problem. Its extent can be quantified by the electroweak fine-tuning parameter Δ_{EW} [13] which is defined in the following way: First, minimization of the MSSM potential leads to the relation

$$\frac{m_Z^2}{2} = \frac{(m_{H_d}^2 + \Sigma_d^d) - (m_{H_u}^2 + \Sigma_u^u)\tan^2\beta}{\tan^2\beta - 1} - \mu^2\,, \tag{1.20}$$

where H_d and H_u refer to the two supersymmetric Higgs doublets, $\tan\beta = v_u/v_d$ to the ratio of their vacuum expectation values, and Σ_d^d and Σ_u^u are loop corrections. Now, the electroweak fine-tuning parameter is defined with the individual contributions C_i in the expanded version of Eq. 1.20 to be

$$\Delta_{\mathrm{EW}} = \max_i \frac{|C_i|}{m_Z^2/2}\,, \tag{1.21}$$

with $C_{H_d} = m^2_{H_d}/(\tan^2\beta - 1)$, $C_{H_u} = -m^2_{H_u}\tan^2\beta/(\tan^2\beta - 1)$, $C_\mu = -\mu^2$, and $C_{\Sigma^d_d(k)} = -\Sigma^d_d(k)/(\tan^2\beta - 1)$, $C_{\Sigma^u_u(k)} = -\Sigma^u_u(k)\tan^2\beta/(\tan^2\beta - 1)$ where k labels the various loop contributions.

Natural realizations of SUSY with low fine-tuning require μ to be close to the SM Z boson mass and are thus generally characterized by higgsino-like charginos and neutralinos with masses $\lesssim 200\,\text{GeV}$, third-generation squarks with masses $\lesssim 1.5\,\text{TeV}$ and gluinos with $m_{\tilde{g}} \lesssim 3\,\text{TeV}$ [13]. Part II of this thesis presents a search for such models, targeting the decay of higgsino-like charginos featuring low-momentum, isolated and mildly displaced pions along with a large magnitude of missing transverse momentum.

2 The CMS Experiment at the Large Hadron Collider

In order to experimentally scrutinize the SM and examine its theorized extensions, arguably the best setup is provided by high-energy particle accelerators. As the world's largest and most powerful accelerator, the Large Hadron Collider (LHC) enables unparalleled insights into the subatomic world. It is located at the CERN research facility in and around Geneva, Switzerland, and hosts four main experiments residing at four interaction points along its circular path. The research for this thesis was performed on simulations for and data recorded by one of those experiments, the Compact Muon Solenoid (CMS), which is introduced in this chapter.

2.1 The Large Hadron Collider

The LHC is a circular particle accelerator with a circumference of 27 km designed to operate at a proton-proton center-of-mass energy of up to $\sqrt{s} = 14\,\text{TeV}$. It is embedded in the CERN accelerator complex as the final part, in which the particles gain their highest energies after having gone through a chain of pre-accelerators as depicted in Fig. 2.1. The particles (most of the time protons but also heavy lead ions) pass through the LHC ring in two opposing beams and are brought to collision at four interaction points. At each of those points, a particle detector is located: ATLAS, ALICE, CMS, and LHCb in a clockwise direction from the CERN main campus.

Apart from the center-of-mass energy, the (instantaneous) luminosity $\mathcal{L}$ is a determining property of a particle collider. It can be used to calculate the rate of a given process with cross section σ:

$$\frac{dN}{dt} = \mathcal{L} \times \sigma\,. \tag{2.1}$$

By integrating over a period of time, the integrated luminosity L is given by

$$L = \int \mathcal{L}\, dt\,. \tag{2.2}$$

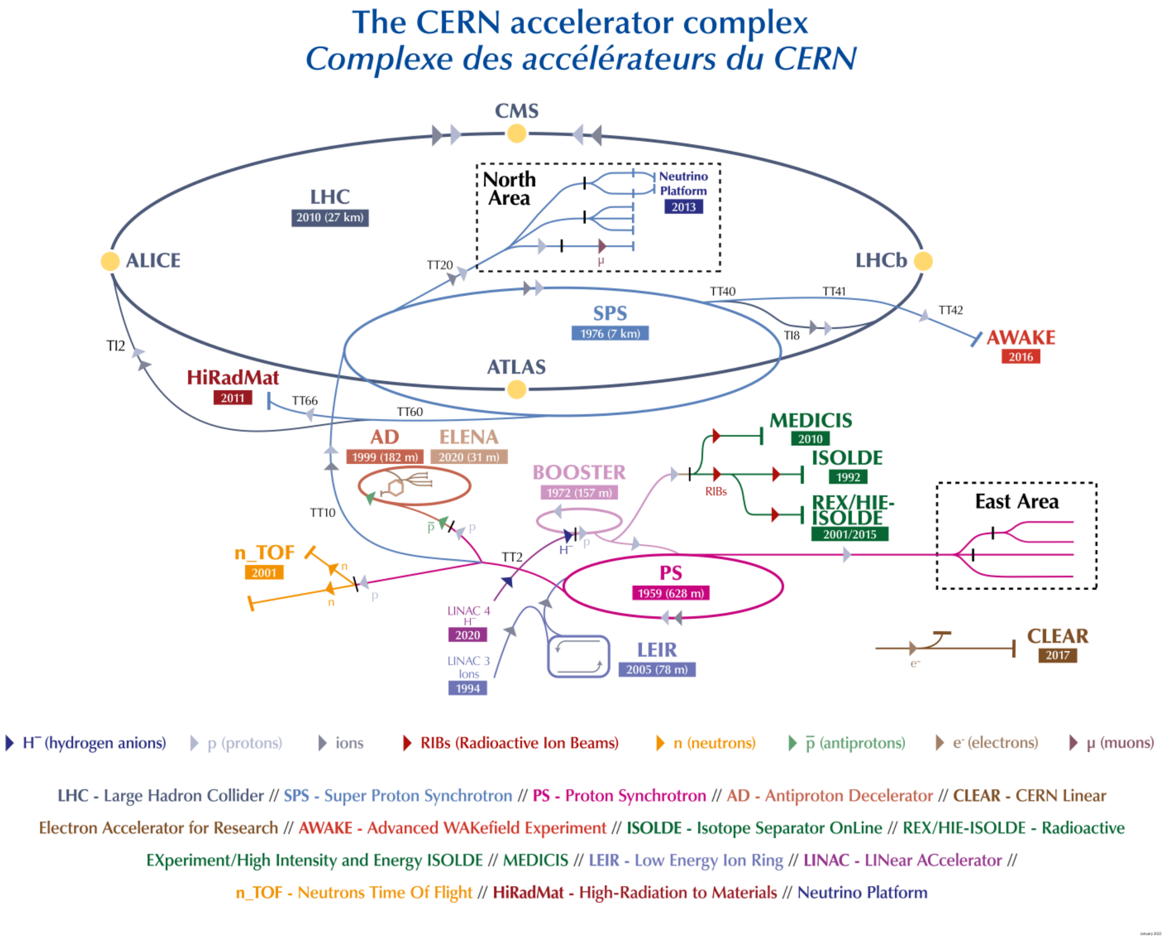

Figure 2.1: Schematic of the CERN accelerator complex [14].

Time periods over which the LHC collides particles with a similar set of conditions are called *Runs*, where Run 1 refers to the years 2010 through 2012 with collisions at $\sqrt{s} = 7\,\text{TeV}$ and $8\,\text{TeV}$, and a total delivered luminosity of approximately $30\,\text{fb}^{-1}$. After the first long shutdown, Run 2 started in 2015 and lasted until 2018 with $\sqrt{s} = 13\,\text{TeV}$ and $L \approx 150\,\text{fb}^{-1}$. The currently ongoing Run 3 started in 2022 and is characterized by the highest ever center-of-mass energy of 13.6 TeV. Also, the mean number of hard interactions per beam crossing (referred to as *pileup*) is greater than in all previous years as can be seen in Fig. 2.2.

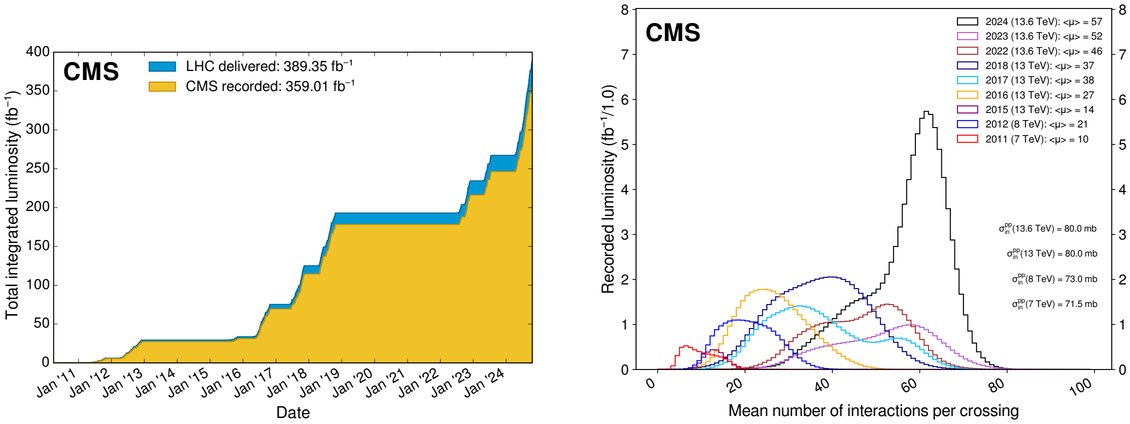

Figure 2.2: Left: Integrated luminosity delivered by the LHC and recorded by CMS cumulative over the course of the three Runs. Right: Breakdown of recorded luminosity by year and mean number of interactions per crossing [15].

2.2 The CMS Experiment

The Compact Muon Solenoid (CMS) detector is one of the two multi-purpose particle detectors at the LHC, located approximately 100 meters below ground near the French village of Cessy. It is built like a cylindrical onion with multiple sub-detectors layered around the interaction point and measures 29 meters in length and 15 meters in diameter (a big onion). Figure 2.3 shows an overview of the detector components along with some of their most important features. In the following, a brief summary of the components is given; more details can be found in Ref. [16].

The eponymous superconducting **solenoid magnet** generates a magnetic field of 3.8 T in which the trajectories of charged particles are bent according to their charge and momentum. Enclosed by the magnet, the tracking detectors and calorimeters provide momentum and energy measurements. The innermost part, the **pixel tracker** — upgraded in 2016/2017 as described in Ref. [18] — consists of 124 million silicon sensors

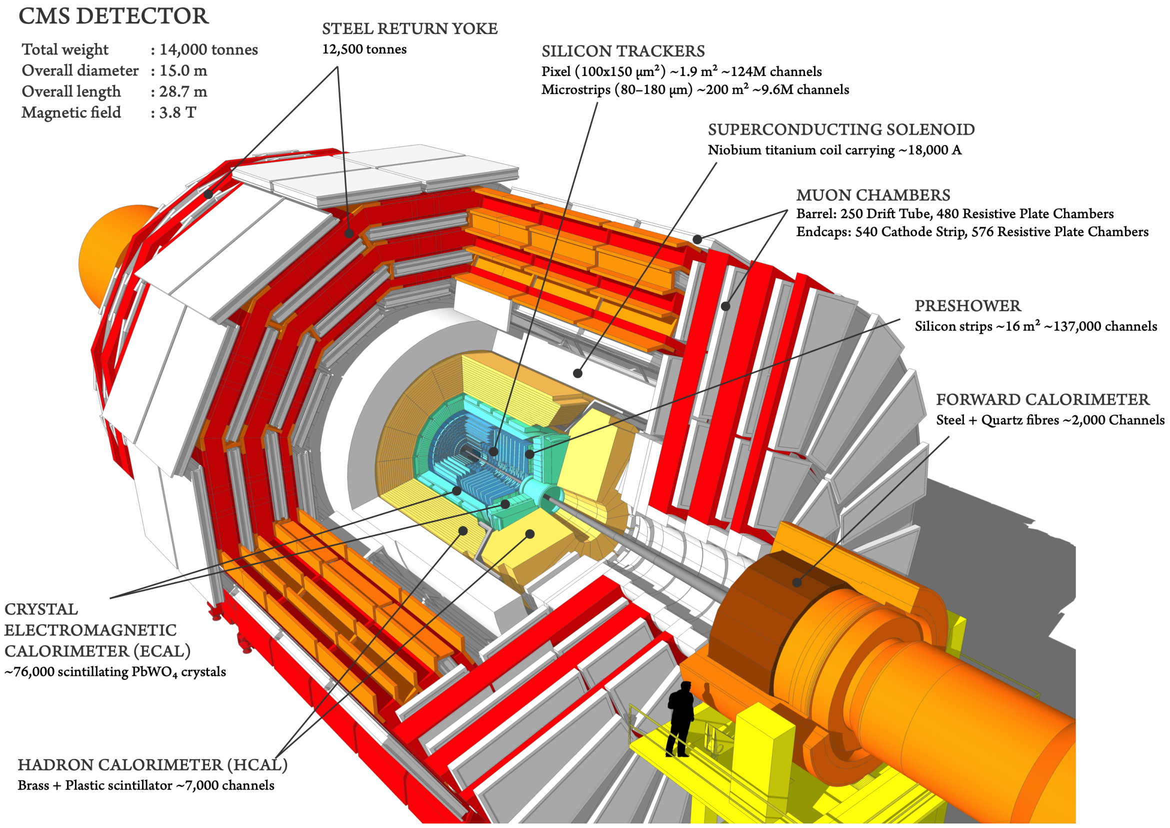

Figure 2.3: A cutaway view of the CMS detector with its various subdetectors and important features [17].

sized $100 \times 150\,\mu\mathrm{m}^2$. They are arranged in concentric cylindrical layers around the beam pipe (referred to as the barrel part, BPIX) and in disks in the forward region (FPIX) as seen in Fig. 2.4. The pixel tracker is surrounded by a **silicon strip detector**, adding 9 million channels to the tracking system. The setup achieves a resolution of the transverse impact parameter with respect to the primary interaction vertex as precise as $20\,\mu\mathrm{m}$ for tracks with transverse momentum $p_\mathrm{T} \gtrsim 10\,\mathrm{GeV}$ ($60\,\mu\mathrm{m}$ for $p_\mathrm{T} \approx 1\,\mathrm{GeV}$) [19]. The resolution of the transverse (longitudinal) position of the primary interaction vertex reaches $10\,\mu\mathrm{m}$ ($15\,\mu\mathrm{m}$) [20].

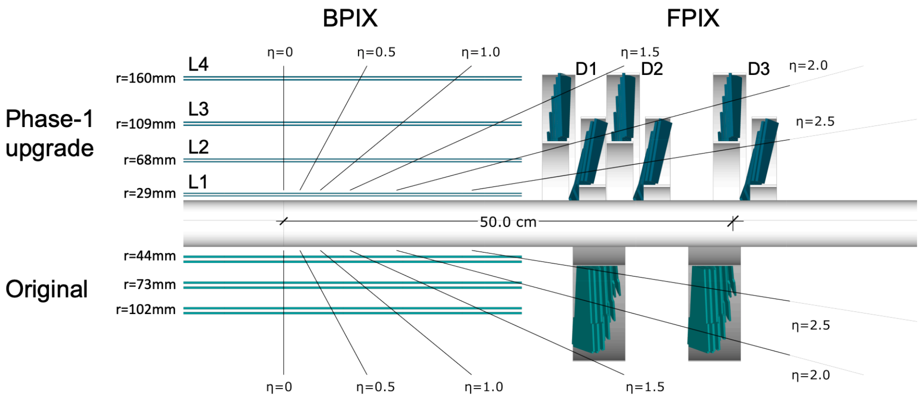

Figure 2.4: Layout of the CMS pixel tracker before (bottom) and after (top) the Phase-1 upgrade between the 2016 and 2017 data taking [18].

Apart from the tracker, the volume within the solenoid magnet also houses two calorimeter systems, designed to stop emerging particles and measure their energies. The **electromagnetic calorimeter** consists of 76,000 scintillator crystals made from lead tungstate ($\mathrm{PbWO_4}$), that act both as an absorber and active material. The **hadron calorimeter** on the other hand is built as a sampling calorimeter with alternating layers of brass as the absorber and plastic scintillator as the active material. Both calorimeter systems include barrel parts and endcaps extending the coverage to regions close to the beam pipe.

The outermost part of the detector is the **muon system** which is embedded in the steel return yoke outside the magnet coil. It consists of various types of gas-ionization chambers (drift tubes, resistive plate chambers, and cathode strip chambers) providing measurements of muon tracks passing through the detector.

As it is not feasible to store the detector response of all occurring collisions, a two-tiered **trigger** system is employed, which reduces the rate of events to be read out from 40 MHz to $\approx 1\,\mathrm{kHz}$. It consists of the Level-1 (L1) trigger implemented in custom hardware which enables a latency of $4\,\mu\mathrm{s}$ [21]. The trigger decision is made by combining information from the calorimeters and the muon systems, yielding a simplified description of the full event.

The second tier, the High-Level Trigger (HLT), uses the full detector information and is run on a farm of CPU processors. Upon selection by the HLT, events are transferred to the Tier 0 data center for reconstruction and storage.

The origin of the **coordinate system** used for the CMS experiment sits at the nominal interaction point, that is, the center of the detector. The x-axis then points towards the center of the LHC ring, the y-axis upward, and the z-axis along the beam pipe towards the Jura Mountains. Therefore, the transverse momentum is $p_\mathrm{T} = \sqrt{p_x^2 + p_y^2}$. As shown in Fig. 2.5, the azimuthal angle φ is the angle in the x-y (transverse) plane to the x-axis and the polar angle θ in the z-y plane to the z-axis. Most often, to state the direction of a particle, the polar angle is transformed to the *pseudorapidity* η according to

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) \tag{2.3}$$

in order to obtain invariant differences under Lorentz-boosts along the z-axis. To express the angular distance between particles, the distance measure $\Delta R = \sqrt{\Delta\varphi^2 + \Delta\eta^2}$ is typically used.

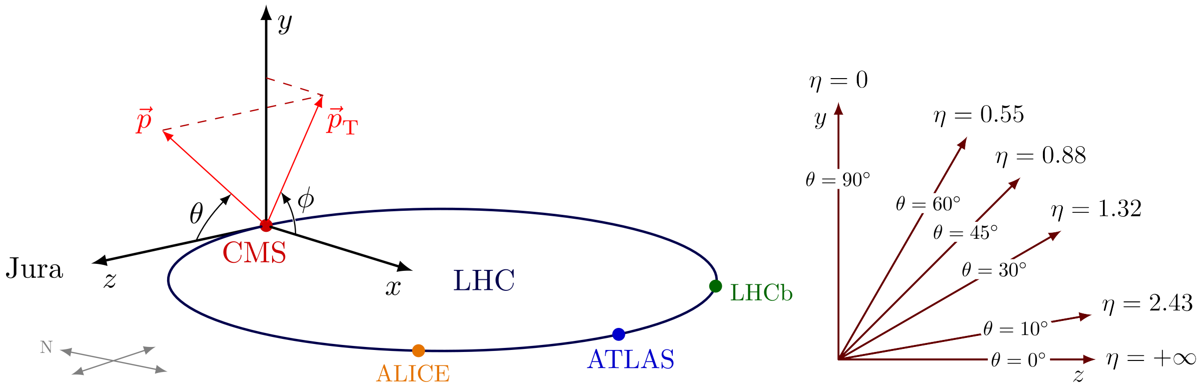

Figure 2.5: The CMS coordinate system and a visualization of the definition of pseudorapidity [10].

2.2.1 Object Reconstruction

In each collision event, a multitude of particles are created both from the primary (highest-energetic) proton-proton collision as well as from additional pileup interactions. The aim of the global event reconstruction is to identify each of those particles by combining information from all detector sub-systems. In CMS, the Particle Flow (PF) algorithm is used for this task [22]. It makes use of the fact that different particle types leave different traces in the detector. For example, as shown in Fig. 2.6, electrons leave a track in the silicon tracker and afterwards deposit their energy in the electromagnetic calorimeter.

The PF algorithm starts by linking tracks reconstructed in the tracker to energy clusters in the calorimeters, forming so-called PF blocks. Additional ECAL energy deposits are linked to the block if they are located tangentially to the track, capturing bremsstrahlung photons. As the first physics objects, muons are identified if a PF block has an associated track in the muon chambers and the various elements fulfill certain quality and consistency criteria. The elements used are then "locked" and not considered in the following steps of the algorithm. Next, electrons and photons are constructed from PF blocks comprised of clusters in the ECAL with (electrons) and without (photons) matching tracks. Energy deposits in the ECAL and HCAL without a matching track are interpreted as neutral hadrons, whereas charged hadrons are identified by combining tracks and calorimeter clusters.

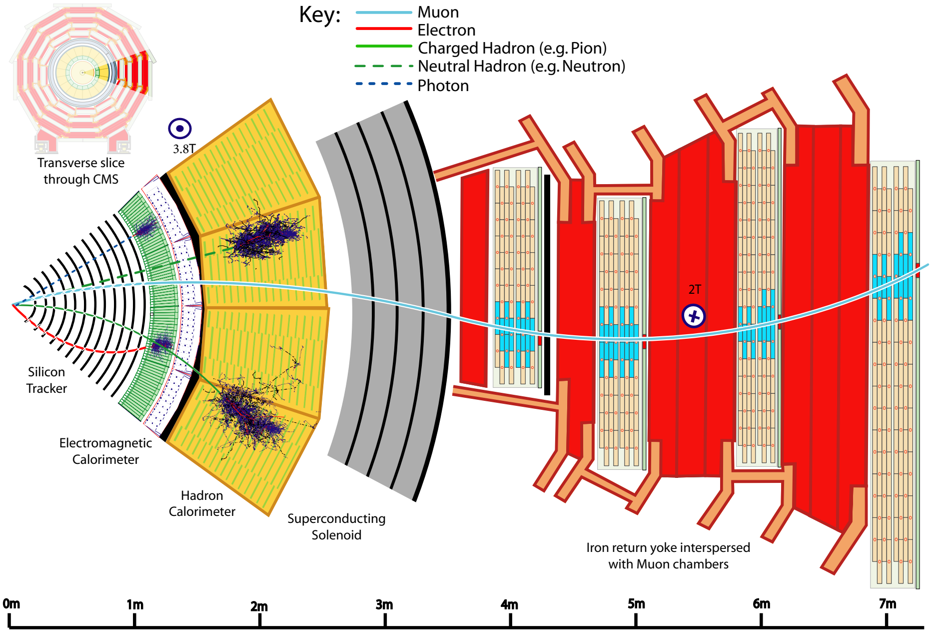

Figure 2.6: Transverse slice through the CMS detector with characteristic traces left by different particle species [22].

In the following, the reconstruction of tracks and the formation of jets is described in more detail.

Tracks To reconstruct the helical paths of all charged particles in an event, an iterative algorithm based on Kalman filtering is run on the tracker hits [20]. Each iteration consists of three steps: In the first step, referred to as *seeding*, two or three hits are identified as a starting point for the next steps. The track *finding* is then performed by extrapolating the track candidate from its seed along the expected path and progressively adding hits. As a third step, track *fitting* is performed to constrain the track parameters and get their uncertainties. Finally, track candidates are kept or discarded depending on the number of hits and the quality of the fit. These three steps are run multiple times iteratively with different seeding configurations and quality criteria, targeting first the reconstruction of high-momentum prompt tracks with later iterations tailored to displaced and/or lower-momentum tracks, possibly with missing hits.

The standard track parameterization used in CMS is based on five quantities defined at the *reference point* of the track (x_0, y_0, z_0), which is the point of closest approach to the nominal interaction point, i.e., the center of the detector:

1. Azimuthal angle φ_0
2. Dip angle $\lambda = \frac{\pi}{2} - \theta$
3. Signed inverse momentum $q/|\vec{p}|$
4. Signed minimal distance in the transverse plane between the straight line passing through (x_0, y_0) with angle φ and the point $(0, 0)$:

$$\mathrm{dxy}_0 = -x_0 \sin(\varphi) + y_0 \cos(\varphi)$$

5. Signed minimal distance in the s-z-plane between the straight line passing through (x_0, y_0, z_0) with angles φ and λ and the point $(s = 0, z = 0)$. The s-axis is defined by the projection of this straight line onto the transverse plane:

$$\mathrm{dsz}_0 = z_0 \cos(\lambda) - (x_0 \cos(\varphi) + y_0 \sin(\varphi)) \sin(\lambda)\,.$$

See Section 3.1 of Part II for further discussion of soft and displaced tracks.

Jets The objects reconstructed by the PF algorithm, collectively referred to as *PF candidates*, are also used to construct higher-level objects like jets. For the clustering of jets, the anti-k_T algorithm [23] is run on all reconstructed PF candidates. Most often, the distance parameter R is set to 0.4 (AK4 jets) or 0.8 (AK8 jets).

An important aspect of reconstructing jets is the rejection of particles from pileup interactions. This can be achieved by omitting PF candidates with a track which is associated to a pileup vertex. This procedure is effective in rejecting charged hadrons from pileup interactions — and is thus known as charged hadron subtraction (CHS) — but falls short in covering neutral particles and particles outside the tracker acceptance.

2.3 Simulation

A crucial part of all LHC experiments, like all modern high-energy physics experiments, is the accurate simulation of particle collisions (the *event*), the detector response, and subsequent reconstruction. This is realized using stochastic Monte Carlo (MC) techniques and involves multiple steps and dedicated software applications.

GEN In the first step, the collision of two protons is simulated using an *event generator* like MadGraph [24] or pythia [25]. Mainly due to the composite nature of protons, this step includes the simulation of many individual processes as depicted in Fig. 2.7.

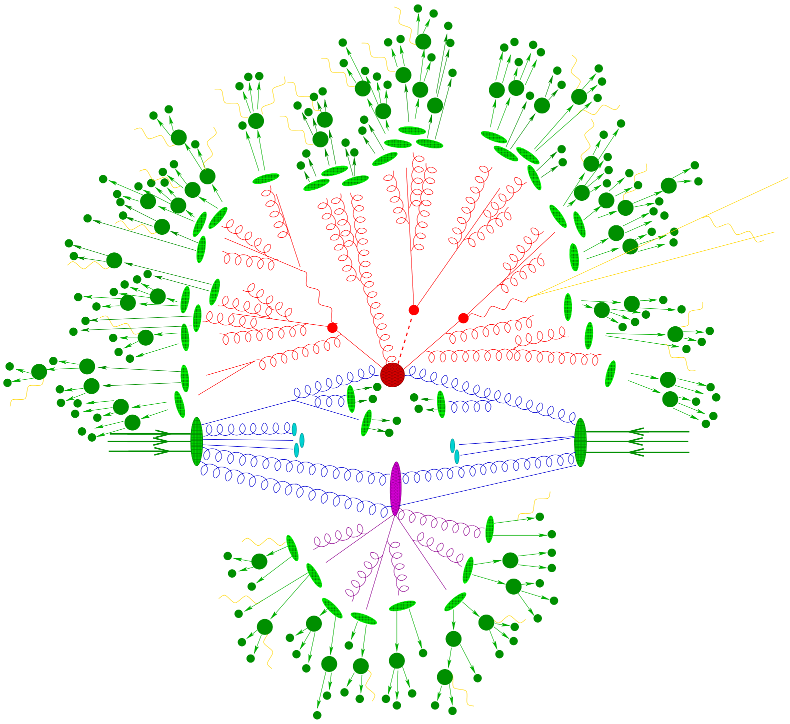

Figure 2.7: Diagrammatic representation of an LHC event showing the collision of two protons (denoted by three incoming green lines from left and right). The hard interaction (large red circle) is followed by the decay of the produced particles (small red circles), the hadronization marked in light green and the hadron decays (dark green). Additionally to the hard interaction, the remaining initial state partons also interact (purple blob). Photon radiation is marked with yellow curly lines [26].

SIM The next step involves the simulation of the passage of the final state particles as produced by the GEN step through the detector. It relies on a detailed representation of the detector geometry and a precise simulation of the interaction with the detector material, implemented, for example, in GEANT4 [27–29].

DIGI & RECO Subsequently, the simulated detector hits are digitized. The output of this step is equivalent to what is recorded by the detector in the real experiment. The event reconstruction as described in Section 2.2.1 is then also run on the simulated event, yielding the objects that are used for further analysis.

2.3.1 The CMS Fast Simulation Framework

In CMS, two simulation chains are used that trade off between accuracy and speed. The fast simulation application (*FastSim*) [30, 31] is, including all steps, approximately 10 times faster than the traditional high-fidelity GEANT4-based full simulation (*FullSim*). This speed-up is possible by a number of simplifications, e.g.,

- fast calorimetry based on GFLASH [32],
- simplified detector geometry (infinitesimally thin layers),
- parametrized material interactions,
- fast tracking in the RECO step using information from the event generator.

Figure 2.8 shows the CPU time per event consumed by the SIM step (in which the speed advantage of FastSim is most prominent) for FullSim and FastSim for two different physics processes and for different versions of the CMS software framework (CMSSW).

Generally, the FastSim output is in good agreement (within 10%) with the respective simulated data using FullSim, but the speed-up also comes at the price of decreased accuracy in some final analysis observables. In Part I of this thesis, a method to *refine* the output of the FastSim chain and thus improve the agreement with FullSim is presented. This effort aims at expanding the usage of FastSim within the collaboration which, in turn, is an important aspect of staying within the computing budget for the coming Runs of LHC operation featuring higher luminosity and more granular detectors, see Fig. 2.9.

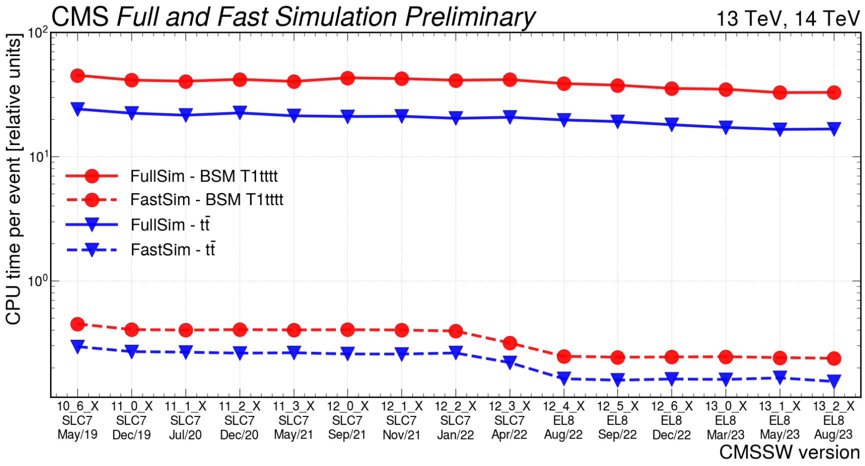

Figure 2.8: CPU time performance of CMS FullSim and FastSim for the SIM step, simulating two different physics processes. The numbers are given for different computer architectures and versions of CMSSW [33].

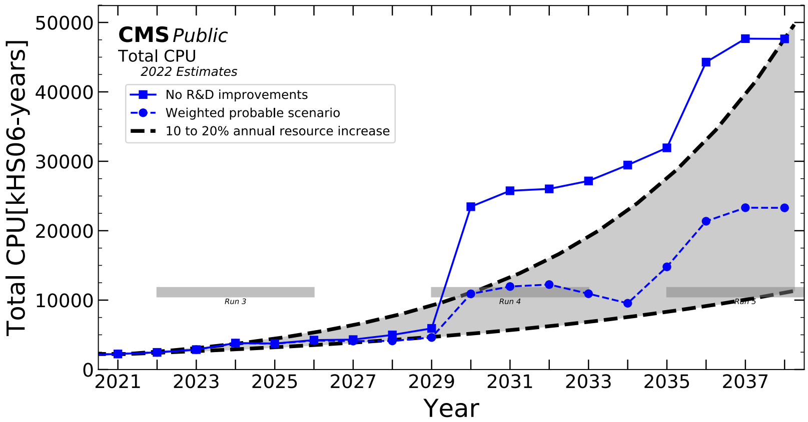

Figure 2.9: Projection of the total CPU time needed in CMS considering two scenarios: a baseline with no R&D and a scenario with the most probable outcome of ongoing R&D activities (blue lines). The gray shaded area shows the projected available resources [34].

Part I

Machine Learning-Based Refinement of Simulations

As introduced in Section 2.3 of the Introduction, the CMS Collaboration uses a fast simulation chain (FastSim) as an alternative to the high-fidelity yet more resource-intensive full simulation framework (FullSim). The usefulness of FastSim and the positive impact of its usage on the computing budget rely on its accuracy. Ongoing efforts to improve the accuracy of FastSim aim to optimize its internal algorithms and the parameterizations used. In contrast to these approaches, this part introduces a method that uses machine learning (ML) to apply a post-hoc modification to the FastSim output improving its accuracy. This *refinement* strategy makes maximum use of the domain knowledge contained in the FastSim framework and applies a relatively small deterministic residual correction. This technique and corresponding machine ML architecture are much lighter than many other applications of (generative) ML for LHC simulations that attempt to learn the feature space from random noise. The task can be understood as a general distribution matching or morphing problem which is approached by training a regression neural network with a dedicated ensemble-based loss function. The method is described in detail in Chapter 1. Chapter 2 shows that the refinement, applied to various use cases, leads to a considerably improved agreement with the target distribution not only in one-dimensional projections but also considering correlations between observables. Beyond this application, Part II of this thesis also makes use of the method in the context of refining Monte Carlo simulation targeting real data.

1 Refinement Method

The basic concept of the refinement method is the application of a regression neural network as a post-hoc correction to selected observables produced by a given simulation chain. In doing so, the agreement of the simulation with a target is improved. The agreement can be measured in one-dimensional marginal distributions as well as in correlations within the set of refined and additional observables. Importantly, the method does not apply weights in order to improve the distribution-level agreement — like classical histogram-based techniques or, e.g., the DCTR approach [35] — but instead, the values of the observables are modified. This preserves the statistical power of the input data set[1] and enables a multidimensional and unbinned correction which is not limited to the domain of the input, like it is the case when applying weights. However, this approach does not inherently correct for efficiency biases, meaning that simulated data points are always shifted in the multidimensional feature space but cannot be disregarded or down-weighted. Other realizations of the refinement idea can be found in Refs. [36–39].

This application of machine learning for high-energy physics simulations is different from many others, e.g., generative modeling for calorimeter simulations (see [40] for a review), as it builds on the established, classical (fast) simulation engine and its physics-based domain knowledge. It does not add additional stochasticity to the simulation, but is of deterministic nature ensuring traceability. It is assumed that the simulation which is refined already contains the needed stochasticity.

The vector of features to be refined is denoted by $\mathbf{x}'$ and after refinement by $\hat{\mathbf{x}}(\theta)$, the latter depending on the network parameters θ. The vector $\mathbf{x}$ refers to the corresponding features in the target data set. Each element of $\mathbf{x}'$ is taken as input to the refiner network, possibly along with conditioning variables $\mathbf{y}'$, concatenated as $\mathbf{a}' = (\mathbf{x}', \mathbf{y}')^\intercal$. In the presented studies, additional "hidden" observables of the input (target) dataset $\mathbf{h}'$ ($\mathbf{h}$) are used as a proof of principle of the method. Although they are not directly refined, the refined correlations $\mathrm{corr}(\hat{\mathbf{x}}, \mathbf{h}')$ are studied regarding their agreement with $\mathrm{corr}(\mathbf{x}, \mathbf{h})$.

---

[1] A weighted data set composed of N samples with weights w_i (with mean $\overline{w}$ and variance $s_w^2 = \frac{1}{N}\sum_{i=1}^{N}(w_i - \overline{w})^2$) has the same relative statistical fluctuations as a data set composed of $N_{\text{eff.}}$ unweighted samples, where $N_{\text{eff.}} = (\sum_i w_i)^2 / \sum_i w_i^2 = N/(1 + (s_w^2/\overline{w})^2) < N$.

1.1 Network Architecture

The refiner network is built as a residual network, which means that before training it acts as the identity function. This is achieved by using *residual blocks* consisting of multiple fully-connected linear layers supplemented by a *skip connection* which adds the input of the first layer of a block to the output of the last layer, see Fig. 1.1. The trainable network parameters are initialized according to the Fixup method [41]: All biases are set to zero before training, as are the weights of all linear layers except for the first of each residual block which are initialized by the Kaiming initialization [42]. The architecture is inspired by the ResNet model [43, 44].

Additional pre- and post-processing layers might be added to the network to apply transformations to the variables, e.g., to increase their dynamic range or to ensure boundary conditions given by the problem at hand.

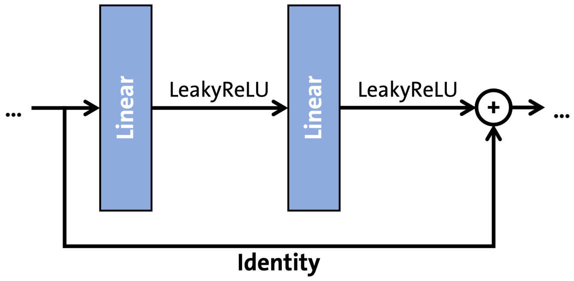

Figure 1.1: Depiction of a residual block with two linear layers and a skip connection used to initialize the refinement network as the identity mapping.

1.2 Loss Functions

The method makes use of two types of loss function to measure the similarity between the network output and the target, the maximum mean discrepancy and the mean squared error. First, estimators of the **maximum mean discrepancy** (MMD) are used as ensemble-based distance measures. The MMD is an integral probability metric that measures the distance between two distributions via their embeddings in a reproducing kernel Hilbert space [45].

Given two samples from the distributions A and $\hat{A}$, denoted by $\{\mathbf{a}_i\}_{i=1,...m}$ and $\{\hat{\mathbf{a}}_i\}_{i=1,...m}$, with vectors $\mathbf{a}_i = (a_i^1, ..., a_i^{n_a})^\intercal$ and $\hat{\mathbf{a}}_i = (\hat{a}_i^1, ..., \hat{a}_i^{n_a})^\intercal$ (where n_a is the number of dimensions), two estimators of the MMD can be calculated; the biased $\mathrm{MMD_b}$ (Eq. 1.1) and the unbiased $\mathrm{MMD_u}$ (Eq. 1.2):

$$\mathrm{MMD_b}(\theta) = \frac{1}{m^2}\sum_{i,j=1}^{m} k(\mathbf{a}_i, \mathbf{a}_j) + \frac{1}{m^2}\sum_{i,j=1}^{m} k(\hat{\mathbf{a}}_i(\theta), \hat{\mathbf{a}}_j(\theta)) - \frac{2}{m^2}\sum_{i,j=1}^{m} k(\mathbf{a}_i, \hat{\mathbf{a}}_j(\theta))\,, \tag{1.1}$$

$$\begin{aligned}\mathrm{MMD_u}(\theta) =& \frac{1}{m(m-1)}\sum_{i=1}^{m}\sum_{i\neq j}^{m} k(\mathbf{a}_i, \mathbf{a}_j) + \frac{1}{m(m-1)}\sum_{i=1}^{m}\sum_{i\neq j}^{m} k(\hat{\mathbf{a}}_i(\theta), \hat{\mathbf{a}}_j(\theta)) \\ &- \frac{2}{m^2}\sum_{i,j=1}^{m} k(\mathbf{a}_i, \hat{\mathbf{a}}_j(\theta))\,.\end{aligned} \tag{1.2}$$

Both estimators use a kernel function k which is chosen to be the Gaussian kernel

$$k(\mathbf{a}, \hat{\mathbf{a}}) = \exp\left(-\sum_{l=1}^{n_a} \frac{(\hat{a}^l - a^l)^2}{\sigma_l^2}\right)\,, \tag{1.3}$$

with bandwidth set by the median heuristic for each dimension: $\sigma_l = \mathrm{median}\{\|a_i^{\prime l} - a_j^l\| : i, j \in [m]\}$ [45]; this ensures similar influence from each dimension with possibly different scales. Choosing the median distance is reasonable since a too large bandwidth would put all possible values of $k(\mathbf{a}, \hat{\mathbf{a}})$ close to one, suppressing the gradient of the MMD with respect to the network weights during training, and a too small bandwidth decreases the “effective batch size” since most terms in the sums are very close to zero.

The first (second) terms in Eq. 1.1 and Eq. 1.2 are sums over all possible pairs of data points within the target (output) data set. For each pair, the kernel function is evaluated. The difference between the biased and unbiased estimator is that in the first case, the “diagonal elements”, pairings of a data point with itself, are included. For the third terms, which are subtracted from the sum of the first two, all output data points are paired with all target data points and the kernel function is evaluated. The result of the subtraction is a large number if the multidimensional overlap between the output and target data sets is small and a small number when the overlap is great. For the biased estimator, the *lower bound* of $\mathrm{MMD_b}$ is zero whereas for the unbiased estimator, the *expectation value* of $\mathrm{MMD_u}$ for two iid samples is zero, allowing $\mathrm{MMD_u}$ to be negative.

Furthermore, the **mean squared error** (MSE) is used as a pairwise distance measure. It requires an element-by-element matching of the output and target samples. Additionally to the standard MSE (Eq. 1.4), the Huber loss (Eq. 1.5) might be used which combines the MSE and the mean absolute error (MAE), reducing the influence of outliers.

$$\mathrm{MSE}(\theta) = \frac{1}{m} \sum_{i=1}^{m} ||\hat{\mathbf{a}}_i(\theta) - \mathbf{a}_i||^2 . \tag{1.4}$$

$$\mathrm{Huber}(\theta) = \frac{1}{m} \sum_{i=1}^{m} h_i \, , \tag{1.5}$$

$$h_i = \begin{cases} 0.5||\hat{\mathbf{a}}_i(\theta) - \mathbf{a}_i||^2 & \text{if } ||\hat{\mathbf{a}}_i(\theta) - \mathbf{a}_i|| < \delta \\ \delta \left(||\hat{\mathbf{a}}_i(\theta) - \mathbf{a}_i|| - 0.5\delta \right) & \text{otherwise} \end{cases} \qquad \text{with } \delta = \text{const} \, .$$

1.3 Training Schemes

Fundamentally, the training of the refinement network depends on whether the input and target data sets can be *matched*. This means that a given sample of the input data set has a corresponding sample in the target data set, sharing, e.g., a common *ground truth*. If this is possible, the pairwise MSE (or Huber) loss function can be meaningfully defined. Otherwise, only the ensemble-based MMD loss can be used.

This section describes various training schemes using one or multiple loss terms, introducing important algorithmic aspects and possible (dis)advantages. Figure 1.2 summarizes the training setup of the refinement network.

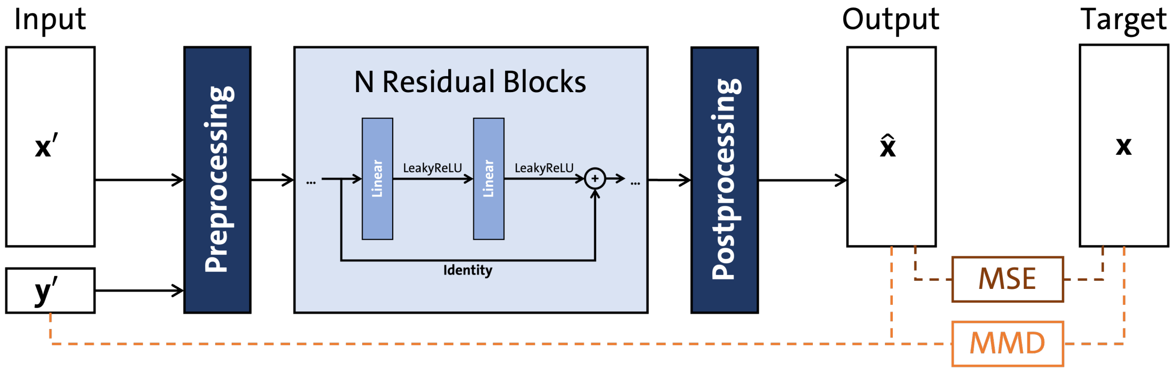

Figure 1.2: Sketch of the refinement network and the training setup. The variables to be refined are denoted by $\mathbf{x}'$ and after refinement by $\hat{\mathbf{x}}$. The loss functions MSE and MMD are used to compare the output of the network to the target $\mathbf{x}$. Additional variables $\mathbf{y}'$ can be used to condition the network and as extra dimensions in the MMD loss to include their correlations to the refined variables in the training objective.

1.3.1 Single Loss Term

Only MMD

The *vanilla* training scheme using only the MMD loss has the advantage that it is simple to implement and efficient in correcting the distribution-level agreement between the refined and target data sets. Furthermore, it does not require a matching between the input and target data set and is therefore also suitable for the refinement of a simulation output with respect to distributions of real data. For the training, the biased estimator is preferred over the unbiased estimator, since it is bounded from below at zero. A possible shortcoming is that if the input and target data sets exhibit a large discrepancy in an additional dimension which is not an input to the MMD, this discrepancy can lead to incorrect morphing in the superspace including the additional dimension.

Only MSE

Training the refinement network with only the MSE loss leads to undesired results. This is due to the independent stochastic nature of the simulation chains, which leads to many data points where a feature value far from the mean is matched to a target data point likely closer to the mean. Therefore the input sample will be modified towards less extreme values. This *regression to the mean* effect leads to an underrepresentation of samples in the tails of all feature distributions. Importantly though, deterministic biases between the input and target data sets can be corrected via the pairwise matching manifest in the MSE loss, which indirectly encodes information about correlations among refinable and hidden observables. Therefore, a pairwise loss function like the MSE can be used together with the ensemble-based MMD loss as described in Section 1.3.2.

1.3.2 Multiple Loss Terms

Training a neural network with multiple, possibly counteracting, loss terms is a multi-objective optimization which, in general, has no solution that minimizes all objective functions simultaneously. The set of all *optimal* solutions X out of all possible solutions Y defines the *Pareto front.* In the bi-objective case with loss functions f and g, a solution $\theta^\star \in X$ is considered optimal if there is no solution $\theta \in Y$ with $f(\theta) < f(\theta^\star)$ and $g(\theta) < g(\theta^\star)$. Solutions may exist with lower values of f or g but not both, as visualized in Figure 1.3. The shape of the Pareto front for a given (neural network) optimization problem is a priori unknown, it can be convex, concave, or a mixture.

Combinations of multiple loss terms are omnipresent in the ML literature, be it that models are trained for multiple tasks, weight regularization terms are added, or in the case of the generative adversarial network setup, where a generator and a discriminator network are trained simultaneously. In the training of the refinement network, the MMD and MSE can be identified as two (competing) loss functions, addressing stochastic and deterministic biases respectively. The following sections describe ways in which the two loss functions can be combined.

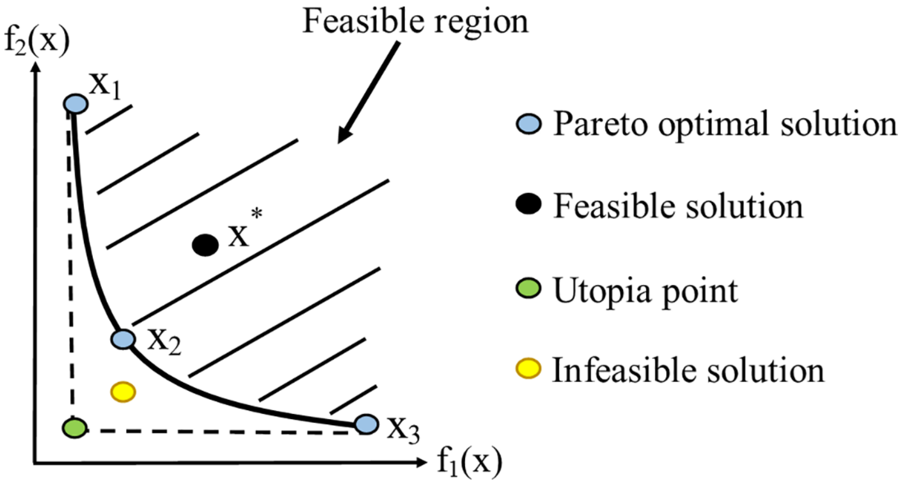

Figure 1.3: Visualization of a convex Pareto front in the bi-objective optimization with objective functions f_1 and f_2 [46].

Fixed Weights

An obvious choice when combining multiple loss terms is to use their weighted sum as the objective function. The weights may be chosen to strike a desired balance between the two loss functions. However, this choice does not necessarily allow convergence to all points on the Pareto front neither does it allow directly choosing the desired point on the Pareto front [47]. Essentially, using the sum of multiple losses with fixed weights transforms the problem into a single-loss optimization.

MDMM Algorithm

The modified differential method of multipliers (MDMM) algorithm is a solution to the multi-loss optimization problem which reformulates the network training as a constrained Lagrangian optimization [48]. Identifying a primary loss $f(\theta)$ and an additional loss

$g(\theta)$, which both are functions of the network parameters θ, the Lagrangian is written as

$$\mathcal{L}(\theta,\lambda) = f(\theta) - \lambda\left(\varepsilon - g(\theta)\right) . \tag{1.6}$$

The objective is to minimize the primary loss subject to the constraint $g(\theta) = \varepsilon$. Importantly, the Lagrange multiplier λ is not a constant weight but is updated during training via gradient *ascent* in contrast to the network parameters θ which are optimized, as usual, via gradient *descent*. Hence, the Lagrangian is minimized with respect to the network parameters and maximized with respect to the Lagrange multiplier. In order to ensure convergence, an additional damping term $\frac{\delta}{2}\left(\varepsilon - g(\theta)\right)^2$ is added to the Lagrangian. By choosing the value of ε, the desired location on the Pareto front can be converged on, as mathematically formalized in the Karush–Kuhn–Tucker conditions [49, 50].

For the application of the MDMM algorithm to the training of the refinement network, either the MMD or the MSE are treated as the primary loss. Both choices are elaborated in the following.

Primary Loss MMD Choosing the MMD as the primary loss while placing a constraint on the MSE is well motivated since the aim of the refinement is the agreement of the output and target *distributions*. However, the corresponding target value for the MSE is not apriori calculable, and typically requires empirical tuning.

Primary Loss MSE Using the MSE as the primary loss and the unbiased estimator of the MMD as the constrained loss functions comes with the advantage that the target value ε has a natural choice which is zero, the expectation value for two iid samples. It is important to choose an appropriate starting value of the Lagrange multiplier $\lambda \neq 0$ such that the influence of the MSE at the start of the training does not lead to an overcorrection of values in the tails but is balanced by the MMD constraint from the beginning.

1.3.3 Two-Stage Training

By combining two of the previously introduced training schemes, the performance of the refinement network can be further improved [1]. In the first stage, the MDMM algorithm using MSE as the primary loss and the constraint $\mathrm{MMD_u}(\theta) = 0$, combining the advantages of both loss terms. After convergence of the algorithm, to fine-tune the network, only the biased estimator of the MMD ($\mathrm{MMD_b}$) is used. This ensures that for the second stage, no “pressure” from the MSE is applied on the tails.

2 Applications

2.1 Refinement of Jet Flavor Tagging Observables Using CMS FastSim

The refinement methodology has been applied to the CMS FastSim chain to improve its agreement with the FullSim output. The results are presented in this chapter and have been published in Ref. [2], where further details can be found.

The simulated data set consists of events in which pairs of gluinos, supersymmetric partners of gluons, are produced. Each gluino subsequently decays to a top quark pair and a neutralino. This corresponds to the supersymmetric simplified model T1tttt [51]. The generator-level events are processed twice, once using FastSim (the input data set to be refined) and once with FullSim (the target data set). The features to be refined are four jet flavor tagging observables:

$$\mathbf{x} = \big(\mathrm{B}, \quad \mathrm{CvB}, \quad \mathrm{CvL}, \quad \mathrm{QG}\big)^{\intercal} \ .$$

They are computed for each jet with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$ from the output of the DeepJet algorithm [52], a multiclass neural network with six output nodes, activated with a softmax function. The nodes correspond to jets containing hadronically (leptonically) decaying b hadrons — labeled b (lepb), jets containing two b hadrons (bb), and jets from c quarks (c), light quarks (uds), and gluons (g). From those values, the four discriminator observables are defined as

$$\mathrm{B} = \mathrm{b} + \mathrm{bb} + \mathrm{lepb}\,, \ \mathrm{CvB} = \frac{\mathrm{c}}{\mathrm{c} + \mathrm{b} + \mathrm{bb} + \mathrm{lepb}}\,, \ \mathrm{CvL} = \frac{\mathrm{c}}{\mathrm{c} + \mathrm{uds} + \mathrm{g}}\,, \ \mathrm{QG} = \frac{\mathrm{g}}{\mathrm{g} + \mathrm{uds}}\,.$$

The discriminators are transformed using a logit-transformation and an additional layer is added to the postprocessing block of the refiner network which ensures that the refinement preserves the sum of the DeepJet output nodes (which is unity due to the softmax activation). Additional conditioning variables are the true transverse momentum of the simulated jet $p_{\mathrm{T}}^{\mathrm{GEN}}$, pseudorapidity η^{GEN}, and the flavor of the underlying hadron.

In this application, the refiner network is trained using the MDMM algorithm with the biased estimator of the MMD as the primary loss and a constraint on the Huber loss.

Multiple values of the corresponding target value ε are scanned. The evolutions of the two losses are plotted in Fig. 2.1 along with training versions using only $\mathrm{MMD_b}$, only Huber, or a sum of both losses with fixed weights of one. As expected, the version using only the Huber loss leads to a smaller average disagreement between the matched jet-jet pairs (on the vertical axis) but clearly to a worsening in the distribution-level agreement (on the horizontal axis). Also, using a simple sum of $\mathrm{MMD_b}$ and Huber does not lead to improved distribution-level agreement. However, using only the $\mathrm{MMD_b}$ loss or combining the two losses via the MDMM algorithm leads to an improvement in both loss dimensions.

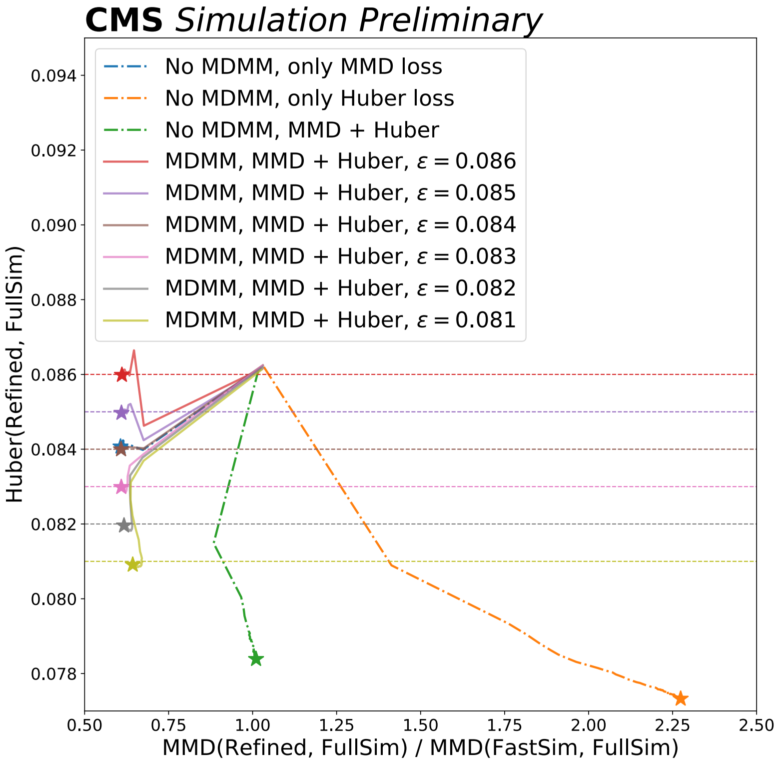

Figure 2.1: Evolution of the two loss terms, $\mathrm{MMD_b}$ and Huber, for different trainings of the refiner network with and without the MDMM algorithm. The values of $\mathrm{MMD_b}$ are normalized to the baseline value comparing unrefined FastSim and FullSim, such that values smaller than 1 indicate improved agreement between the distributions. The Pareto front is observed to exhibit a convex shape [2].

Figure 2.2 shows the distributions of the four DeepJet discriminators for the FullSim data set as well as for the unrefined FastSim and the refined FastSim using the network trained with the MDMM algorithm, setting $\varepsilon = 0.084$. The improvement in the agreement between FastSim and FullSim after refinement is evident. Furthermore, correlations within the set of (refined) observables and additional truth-level features are examined by evaluating the Pearson correlation coefficients as shown in Fig. 2.3. Also by these measures, the accuracy of the FastSim output is improved by the refinement.

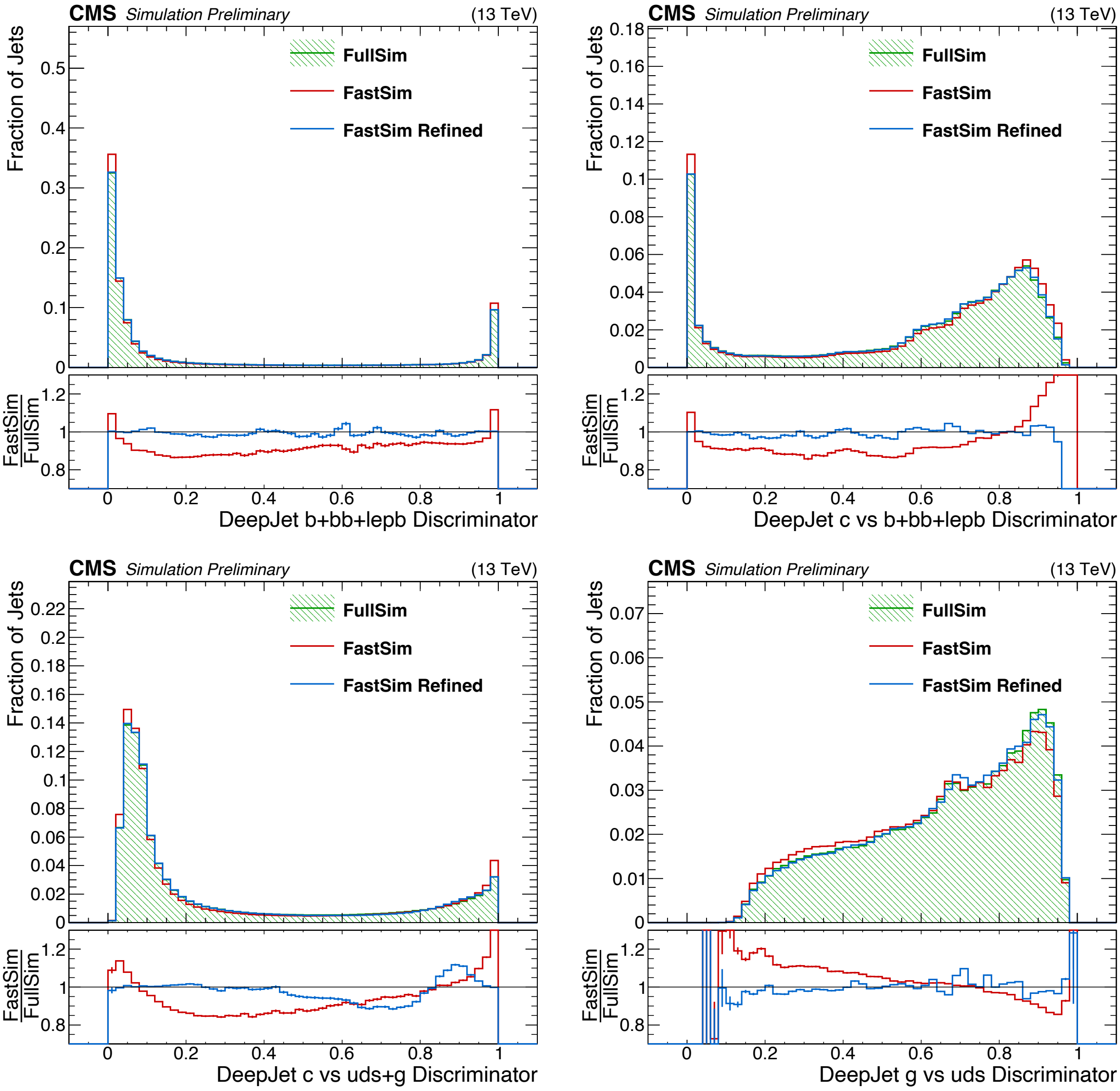

Figure 2.2: Distributions of the four DeepJet discriminators B (upper left), CvB (upper right), CvL (lower left), and QG (lower right) for FullSim, FastSim, and refined FastSim [2].

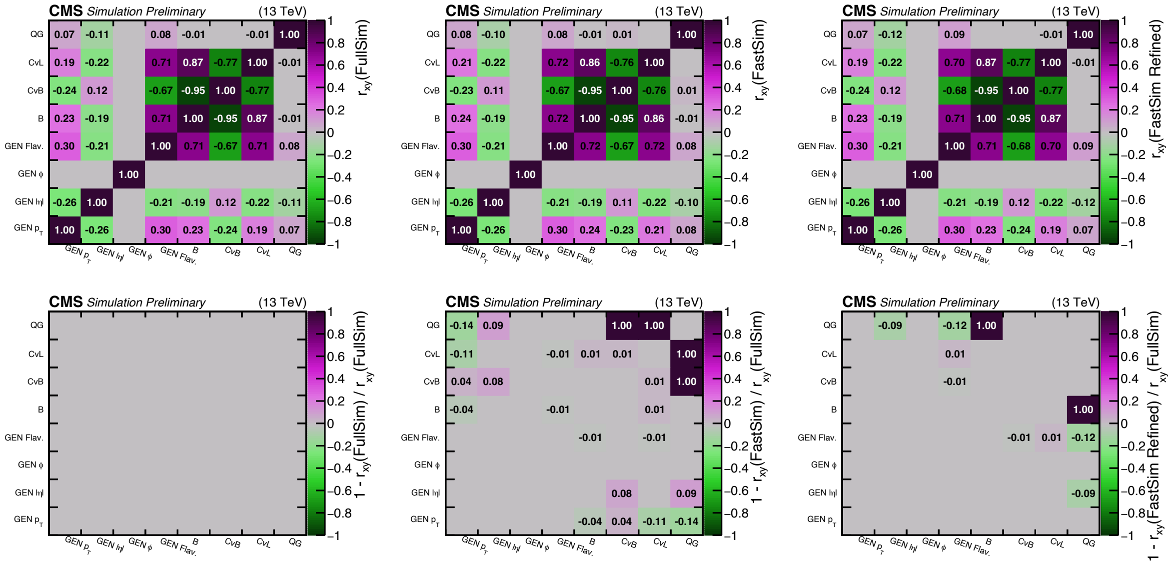

Figure 2.3: Upper row: Pearson correlation coefficients for a set of jet observables for FullSim (left), FastSim (center), and refined FastSim (right). Lower row: relative difference between each correlation factor and the corresponding FullSim value. All entries in the bottom left plot are zero by construction and a relative difference of 100% indicates cases where the correlation factor for FullSim or for (refined) FastSim, as shown in the upper plot, is zero [2].

2.2 Refinement of Jet Substructure Observables Using Delphes Simulation

A study has been performed using Delphes [53], a parametric detector simulation framework, to showcase the application of the refinement method to jet substructure observables. A ground truth data set based on the pair production of SM top quarks in proton-proton collisions is processed twice with Delphes: once with the default CMS detector implementation yielding the data set that is treated as FullSim and once with a manually flawed implementation emulating a FastSim data set. The features to be refined are chosen to be three ratios of the N-subjettiness τ_N [54], variables used to establish information about the initial particle from which a given jet originated.

The two-stage training as described in Section 1.3.3 is used to train the refiner network. The corresponding learning curves can be seen in Fig. 2.4, where the switch from the first to the second training stage is visible after 118 epochs. At this point, $\mathrm{MMD_u}$ has converged to its target value of $\varepsilon = 0$ and the corresponding Lagrange multiplier λ does not change anymore. In the second stage, by removing MSE as a training objective, the MMD estimators can be further minimized. The figure also shows the evolution of a metric called *omniscient* MMD which includes not only the refined dimensions but also additional (hidden) observables. The omniscient MMD is not used for the training of the network but for evaluation purposes, containing information about correlations between the refined and hidden observables. The hidden observables include the jet mass, p_{T}, η, the distance to the closest neighbor jet $\mathrm{dR} = \sqrt{\mathrm{d}\eta^2 + \mathrm{d}\varphi^2}$ (with azimuthal angle φ), and the numbers of charged and neutral jet constituents $N(\mathrm{ch})$ and $N(\mathrm{ne})$.

The results of applying the refiner network to the input data set are presented in Figs. 2.5 and 2.6. A clear improvement in the agreement with the target data set can be observed. These results have been published as a major article [1] containing further details.

2.3 Refinement of Simulation to Data

In both of the previously shown applications, the refinement method is employed to refine the output of a fast simulation chain with respect to a more precise full simulation output. Another well-motivated use case of the method is the refinement of a simulated sample with respect to distributions observed in real-world data. This is realized in Part II of this work (see Section 6.2.2) in the context of a search for new physics using data recorded by the CMS experiment at the LHC.

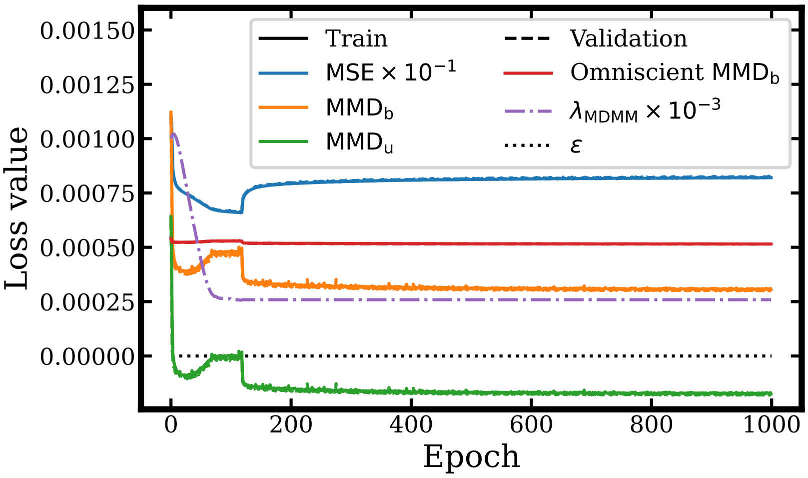

Figure 2.4: Evolution of the MSE, (omniscient) MMD estimators, and Lagrange multiplier λ during training of the refiner network used in the Delphes application. The switch from the first to the second training stage is visible after 118 epochs [1].

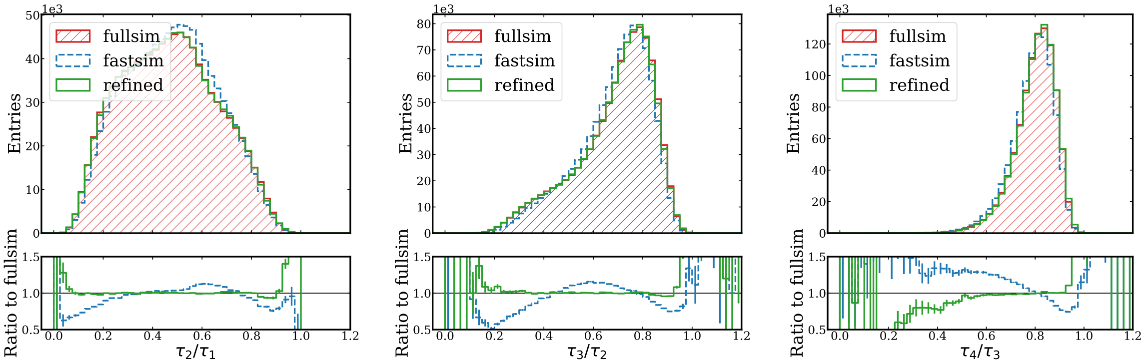

Figure 2.5: Marginal distributions of the three N-subjettiness ratios for the target ("fullsim"), input ("fastsim"), and refined data sets. The lower panels show the relative differences of bin counts to the target distribution [1].

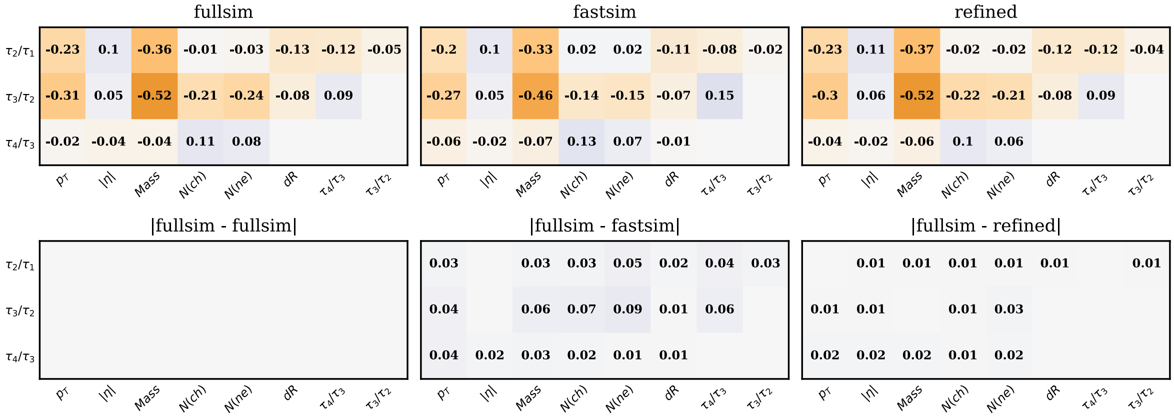

Figure 2.6: Upper row: Pearson correlation coefficients for a set of jet observables for the target “fullsim” (left), input “fastsim” (center), and refined (right) data sets. Lower row: relative difference between each correlation factor and the corresponding target value [1].

3 Summary and Outlook

In this Part, a method to refine output observables of a simulation chain with respect to a target distribution was introduced and illuminated. It makes use of a regression neural network, trained to provide a residual correction (morphing) to the input data set and trained with a dedicated ensemble-based loss function — the maximum mean discrepancy (MMD). The MMD can be complemented by a pair-based loss function like the MSE, where the combination of the two objective functions is realized by the modified differential method of multipliers (MDMM), assuring convergence on the Pareto front.

Two applications have been presented in which the data set to be refined corresponds to the output of *fast* simulation applications, which is modified to better agree with the output of a more accurate, yet also more resource-intensive, *full* simulation program. In both cases, a substantial improvement of the modeling, also regarding correlations to hidden features, is shown. This concept of using machine learning for the refinement of available physics-informed simulations is a promising way forward facing increasing computational challenges in science.

The impact of the choice of kernel function in the MMD loss remains an open topic. For example, using another radial basis function like the inverse multiquadric kernel, as suggested in Ref. [55] due to its heavier tails, might be beneficial for the refinement. Furthermore, combining the refinement approach with a weight-based technique could lead to an improved performance since data points could be weighted to directly correct for the mismodeling of (reconstruction) efficiencies additionally to the refinement correction which shifts the data point within the feature space.

Part II

Search for Natural Supersymmetry with Low-Momentum and Displaced Tracks

The search for physics beyond the Standard Model (BSM) is one of the key objectives of experiments at the LHC. A particularly well-motivated class of BSM scenarios is based on the concept of natural supersymmetry, featuring higgsino-like charginos and neutralinos, discussed in Section 1.3 of the Introduction. In this Part, a search for such models, performed at the CMS experiment, is presented. The crucial element of the analysis is the usage of low-momentum, isolated, and displaced tracks to gain sensitivity to the signal. After giving a more detailed description of the signal model and the analysis strategy in Chapter 1, the data sets and physics objects used are introduced in Chapters 2 and 3. The multiclass neural network employed to distinguish signal from background tracks is explained in Chapter 4, followed by the selection of events and the definition of signal regions in Chapter 5. To obtain a reliable estimate of the SM background yields in the signal regions, dedicated corrections are applied to the simulated MC samples, including the application of the refinement methodology introduced in Part I, as explained in Chapter 6. Based on the systematic uncertainties compiled in Chapter 7, the results are presented in Chapter 8. The search excludes previously unexplored phase space regions with chargino masses up to 180 GeV.

1 Compressed Higgsinos

As explained in Section 1.3 of the Introduction, naturalness arguments call for the higgsino mass parameter in supersymmetric extensions of the SM to be on the order of the Z boson mass. This leads to the configuration that the three lightest electroweakinos $\widetilde{\chi}_2^0$, $\widetilde{\chi}_1^\pm$, and $\widetilde{\chi}_1^0$ (the lightest supersymmetric particle, LSP) are higgsino-like and thus referred to as higgsinos. The mass differences between the higgsinos $\Delta m^\pm \equiv m(\widetilde{\chi}_1^\pm) - m(\widetilde{\chi}_1^0)$ and $\Delta m^0 \equiv m(\widetilde{\chi}_2^0) - m(\widetilde{\chi}_1^0)$ are small — on the order of 1 GeV— characterizing the *compressed* mass spectrum. Corresponding to the limit of large $\tan\beta$ in Ref. [56], the relation $\Delta m^0 = 2\Delta m^\pm$ holds.

If R-parity is conserved, the production of such higgsinos happens in pairs and $\widetilde{\chi}_2^0$ and $\widetilde{\chi}_1^\pm$ decay to $\widetilde{\chi}_1^0$ and a highly virtual Z or W boson, respectively, see Fig. 1.1. As described in Ref. [56] and [57], the dominant decay of the higgsino-like chargino $\widetilde{\chi}_1^\pm$ is to the LSP and, via an off-shell W boson, a single pion (as shown in Fig. 1.2). The corresponding decay length can be calculated as

$$c\tau(\widetilde{\chi}_1^\pm \to \widetilde{\chi}_1^0\pi^\pm) = 1.1\,\mathrm{cm}\left(\frac{\Delta m^\pm}{300\,\mathrm{MeV}}\right)^{-3}\left(1 - \frac{m(\pi^\pm)^2}{\Delta m^{\pm 2}}\right)^{-1/2} \tag{1.1}$$

and the associated total width is plotted in Fig. 1.2. It can be seen that a mass splitting $\Delta m^\pm \lesssim 1$ GeV can lead to a measurable decay length of the chargino.

1.1 Previous Searches

In searches for electroweakinos with such a compressed mass spectrum, the analysis strategy depends on the specific mass splittings, as visualized in Fig. 1.3. If Δm^0 is larger than ≈ 1 GeV soft, prompt leptons from the decay of the second-lightest neutralino ($\widetilde{\chi}_2^0 \to \widetilde{\chi}_1^0 Z^* \to \widetilde{\chi}_1^0 l^+ l^-$) can be used to gain sensitivity to the signal phase space [58]. On the other hand, if $\Delta m^\pm$ is smaller than ≈ 0.3 GeV, the chargino becomes sufficiently long-lived to reach the third or fourth layer of the tracking detector before decaying, which can give rise to a disappearing track signature [59].

Exclusion limits set by analyses performed by the ATLAS collaboration using those signatures can be seen in Fig. 1.4 as blue and green shaded areas. Searches using these

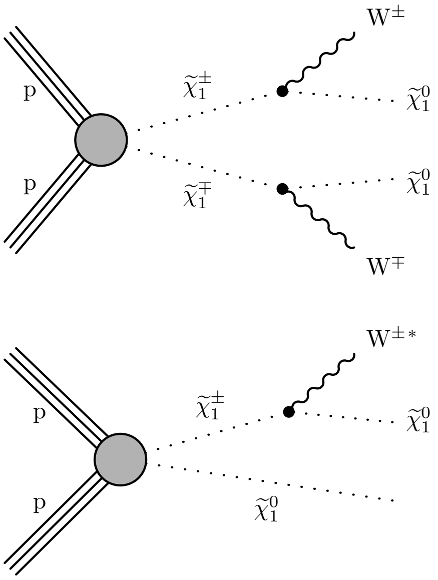

Figure 1.1: Feynman diagrams for electroweakino pair production. Possible processes include the production of two charginos (top left), one chargino along with the second-lightest (top right) or lightest neutralino (bottom left), as well as two neutralinos (bottom right).

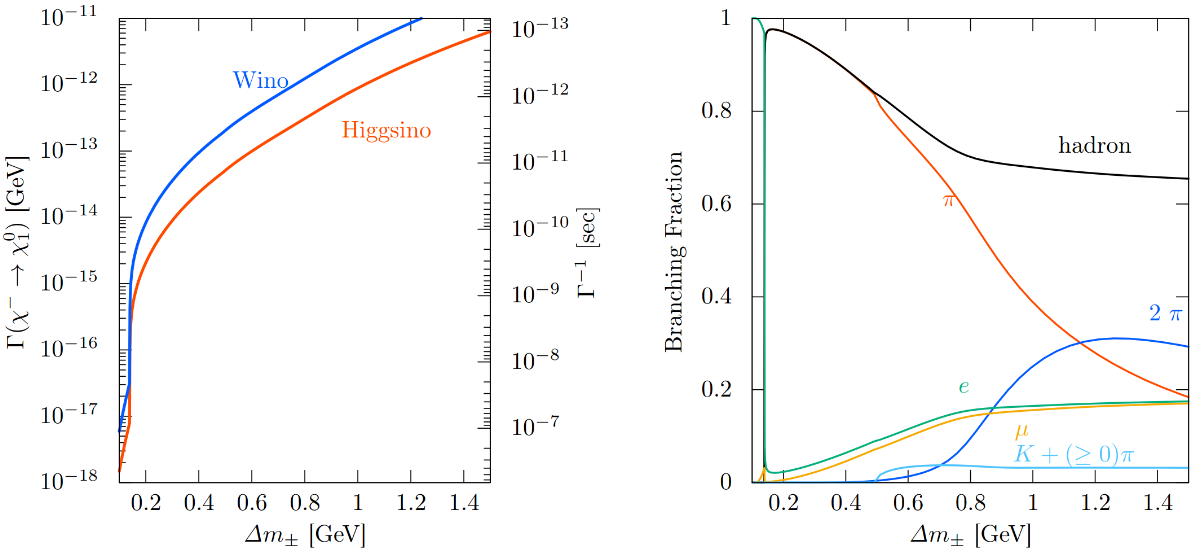

Figure 1.2: Total width (left) and branching fractions (right) of the higgsino-like chargino versus $\Delta m^{\pm}$ [57]. The total width is also shown for the case of a wino-like chargino.

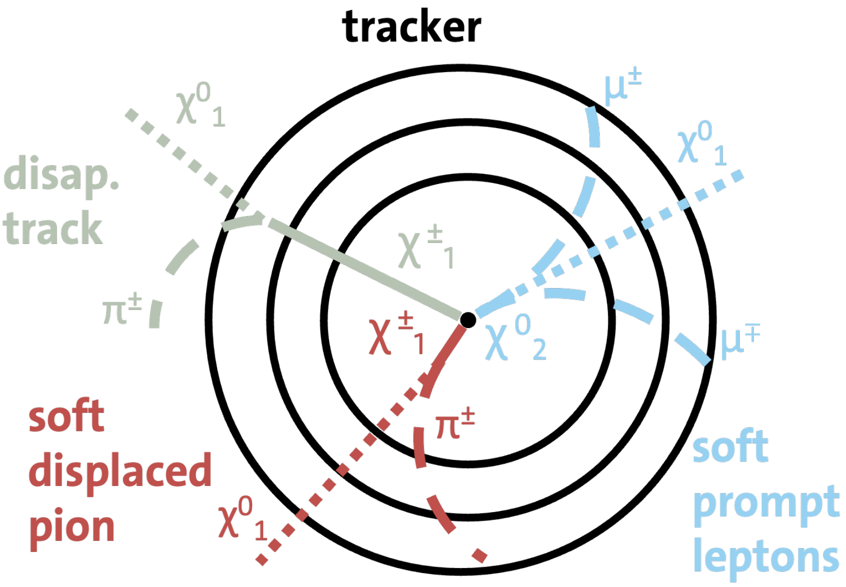

Figure 1.3: Sketch of detector signatures that can be used to search for electroweakinos $\widetilde{\chi}_2^0$, $\widetilde{\chi}_1^\pm$, and $\widetilde{\chi}_1^0$ in different regimes of mass splittings $\Delta m(\widetilde{\chi}_1^\pm, \widetilde{\chi}_1^0)$ and $\Delta m(\widetilde{\chi}_2^0, \widetilde{\chi}_1^0)$. This analysis targets the soft displaced pion signature shown in red.

strategies have also been recently carried out by CMS in final states with disappearing tracks [60] and with pairs of very soft leptons [61]. Exclusion limits derived in the model phase space are shown in Fig. 1.5.

However, the intermediate region, $0.3\,\text{GeV} \lesssim \Delta m^{\pm} \lesssim 1\,\text{GeV}$, cannot be probed by either of these methods and is the target of the search presented in this work. The corresponding analysis strategy is presented in the following.

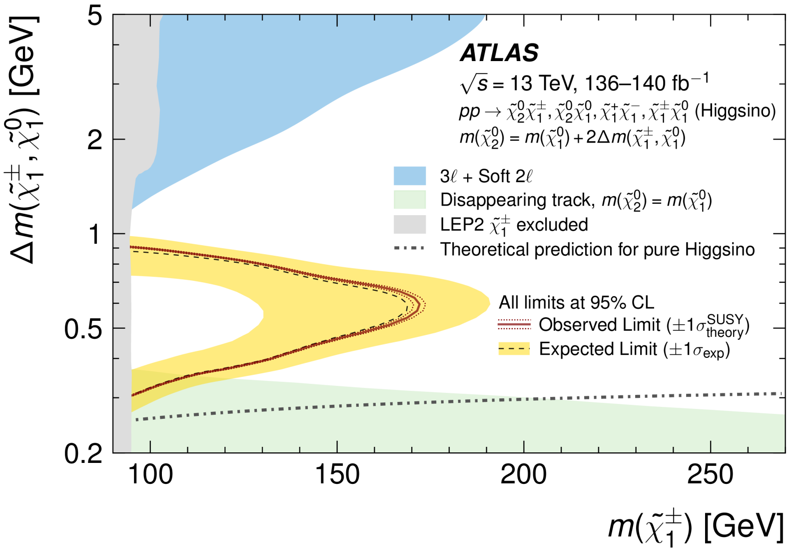

Figure 1.4: Exclusion limits set by the ATLAS Collaboration in the plane of chargino mass and mass splitting between the chargino and the lightest neutralino assuming the higgsino simplified model for different search strategies. The red line corresponds to the result of the soft displaced pion analysis [62].

1.2 Analysis Strategy

This analysis targets natural SUSY models with a compressed higgsino mass spectrum, $0.3\,\text{GeV} \lesssim \Delta m(\widetilde{\chi}_1^{\pm}, \widetilde{\chi}_1^0) \lesssim 1\,\text{GeV}$, using events in which one or two charginos $\widetilde{\chi}_1^{\pm}$ are produced (see Feynman diagrams in Fig. 1.1). After a short but discernible decay length up to $O(1\,\text{cm})$, the charginos decay to the lightest neutralino $\widetilde{\chi}_1^0$ and an off-shell W boson. The W boson in turn most often decays to a single pion $\pi^{\pm}$ which is the final state on which this analysis is focused. The energy transferred to the pion is very small and corresponds to the difference in masses $\Delta m(\widetilde{\chi}_1^{\pm}, \widetilde{\chi}_1^0)$, labeled $\Delta m^{\pm}$.

The signature of such events in the detector is dominated by a large amount of missing transverse energy if the higgsinos recoil against an initial-state radiation jet. This is due

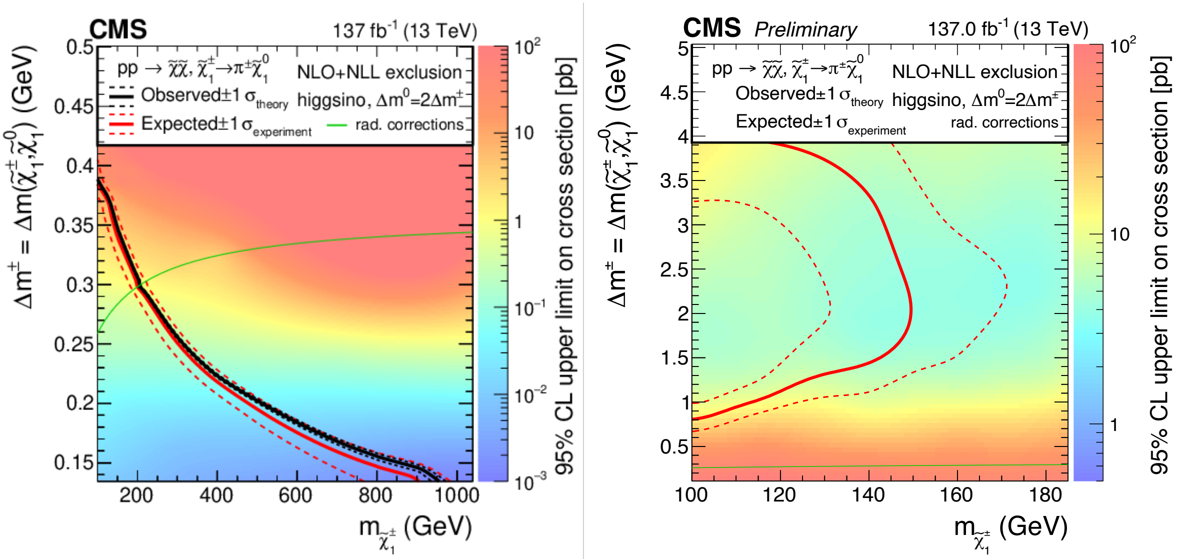

Figure 1.5: Observed and expected limits obtained in the Higgsino model phase space by the analysis of disappearing tracks [60] (left) and soft lepton pairs [61] (right).

to the neutralinos $\tilde{\chi}_1^0$ leaving the detector without depositing their energy. The other signal feature is the pion, which can be reconstructed as a significantly bent track and thus be used to distinguish signal events from SM background. This track can be used to establish sensitivity to the model phase space region, as illustrated by the expected exclusion limits in Fig. 1.6 and the observed limits set by the ATLAS Collaboration shown in Fig. 1.4 as a red line.

In this analysis, the most signal-like track in an event is selected using a neural network classifier (see Section 4) and events are subsequently categorized into various signal regions after passing the baseline “monojet”-like event selection requiring the presence of a high-momentum jet and an imbalance in the transverse momenta of all reconstructed objects (see Section 5.2).

To pass the baseline event selection, reconstructed events necessarily have to exhibit a large magnitude of missing transverse momentum p_T^{miss}. Thus, the SM processes leading to such events primarily feature neutrinos which escape the detector without depositing their energy. Consequently, the dominant SM background to this search is the “invisible” Z boson decay $Z \to \nu\nu$ (labeled “Z(inv)Jets”). The sub-dominant background comes from the $W \to l\nu$ process which is suppressed by vetoing on the presence of leptons. Further minor backgrounds, making up in total less than 1% of all events in the signal regions, include processes in which (pairs of) top quarks (labeled “TTJets/ST” where “ST” stands for single-top) or two gauge bosons WW, WZ, or ZZ (“Diboson”) are produced. These processes constitute the *event-level* background to

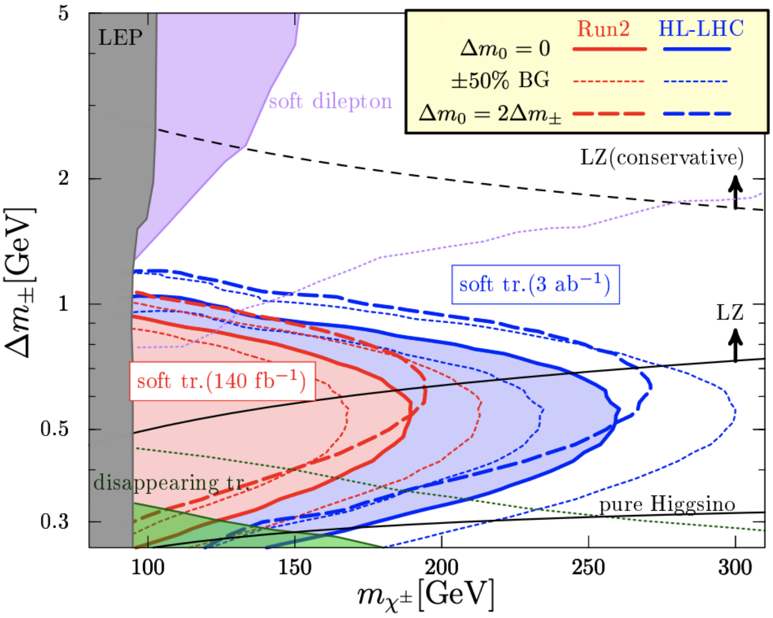

Figure 1.6: Expected exclusion limits for the "T" ($\Delta m^0 = 0$) and "Sandwich" ($\Delta m^0 = 2\Delta m^\pm$) configurations based on phenomenological studies of soft displaced track signature [63].

this search.

In each of those events, numerous soft — and often displaced — tracks are present that can fake the signature of the signal pion. Those tracks are mostly not originating from the hard process but rather stem from the underlying event, pileup interactions, and errors in the reconstruction (fakes). Therefore, the *track-level* background is largely independent of the event-level SM process. However, there are correlations of the track kinematics with the overall event topology, e.g., the direction of the track with respect to $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$.

The expected background event yields in the signal regions are estimated using MC simulations corrected by data (see Section 6). The analysis strategy is summarized in Fig. 1.7.

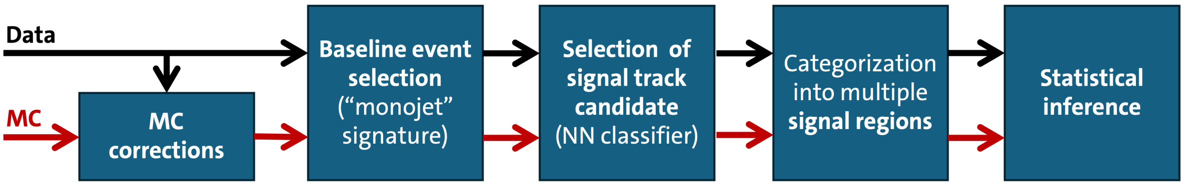

Figure 1.7: Steps involved in the processing of events for data and simulation (MC).

2 Data Sets

2.1 Collected Data

The data sets used in this analysis correspond to the so called `MET` and `SingleMuon` data streams collected by the CMS experiment during the LHC Run 2 in the years 2016, 2017, and 2018. The ultra-legacy reconstruction is used and only events within certified "lumi sections" are considered. Events from the `MET` data stream generally exhibit a large magnitude of missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$ and are therefore used to populate the signal regions whereas the `SingleMuon` data stream is used to construct a control region by selecting $\mathrm{Z} \to \mu\mu$ events, see Section 6.1. The integrated luminosities for each year amount to $36.3\,\mathrm{fb}^{-1}$, $41.5\,\mathrm{fb}^{-1}$, and $59.8\,\mathrm{fb}^{-1}$ for 2016, 2017, and 2018, respectively. Tables A.1 and A.2 in the Appendix list the detailed data set names.

Trigger

For both data streams, a logical OR of the following triggers is used which all use the presence of large missing transverse energy to identify events:

- `HLT_PFMETX_PFMHTX_IDTight_v*` (`X` = 90, 100, 110, 120, 130, 140),
- `HLT_PFMETX_PFMHTX_IDTight_PFHT60_v*` (`X` = 100, 110, 120, 130, 140),
- `HLT_PFMETNoMuX_PFMHTNoMuX_IDTight_v*` (`X` = 90, 100, 110, 120, 130, 140),
- `HLT_PFMETNoMuX_PFMHTNoMuX_IDTight_PFHT60_v*` (`X` = 100, 110, 120, 130, 140).

By requiring an offline cut on $p_{\mathrm{T}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$ of 300 GeV (see Section 3 for definitions of the observables and Section 5.2 for a detailed description of the offline selection cuts), the majority of SM background can be rejected and the trigger turn-on avoided with negligible loss in sensitivity. This family of triggers has been studied in Ref. [64] and is found to be 99% efficient for the above-mentioned offline selection, as can be seen in Fig. 2.1.

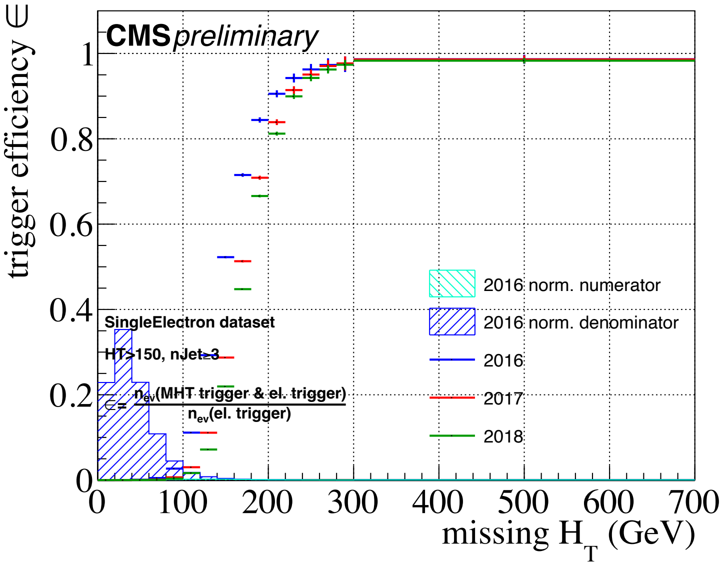

Figure 2.1: Efficiency of the set of $p_{\mathrm{T}}^{\mathrm{miss}}$-$H_{\mathrm{T}}^{\mathrm{miss}}$ cross triggers versus $H_{\mathrm{T}}^{\mathrm{miss}}$ [64].

2.2 Standard Model Backgrounds

Background samples were simulated centrally by the CMS Collaboration and are based on FullSim. They include simulated Z $\rightarrow \nu\nu$, W $\rightarrow$ lν, top quark (pair) production, and WW, WZ, and ZZ production. Also, simulated Z $\rightarrow$ ll events are used for the Z $\rightarrow \mu\mu$ control region. The primary names of the MC background data sets are listed in the Appendix in Table A.3 along with the respective cross section. Table A.4 lists the processing strings for each of the four eras.

2.3 Signal

Signal samples are produced privately using FastSim and FullSim. The simulation has been carried out by Samuel Bein. Event generation is performed using Pythia8.240 [65]. An admixture of all production modes, including neutralino-chargino, chargino-chargino, and neutralino-neutralino are generated simultaneously by setting the Pythia configuration parameter `susy:all = on`. Leading order NNPDF 3.1 PDFs are used during event generation. The production cross sections are computed at NLO plus next-to-leading-log (NLL) precision in a limit of mass-degenerate higgsino $\widetilde{\chi}_2^0$, $\widetilde{\chi}_1^{\pm}$, and $\widetilde{\chi}_1^0$ with all the other superpartners assumed to be heavy and decoupled [66–68].

Table 2.1 lists the simulated signal model points and their cross sections. The model properties are based on Ref. [56] and [57] and further discussed in Section 1.

Table 2.1: Simulated signal model points and cross sections. The second value of $\Delta m^{\pm}$ in each row (bold) corresponds to the case of pure higgsinos with the mass splitting arising only from radiative corrections from the SM particles.

$m(\widetilde{\chi}_1^{\pm})$ (GeV)	$\Delta m^{\pm}$ (GeV)	σ (fb)
100	0.16, **0.26**, 0.36, 0.46, 0.56, 0.76, 0.96, 1.26, 1.76	16797.21
115	0.17, **0.27**, 0.37, 0.47, 0.57, 0.77, 0.97, 1.27, 1.77	10833.94
140	0.18, **0.28**, 0.38, 0.48, 0.58, 0.78, 0.98, 1.28, 1.78	5166.30
160	0.19, **0.29**, 0.39, 0.49, 0.59, 0.79, 0.99, 1.29, 1.79	3109.02
180	0.19, **0.29**, 0.39, 0.49, 0.59, 0.79, 0.99, 1.29, 1.79	2040.76
200	0.20, **0.30**, 0.40, 0.50, 0.60, 0.80, 1.00, 1.30, 1.80	1335.61
250	0.21, **0.31**, 0.41, 0.51, 0.61, 0.81, 1.01, 1.31, 1.81	577.31
300	0.22, **0.32**, 0.42, 0.52, 0.62, 0.82, 1.02, 1.32, 1.82	284.86
500	0.22, **0.32**, 0.42, 0.52, 0.62, 0.82, 1.02, 1.32, 1.82	33.85

3 Physics Objects

This section describes the reconstructed objects used in this analysis. They are listed in Tables 3.1 and 3.2 along with their respective phase space cuts and isolation definitions. The isolation measures, used to purify the object selections, are calculated in the following ways for photons, muons, and electrons:

$$\begin{aligned}
\mathrm{Iso}_\gamma &= \texttt{chargedHadronIso} + \texttt{neutralHadronIso} + \texttt{photonIso}\,,\\
\mathrm{Iso}_\mu &= \texttt{chargedHadronIso} + \max(0, \texttt{neutralHadronIso} + \texttt{photonIso} - 0.5 * \texttt{sumPUPt})\,,\\
\mathrm{Iso}_\mathrm{e} &= \texttt{chargedHadronIso} + \max(0, \texttt{neutralHadronIso} + \texttt{photonIso} - 0.5 * \texttt{sumPUPt})\,,
\end{aligned}$$

where `chargedHadronIso/neutralHadronIso/photonIso/sumPUPt` refers to the sum of transverse energies of charged hadrons associated to the primary vertex/neutral hadrons/photons/charged hadrons associated to pileup vertices in a cone of $\Delta R < 0.3$ around the corresponding object.

More details on the general object reconstruction procedure in CMS including the particle flow (PF) algorithm can be found in Section 2.2.1 of the Introduction.

All objects except for tracks are standard objects with centrally provided selection criteria used in CMS analyses. The preselection for tracks includes quality criteria, e.g., on the goodness of the track fit, and the requirement that the track is associated to a reconstructed Particle Flow candidate. A cut on the pseudorapidity restricts tracks to the acceptance of the tracking detector, whereas cuts on the transverse momentum, the longitudinal impact parameter, and the distance to the closest jet reject tracks which are irrelevant to the signal model phase space. Sections 3.1 and 4 describe the further treatment of tracks beyond this preselection.

Jets

Jet energy corrections (JEC) are applied in data and MC, whereas for MC, the jet energy resolution (JER) is also corrected using centrally provided correction maps. To tag jets as originating from b quarks, the DeepCSV algorithm [69, 70] with loose working point is employed. This is used to veto events with such jets in order to suppress SM background events including top quarks ($\mathrm{t} \rightarrow \mathrm{Wb}$).

Table 3.1: Definition of physics objects used in this analysis - Part 1

	Tracks	Jets	Taus
CMSSW collection	`generalTracks`	`ak4PFJetsCHS`	`hpsPFTauProducer`
Selection criteria	`highPurity` quality && `isPfCand`	`TightLepVeto` ID	`Tight` ID
Phase space cuts	$p_\mathrm{T} < 20\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.4$, $\lvert\mathrm{dz}^\mathrm{PV}\rvert < 1\,\mathrm{cm}$	$p_\mathrm{T} > 30\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.4$	$p_\mathrm{T} > 20\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.3$
Isolation definition	$\Delta R(\text{closest jet}) > 0.4$	—	—

Table 3.2: Definition of physics objects used in this analysis - Part 2

	Photons	Electrons	Muons
CMSSW collection	`gedPhotons`	`gedGsfElectrons`	`muons`
Selection criteria	Cut-based loose ID	Cut-based veto ID	`PFMuon` && (`GlobalMuon` || `TrackerMuon`)
Phase space cuts	$p_\mathrm{T} > 15\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.5$	$p_\mathrm{T} > 10\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.5$	$p_\mathrm{T} > 10\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.4$
Isolation definition	$\mathrm{Iso}_\gamma/p_\mathrm{T} < 0.2$	$\mathrm{Iso}_\mu/p_\mathrm{T} < 0.2$	$\mathrm{Iso}_\mathrm{e}/p_\mathrm{T} < 0.2$

The following event-level observables are defined using the reconstructed and corrected jets:

$$
\begin{aligned}
H_\mathrm{T}^\mathrm{miss} &= \Big| \sum_{\text{jet}\in\{\text{jets with } p_\mathrm{T}>30\,\mathrm{GeV} \text{ and } |\eta|<5.0\}} \vec{p}_\mathrm{T}(\text{jet}) \Big| \\
H_\mathrm{T} &= \sum_{\text{jet}\in\{\text{jets with } p_\mathrm{T}>30\,\mathrm{GeV} \text{ and } |\eta|<2.4\}} |p_\mathrm{T}(\text{jet})| \\
H_\mathrm{T}^5 &= \sum_{\text{jet}\in\{\text{jets with } p_\mathrm{T}>30\,\mathrm{GeV} \text{ and } |\eta|<5.0\}} |p_\mathrm{T}(\text{jet})| \,.
\end{aligned}
$$

Missing Transverse Momentum

The missing transverse momentum $\vec{p}_\mathrm{T}^\mathrm{miss}$ (and its magnitude $p_\mathrm{T}^\mathrm{miss}$) is calculated as the negative vectorial sum of the transverse momenta of all reconstructed PF candidates (corresponding to the `pfMet` CMSSW collection). Jet energy corrections are propagated to $\vec{p}_\mathrm{T}^\mathrm{miss}$ (Type-1 correction) and “MET filters” are applied to veto noisy events (see Section 5.1).

Pileup Reweighting

Simulated events are reweighted according to the number of true pileup interactions in order to better match the distribution observed in data using centrally provided weights.

3.1 Soft and Displaced Tracks

This section describes the treatment of the central element of this analysis — soft and displaced tracks. Importantly, for single displaced tracks, standard requirements on the track's association to the primary vertex or a secondary decay vertex cannot be used. Instead, the fitted tracks are extrapolated along their helical path using a custom helix parameterization. This enables improved matching of soft and displaced reconstructed tracks to their corresponding generator-level particles and allows for the definition of additional track observables.

Helix Extrapolation

Tracks with assigned charge q, azimuthal angle φ_0, polar angle θ (corresponding to dip angle $\lambda = \frac{\pi}{2} - \theta$ and pseudorapidity $\eta = -\ln(\tan(\frac{\theta}{2}))$), and momentum $\vec{p}$ (with magnitude $p = |\vec{p}|$ and thus transverse momentum $p_\mathrm{T} = p\cos(\lambda)$), are extrapolated from their reference point (x_0, y_0, z_0), which is the point of closest approach to the center of the CMS detector, along a helical path parameterized by t. The parametrization is:

$$\begin{aligned} x(t) &= x_0 + R_1 \cos(\lambda) p q^{-1} \left(\sin(\varphi_0) - \sin(\varphi_0 - t)\right) \\ y(t) &= y_0 + R_1 \cos(\lambda) p q^{-1} \left(\cos(\varphi_0 - t) - \cos(\varphi_0)\right) \\ z(t) &= z_0 + R_1 \sin(\lambda) p q^{-1} t\,. \end{aligned} \tag{3.1}$$

The constant R_1 corresponds to the radius of the track of a particle with charge $q = 1e$ and $p_\mathrm{T} = 1\,\mathrm{GeV}$ in a magnetic field of $3.8\,\mathrm{T}$,

$$R_1 = \frac{p_\mathrm{T}}{q \cdot B} = \frac{1\,\mathrm{GeV}c^{-1}}{1e \cdot 3.8\,\mathrm{T}} \approx 87.78\,\mathrm{cm}\,.$$

Along this path, both the curvature (and therefore p_T) and the pseudorapidity are constant. The azimuthal angle changes according to $\varphi(t) = \varphi_0 - t$. This parameterization assumes a constant magnetic field, which is a good approximation in the region close to the luminous region within the beam line.

Track Matching

It is found that the matching of tracks to generator-level (GEN) objects using the distance measure $\Delta R = \sqrt{\Delta\varphi^2 + \Delta\eta^2}$ can be inefficient for very soft and displaced objects. This is due to a shifting frame of reference. Namely, φ is defined at the track's reference point. If this point is distinctly different from the origin vertex of the generator-level particle, also the azimuthal angle φ of the track and the generator-level particle are different. This can also be due to subtle misaligment or material interactions of the charged particle in the detector. The issue is illustrated in Fig. 3.1 for the signal process.

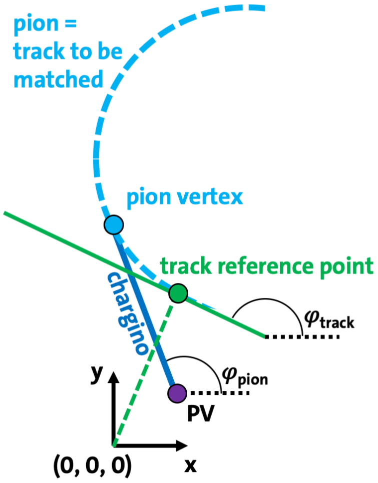

Figure 3.1: Sketch visualizing how the standard ΔR matching breaks down for very soft and displaced tracks due to different definitions of φ.

To recover matching efficiency for soft and displaced particles, a helical matching procedure is developed. It uses the vertex of the generated particle in addition to its trajectory. By extrapolating the helix of a given track (Eq. 3.1), the parameter value t_{min} can be determined that corresponds to the distance of closest approach Δxyz of the track helix to said vertex. Then, by extrapolating the track's azimuthal angle to the point of closest approach, the extrapolated distance measure

$$\Delta R^{\text{extrapolated}} = \sqrt{(\varphi_{\text{GEN}} - \varphi_{\text{track}}(t_{\text{min}}))^2 + (\eta_{\text{GEN}} - \eta_{\text{track}})^2}$$

is computed and the matching criterion is defined to be

$$\Delta xyz < 0.2\,\text{cm} \quad \text{and} \quad \Delta R^{\text{extrapolated}} < 0.05\,.$$

To reduce the computational load without a loss of matching efficiency, the helical matching to a generator-level particle is performed only for the subset of tracks which have been assigned the particle's charge and fulfill $|p_{\mathrm{T,GEN}} - p_{\mathrm{T,track}}|/p_{\mathrm{T,GEN}} < 0.2$, as well as $|\eta_{\mathrm{GEN}} - \eta_{\mathrm{track}}| < 0.1$ and $|\varphi_{\mathrm{GEN}} - \varphi_{0,\mathrm{track}}| < 1.57$. As a fallback in cases where the helical matching leads to no matching track, the simple $\Delta R < 0.02$ matching is attempted on all tracks.

The helical matching procedure is validated by comparing the Δxyz versus $\Delta R^{\mathrm{extrapolated}}$ distributions for signal pions and tracks with the same charge and with opposite charge. The distribution with the wrongly charged track does not show any accumulation in the matching region, indicating that the matching criteria effectively identify the correct tracks and reject random pairings. This can be seen in Fig. 3.2.

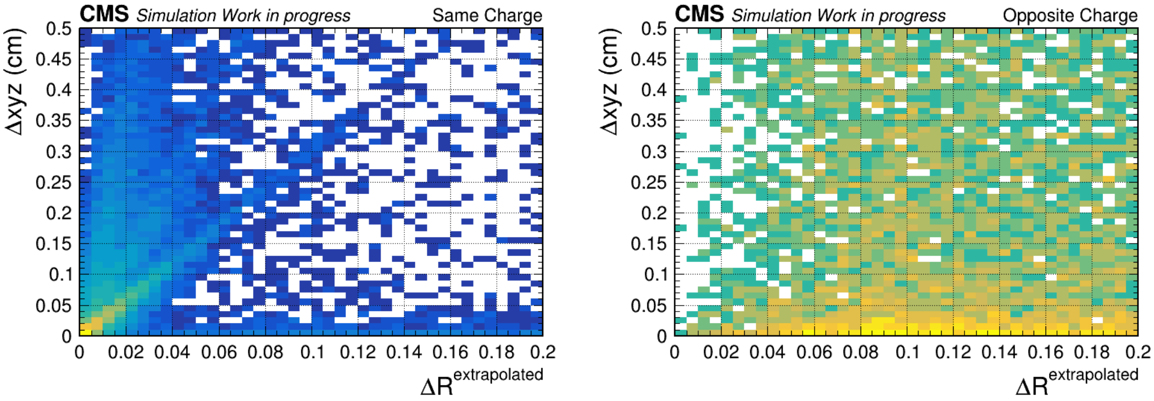

Figure 3.2: Two-dimensional distributions of Δxyz versus $\Delta R^{\mathrm{extrapolated}}$ for generator-level signal pions paired with tracks of the same charge (left) and opposite charge (right). The distributions refer to signal events simulated for 2018 (Phase 1 tracker geometry) using the sum of all signal model points.

The lower rows of Figs. 3.3 and 3.4 show a comparison of the pion matching efficiencies for the standard $\Delta R < 0.02$ matching procedure with the newly developed helical matching for 2016 (Phase 0 tracker) and 2018 (Phase 1 tracker). The matching efficiency is defined as the number of signal pions with a matching reconstructed track (after preselection) divided by all signal pions, and is plotted for different chargino transverse decay lengths and pion transverse momenta. The improvement for transverse decay lengths of the chargino larger than approximately 1 cm and pion transverse momenta lower than roughly 2 GeV are clearly visible. Also plotted in the upper parts are the matching efficiencies, using the procedure outlined above, versus various variables of interest. Complementing Fig. 3.4, Fig. 3.6 shows the efficiency for a wider range of chargino transverse decay lengths.

As can be seen in Fig. 3.5, the efficiency of reconstructing and matching the signal pion is largely independent of the chargino mass $m(\widetilde{\chi}_1^{\pm})$. This is due to fact that the chargino decay length as well as the pion kinematics are rather determined by the mass splitting $\Delta m(\widetilde{\chi}_1^\pm, \widetilde{\chi}_1^0)$.

Track Observables

The helix extrapolation is further used to define a set of custom track observables that serve to improve the final analysis sensitivity. They are listed in Table 3.3.

Table 3.3: Custom track observables defined using the track helix extrapolation described in Section 3.1. All quantities are defined with respect to (a) the leading primary vertex, (b) the closest primary vertex from pileup interactions, (c) the primary vertex associated to the track during reconstruction, and (d) the closest primary vertex excluding the associated vertex.

Observable	Description
IP	Distance in 3D (impact parameter) between a track and a vertex, evaluated at the point of closest approach in 3D between the track helix and the vertex.
IP Significance	Impact parameter significance, IP/σ(IP), with uncertainty σ(IP) calculated by propagating the covariance matrices of the track and the vertex.
IPxy	Distance in the transverse plane (transverse impact parameter) between a track and a vertex, evaluated at the point of closest approach in 3D between the track helix and the vertex.
IPxy Significance	Transverse impact parameter significance, IPxy/σ(IPxy), with uncertainty σ(IPxy) calculated by propagating the covariance matrices of the track and the vertex.
IPz	Distance along the z-axis (longitudinal impact parameter) between a track and the primary vertex, evaluated at the point of closest approach in 3D between the track helix and the primary vertex.
IPz Significance	Transverse impact parameter significance, IPz/σ(IPz), with uncertainty σ(IPz) calculated by propagating the covariance matrices of the track and the vertex.

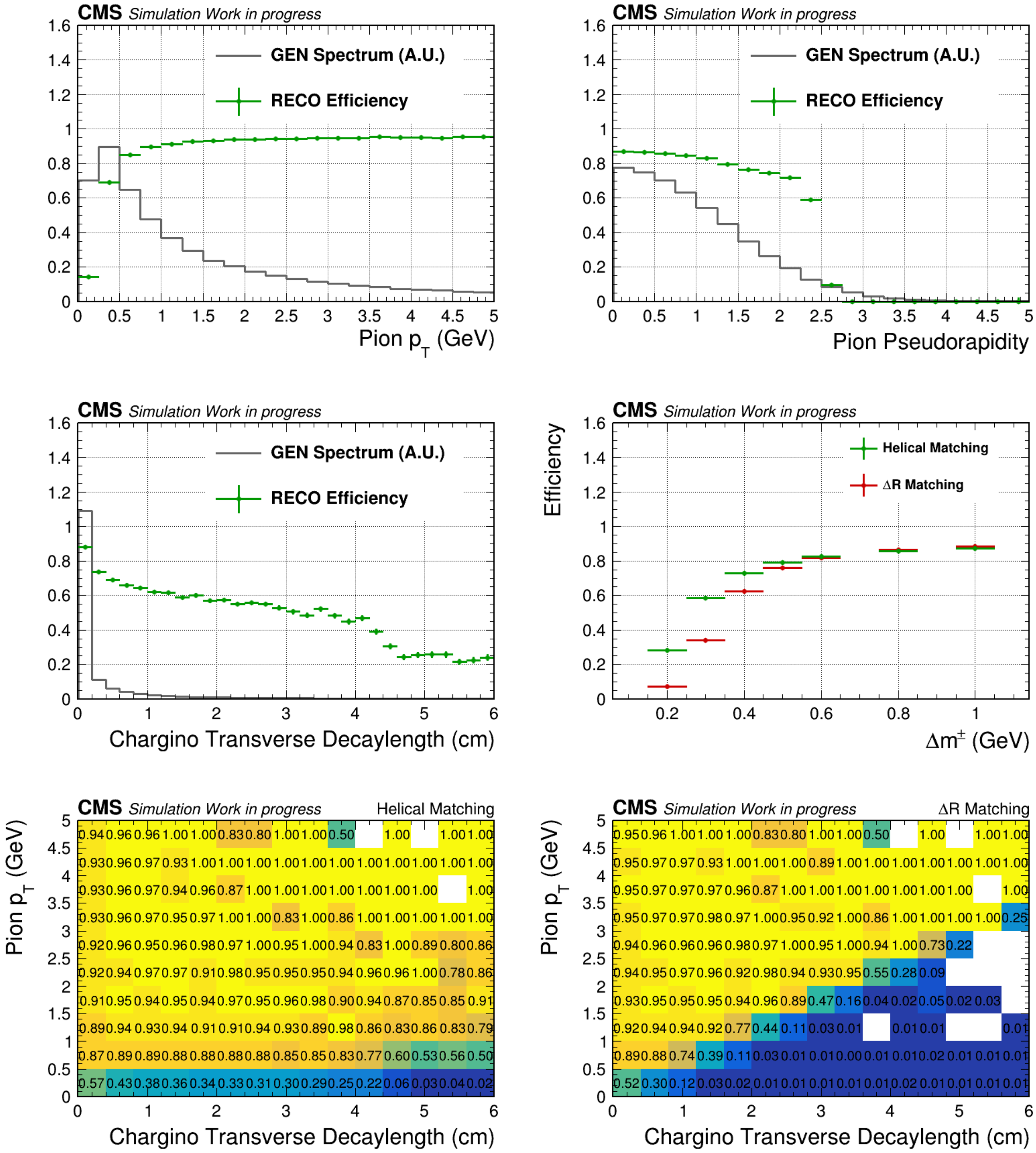

Figure 3.3: Pion reconstruction and matching ("RECO") efficiencies versus pion p_{T} (top left), pion pseudorapidity η (top right), chargino transverse decaylength (center left). Comparisons between the standard ΔR matching and the helical matching efficiencies versus $\Delta m^{\pm}$ (center right) and in two dimensions (pion p_{T} versus chargino transverse decaylength, bottom row). All distributions refer to signal events simulated for 2016 (Phase 0 tracker geometry) using the sum of all signal model points.

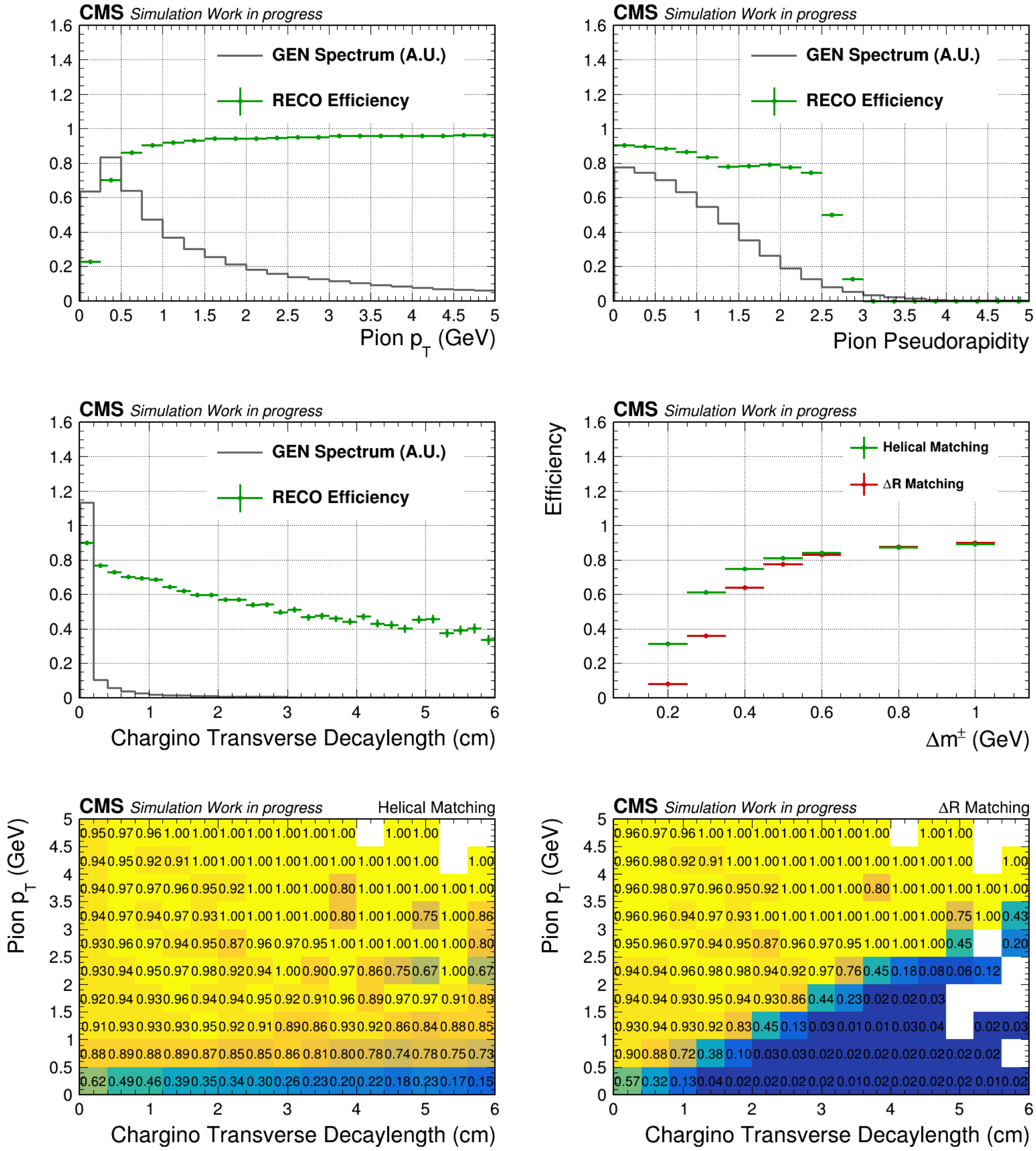

Figure 3.4: Pion reconstruction and matching ("RECO") efficiencies versus pion p_{T} (top left), pion pseudorapidity η (top right), chargino transverse decaylength (center left). Comparisons between the standard ΔR matching and the helical matching efficiencies versus $\Delta m^{\pm}$ (center right) and in two dimensions (pion p_{T} versus chargino transverse decaylength, bottom row). All distributions refer to signal events simulated for 2018 (Phase 1 tracker geometry) using the sum of all signal model points.

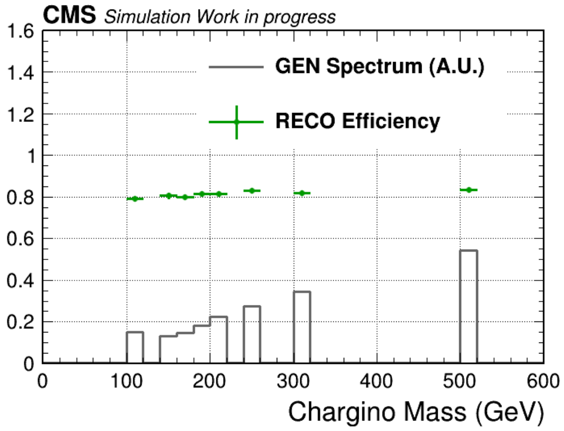

Figure 3.5: Pion reconstruction and matching ("RECO") efficiency versus chargino mass for signal events simulated for 2018 (Phase 1 tracker geometry) using the sum of all signal model points.

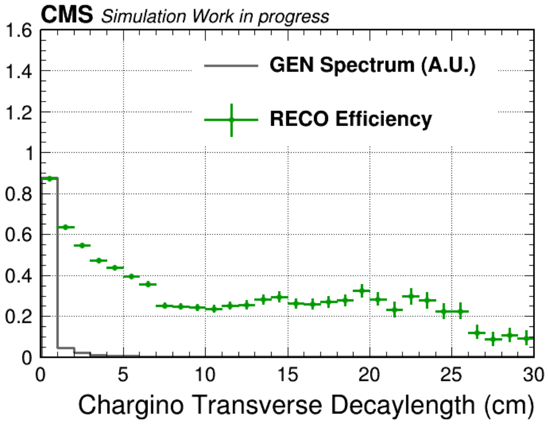

Figure 3.6: Pion reconstruction and matching ("RECO") efficiency versus chargino transverse decaylength for signal events simulated for 2018 (Phase 1 tracker geometry) using the sum of all signal model points. Zoomed out version of the plot in Fig. 3.4.

4 Soft Track Classifier

A multiclass neural network is used to distinguish signal tracks from background tracks, where background tracks refers to all tracks in signal events not originating from the $\widetilde{\chi}_1^{\pm}$ or $\widetilde{\chi}_2^0$ decay, as well as tracks in SM processes. For training and validation of the network, tracks matched to the generator-level pion from the chargino decay (see Section 3.1) are taken from the simulated signal samples and constitute the signal track class. All model points given in Table 2.1 are used and weights are applied to provide a flat distribution of $\Delta m^{\pm}$ and $m(\widetilde{\chi}_1^{\pm})$ for training. For the final results of the analysis, not only the decay to a single pion but all possible decays channels of the chargino are taken into account, see Fig. 1.2.

Background tracks are taken from simulated $\mathrm{Z} \to \nu\nu$ and $\mathrm{W} \to \mathrm{l}\nu$ events, weighted relative to cross section. A matching of background tracks to generator-level particles is also established (if possible), and they are subsequently classified into four classes of background tracks: spurious tracks (without a generator-level match), tracks associated to prompt particles from the primary vertex, tracks associated to secondary particles from the primary vertex, and tracks associated to decay products of τ decays in $\mathrm{W} \to \tau\nu$ events. The data set is divided into a train and a test data set; Table 4.1 lists all track classes with the corresponding numbers of tracks used.

As described in Section 4.1, the characteristics of the signal tracks vary considerably for different values of the signal model parameter $\Delta m^{\pm}$. Therefore, the neural network is parameterized by this mass splitting. The value $\Delta m^{\pm}$ is provided to the network as an additional input along with the observable features. For background tracks, a random value is chosen by sampling the distribution of $\Delta m^{\pm}$ in the signal sample. To account for different characteristics of tracks from different background track classes, the network has multiple output nodes; one for signal tracks and one for each of the four classes of background tracks.

The network uses the 37 variables listed in Table 4.2 and is parameterized by $\Delta m^{\pm}$ and an integer value encoding the era of data taking. Distributions of the input variables for signal and background tracks can be found in Figs. A.2 and A.3 in the Appendix.

Table 4.1: Statistics of the train and test data sets for the different track classes. An independent validation data set is constructed by taking 10% of the tracks of the training data set.

Track class	$N^{\text{train}}_{\text{unweighted}}$	$N^{\text{train}}_{\text{weighted}}$	$N^{\text{test}}_{\text{unweighted}}$
Signal	6×10^5	6.00×10^5	8×10^4
Spurious / no GEN match	1×10^6	3.89×10^5	1×10^5
PV-associated (prompt)	1×10^6	0.90×10^5	1×10^5
PV-associated (secondary)	1×10^6	1.14×10^5	1×10^5
Secondary from τ in $\mathrm{W} \to \tau\nu$	4×10^4	0.06×10^5	7×10^3

4.1 Signal Characteristics

Figures 4.1 and 4.2 show some of the important track variables for different signal model points compared to background tracks. Both background tracks in signal events (dashed histogram) and tracks from SM background events (stacked histogram) are shown. Differences between the two can arise due to different event topologies or might be a sign of biases in the simulation of signal events (using FastSim) compared to FullSim.

In Fig. 4.1 (upper left), it can be seen that the p_{T} of the signal track, as noted above, is on the order of the mass splitting $\Delta m^{\pm}$. Independent of the mass splitting, signal tracks are isolated from any jets in the event (upper right) and point in the direction of $\vec{p}_{\mathrm{T}}^{\text{miss}}$ and opposite to the leading jet (lower two plots).

Another key aspect is the displacement of the signal tracks, as shown by the plots in Fig. 4.2. Most importantly, the upper left plot shows that the smaller $\Delta m^{\pm}$, the longer the chargino lifetime and, therefore, the more displaced the signal track with respect to the leading primary vertex (PV). The displacement is quantified by the impact parameter (IP) significance with respect to the primary vertex both in the transverse and longitudinal direction (see Section 3.1). The distribution for background tracks is bimodal, with one peak at small values attributed to tracks from the PV and another peak at higher values comprising “spurious” tracks including pileup and fakes. When looking at the displacement with respect to the closest pileup vertex, including (upper right) and excluding (lower left) the vertex to which the track is associated to during reconstruction, signal tracks yield large values. In addition, the error on the transverse displacement is smaller for signal tracks compared to background tracks (lower right plot).

Table 4.2: Variables used by the neural network soft track classifier.

Variable	Description
p_T	Transverse momentum of the track.
$\|\eta\|$	Pseudorapidity of the track.
$\|\Delta\varphi(\mathrm{Track}, \vec{p}_\mathrm{T}^\mathrm{\,miss})\|$	Azimuthal angle between the track and the missing transverse momentum vector.
$\|\Delta\varphi(\mathrm{Track}, \mathrm{Leading\ Jet})\|$, $\|\Delta\eta(\mathrm{Track}, \mathrm{Leading\ Jet})\|$	Azimuthal angle and distance in pseudorapidity between the track and the leading jet.
$\log_{10}(\mathrm{dxy})$, $\log_{10}(\mathrm{dz})$	Transverse and longitudinal impact parameters (standard straight line approximation) with respect to (a) the leading primary vertex and (b) the closest primary vertex from pileup interactions.
$\log_{10}(\mathrm{dxy}^\mathrm{Error})$, $\log_{10}(\mathrm{dz}^\mathrm{Error})$	Error on the transverse and longitudinal impact parameters (standard straight line approximation).
$\log_{10}$(IPxy), $\log_{10}$(IPz), $\log_{10}$(IPxy Significance), $\log_{10}$(IPz Significance)	Transverse and longitudinal impact parameters and impact parameter significances (custom helix extrapolation) with respect to (a) the leading primary vertex, (b) the closest primary vertex from pileup interactions, (c) the primary vertex associated to the track during reconstruction, and (d) the closest primary vertex excluding the associated vertex.
Δxy(PV, ass. PV), Δz(PV, ass. PV)	Distance in the transverse plane and along the z-axis between the leading primary vertex and the primary vertex associated to the track (if assigned).
Δxy(PV, ass. SV), Δz(PV, ass. SV)	Distance in the transverse plane and along the z-axis between the leading primary vertex and the secondary vertex associated to the track (if assigned).
Abs. Iso PF	Sum of transverse momenta of PF candidates within a cone of $\Delta R < 0.3$ around the track.
ΔR_min	Distances to the (a) closest jet with $p_\mathrm{T} > 30\,\mathrm{GeV}$, (b) closest jet with $p_\mathrm{T} > 15\,\mathrm{GeV}$, (c) closest track with $p_\mathrm{T} > 5\,\mathrm{GeV}$, and (d) second closest track with $p_\mathrm{T} > 5\,\mathrm{GeV}$.
$p_\mathrm{T}^\mathrm{miss}$	Event-level magnitude of missing transverse momentum.

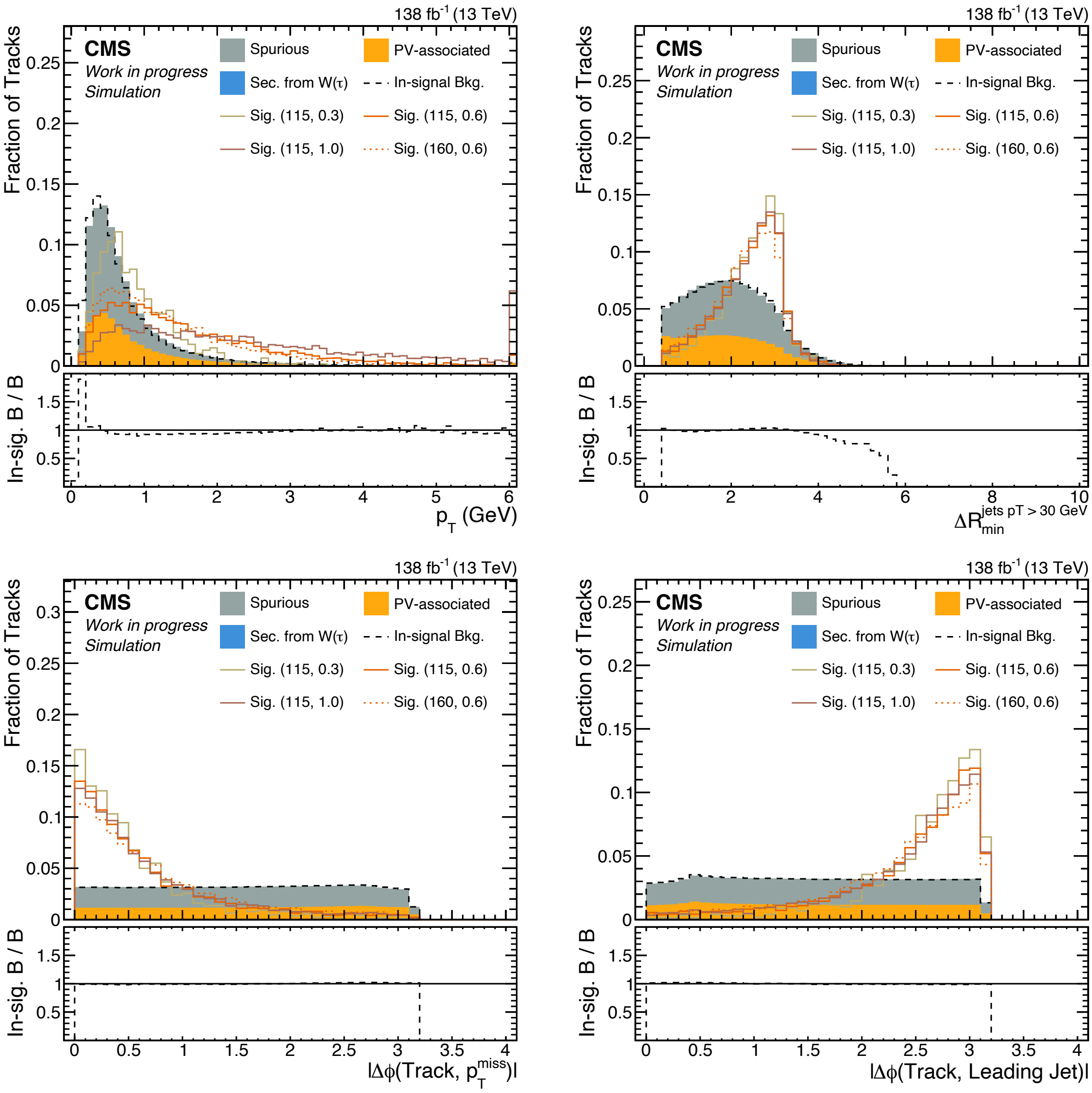

Figure 4.1: Distributions of track observables for background and signal tracks from four example signal model points. The filled histograms show distributions for background tracks in SM background events, whereas the dashed lines show background tracks within signal events, i.e., tracks that are not matched to (decay products of) the SUSY particles. The lower panels show the ratio of in-signal background to SM background. The contribution from tau decay products is too small to be visible.

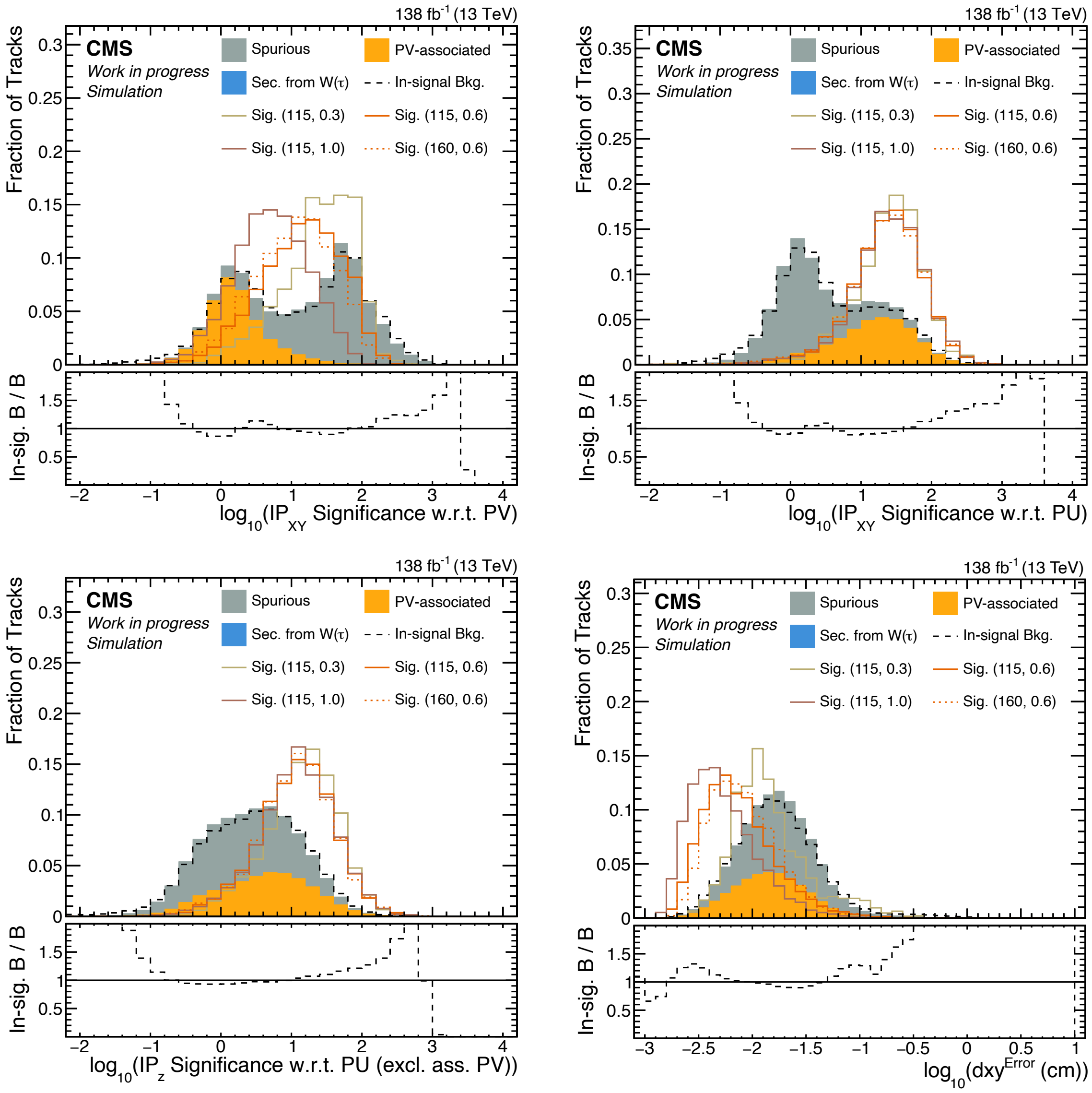

Figure 4.2: Distributions of displacement-related track observables for background and signal tracks from four example signal model points. The filled histograms show distributions for background tracks in SM background events, whereas the dashed lines show background tracks within signal events, i.e., tracks that are not matched to (decay products of) the SUSY particles. The lower panels show the ratio of in-signal background to SM background. The contribution from tau decay products is too small to be visible.

4.2 Architecture and Training

Table 4.3 summarizes the hyperparameters of the soft track classifier neural network and Table 4.4 comprises a list of parameters that determine the network training. Preprocessing layers are added to the network to one-hot-encode the era parameter and to normalize the inputs by subtracting the mean and dividing by the variance of each variable. The classification performance was checked to be robust with respect to the detailed choice of hyperparameters.

Table 4.3: The hyperparameters defining the architecture of the neural network used for multiclass track classification. The number of output dimensions corresponds to the number of track classes.

Hyperparameter	Value
Input dimensions	39
Output dimensions	5
Hidden dimensions	256
Number of hidden layers	3
Activation function (hidden layers)	LeakyReLU ($\alpha = 0.1$)
Activation function (last layer)	Softmax
Batch normalization momentum	0.7
Batch normalization epsilon	1×10^{-3}

Table 4.4: Parameter values used for training the fully-connected multiclass neural network.

Training parameter	Value
Loss function	Categorical cross-entropy
Learning rate	1×10^{-4}
Optimizer	Adam
Number of epochs	2000
Early stopping	100
Batch size	10000
Dropout rate	0.2

4.3 Performance

The track-level output of the signal node evaluated for a given value of $\Delta m^{\pm}$ is denoted as P(Signal | $\Delta m = \ldots$ GeV). Figure 4.4 shows the normalized distributions of P(Signal |

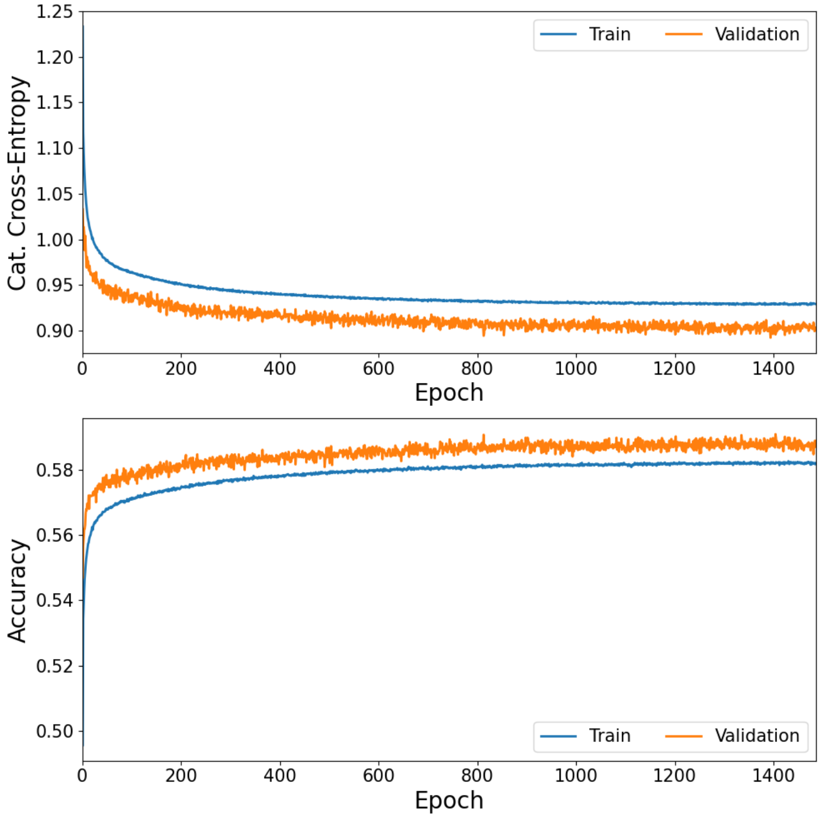

Figure 4.3: Evolution of the categorical cross-entropy loss function (top) and the network accuracy (bottom) over the course of training. Values are given for the train and validation data sets. The slight offset between the two can be explained by the fact that dropout is active for evaluating on the train data set but not for evaluating on the validation data set.

$\Delta m = 0.3\,\mathrm{GeV}$), $\mathrm{P}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$, and $\mathrm{P}(\mathrm{Signal}\,|\,\Delta m = 1.0\,\mathrm{GeV})$ for signal and background tracks from the train and test data sets.

In this analysis, usually, the logit-transformed value is used to enhance the dynamic range of the tails:

$$\widetilde{\mathrm{P}} := \mathrm{logit}\,(\mathrm{P}) = \ln\left(\frac{\mathrm{P}+\epsilon}{1-\mathrm{P}+\epsilon}\right). \tag{4.1}$$

The offset ϵ, taken to be 1×10^{-6}, is added to ensure numerical stability; it confines the possible values of $\widetilde{\mathrm{P}}$ to the range $\left[-\ln(10^6+1), \ln(10^6+1)\right] \approx [-13.82, 13.82]$.

An event-level quantity is defined as the maximum achieved $\widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = \ldots\,\mathrm{GeV})$ within all tracks of a given event: $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = \ldots\,\mathrm{GeV})$ (Fig. 4.5). Note that the event-level $p_{\mathrm{T}}^{\mathrm{miss}}$ is already used as an input to the network.

4.4 Validation

The ROC[1] curves (Fig. 4.6) indicate that the classifier performs well when varying the variables it is parameterized on; $\Delta m^{\pm}$ and era. To substantiate the choice of training one network for all eras combined, the ROC curves for a reference training which only uses the subset of the full training data set that corresponds to 2016 postVFP data are shown. It can be seen that the performance is slightly worse compared to the training with all eras combined.

Also evaluated are the SHAP values [71] for the input variables indicating their importance in classifying a track as signal-like. This is done both for evaluating the network with $\Delta m^{\pm} = 0.3\,\mathrm{GeV}$ (Fig. 4.7) and with $\Delta m^{\pm} = 1.0\,\mathrm{GeV}$ (Fig. 4.8) by selecting 100 tracks from each of the background classes and evaluating the SHAP values with respect to the signal node output. Positive (negative) SHAP values indicate large influence of a variable to classify a given track as (not) signal-like. In the plots, the coloring of each dot depends on the value of the respective variable for the given track ranging from red (large value with respect to the sample mean) to blue (small value with respect to the sample mean). For each background class, the variables are sorted by descending mean absolute SHAP value.

The SHAP values provide in-depth information about the impact of each input variable on the signal node score. Generally, for both the 0.3 and 1 GeV model evaluation, the most important variables for determining signal-like phase space are displacement-related. For the 1 GeV case, smaller displacements and smaller impact parameter significances with respect to the leading primary vertex are associated with more signal-like tracks. For the 0.3 GeV evaluation, on the other hand, larger values of these observables,

---

[1] Receiver operating characteristic

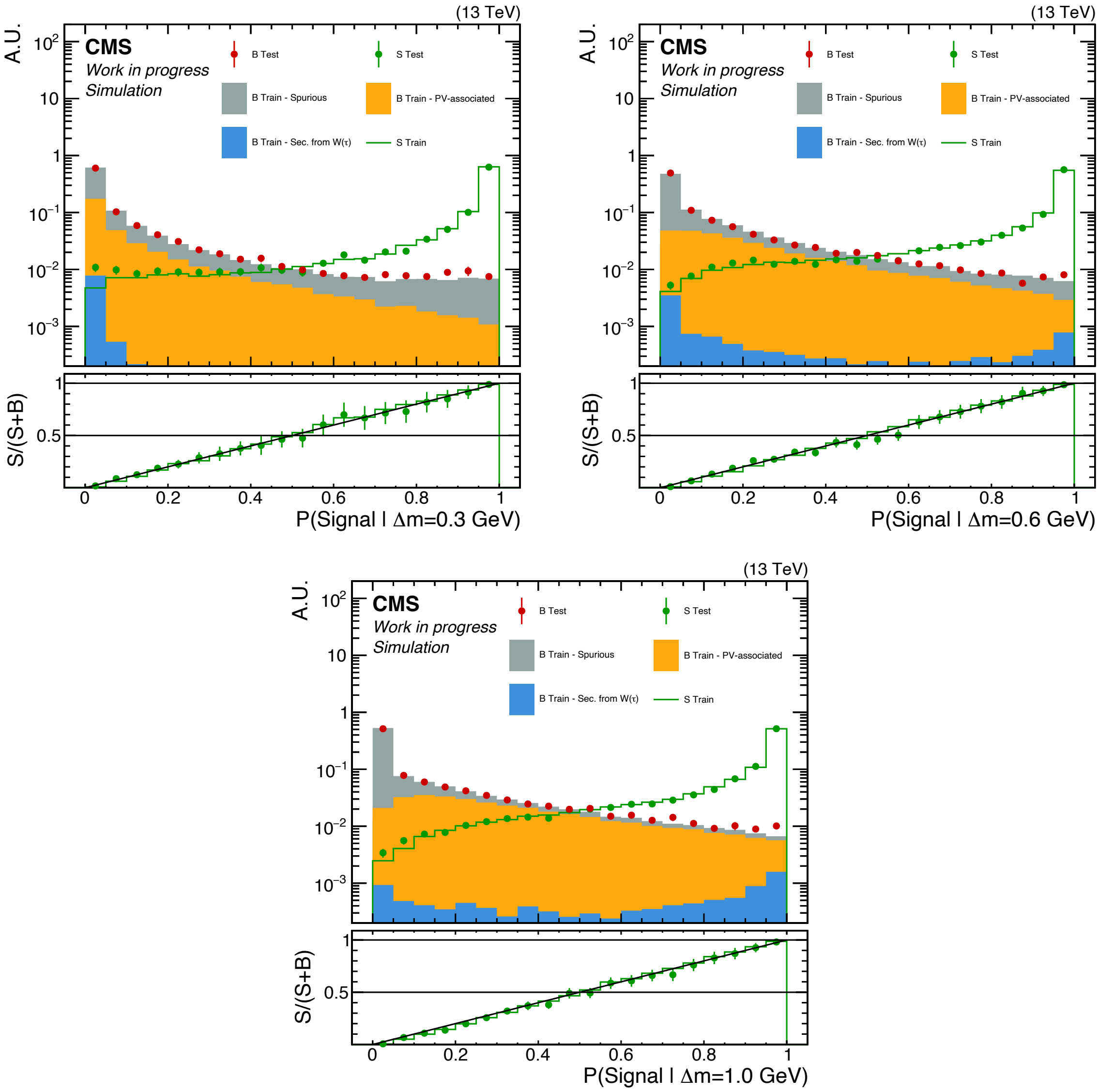

Figure 4.4: Response of the signal node for background (stacked histograms) and signal (green lines) tracks for different $\Delta m^{\pm}$ values used in the evaluation. The set of signal tracks is restricted to the respective mass splitting. Also shown are the distributions for the independent test data sets and, in the lower panels, calibration curves indicating both no signs of overtraining and good calibration.

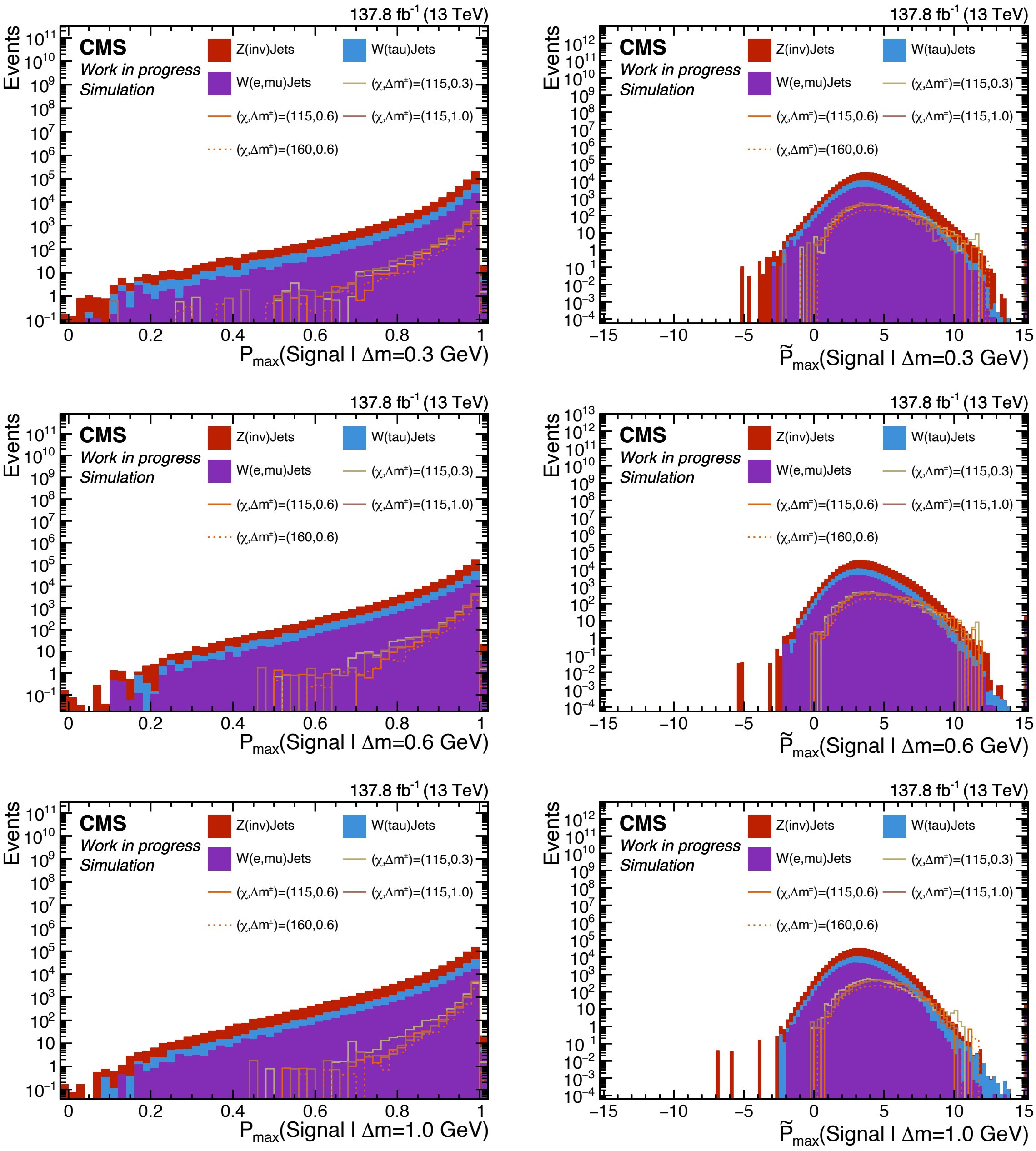

Figure 4.5: Distributions of $\mathrm{P}_{\mathrm{max}}$(Signal) (left column) and $\widetilde{\mathrm{P}}_{\mathrm{max}}$(Signal) (right column) for $\Delta m^{\pm} = 0.3\,\mathrm{GeV}$ (top row), $\Delta m^{\pm} = 0.6\,\mathrm{GeV}$ (center row), and $\Delta m^{\pm} = 1.0\,\mathrm{GeV}$ (bottom row) for signal and background events.

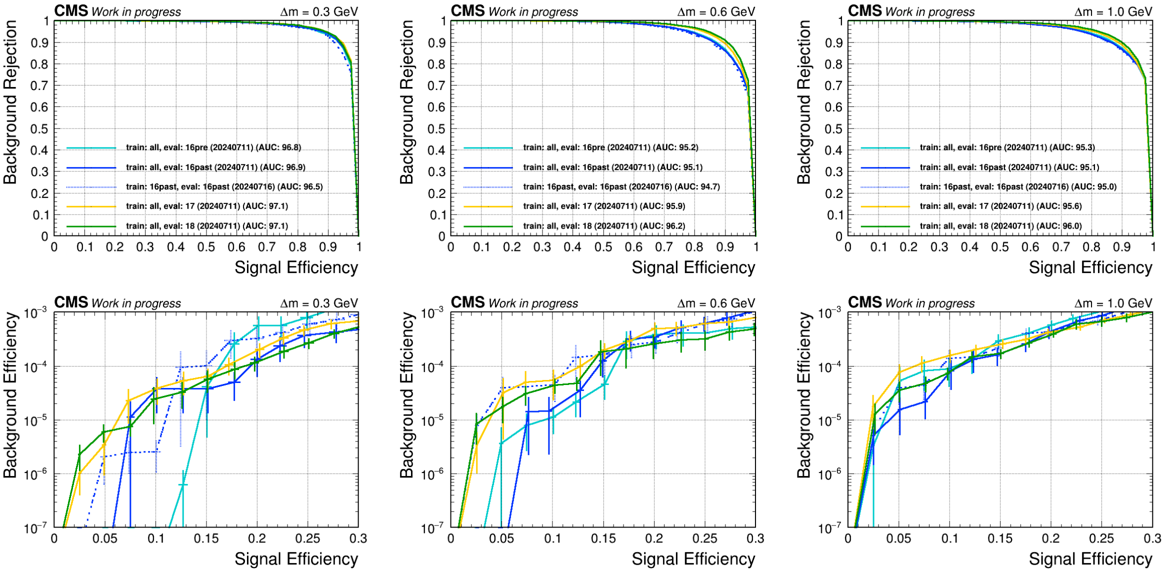

Figure 4.6: Background rejection versus signal efficiency (top row) of the trained neural network classifier for $\Delta m^{\pm} = 0.3\,\text{GeV}$ (left column), $\Delta m^{\pm} = 0.6\,\text{GeV}$ (center column), and $\Delta m^{\pm} = 1.0\,\text{GeV}$ (right column). The bottom row highlights the region of high background rejection (low background efficiency). The curves are plotted for each era separately and a reference training using only 2016 postVFP data is also shown.

Figure 4.7: SHAP values of the NN input variables corresponding to the signal node output evaluated for 100 tracks from each of the Spurious (top left), Prompt (top right), Secondary (bottom left), and Secondary from τ in W $\rightarrow \tau\nu$ (bottom right) background classes. The mass splitting for evaluation is set to $\Delta m^{\pm} = 0.3\,\mathrm{GeV}$.

Figure 4.8: SHAP values of the NN input variables corresponding to the signal node output evaluated for 100 tracks from each of the Spurious (top left), Prompt (top right), Secondary (bottom left), and Secondary from τ in W $\rightarrow \tau\nu$ (bottom right) background classes. The mass splitting for evaluation is set to $\Delta m^{\pm} = 1.0\,\mathrm{GeV}$.

e.g., the transverse impact parameter, lead to more signal-like candidates. Track p_{T} is also important for the 0.3 GeV evaluation with smaller values generally identified as more signal-like. The opposite is true for the 1.0 GeV evaluation where relatively larger values of p_{T} yield higher signal node output scores, except for tracks from the Secondary from τ in $\mathrm{W} \to \tau\nu$ class where smaller values of p_{T} (relative to other tracks within the class) are more signal-like.

High purity of true signal tracks is important when selecting from among all tracks in a signal event. Fig. 4.9 shows that the fraction of signal events in which the highest scoring track is matched to the signal pion (or to another particle from the $\widetilde{\chi}_1^{\pm}$ or $\widetilde{\chi}_2^0$ decay) is greater than 90% for a sufficiently high cut on the max score.

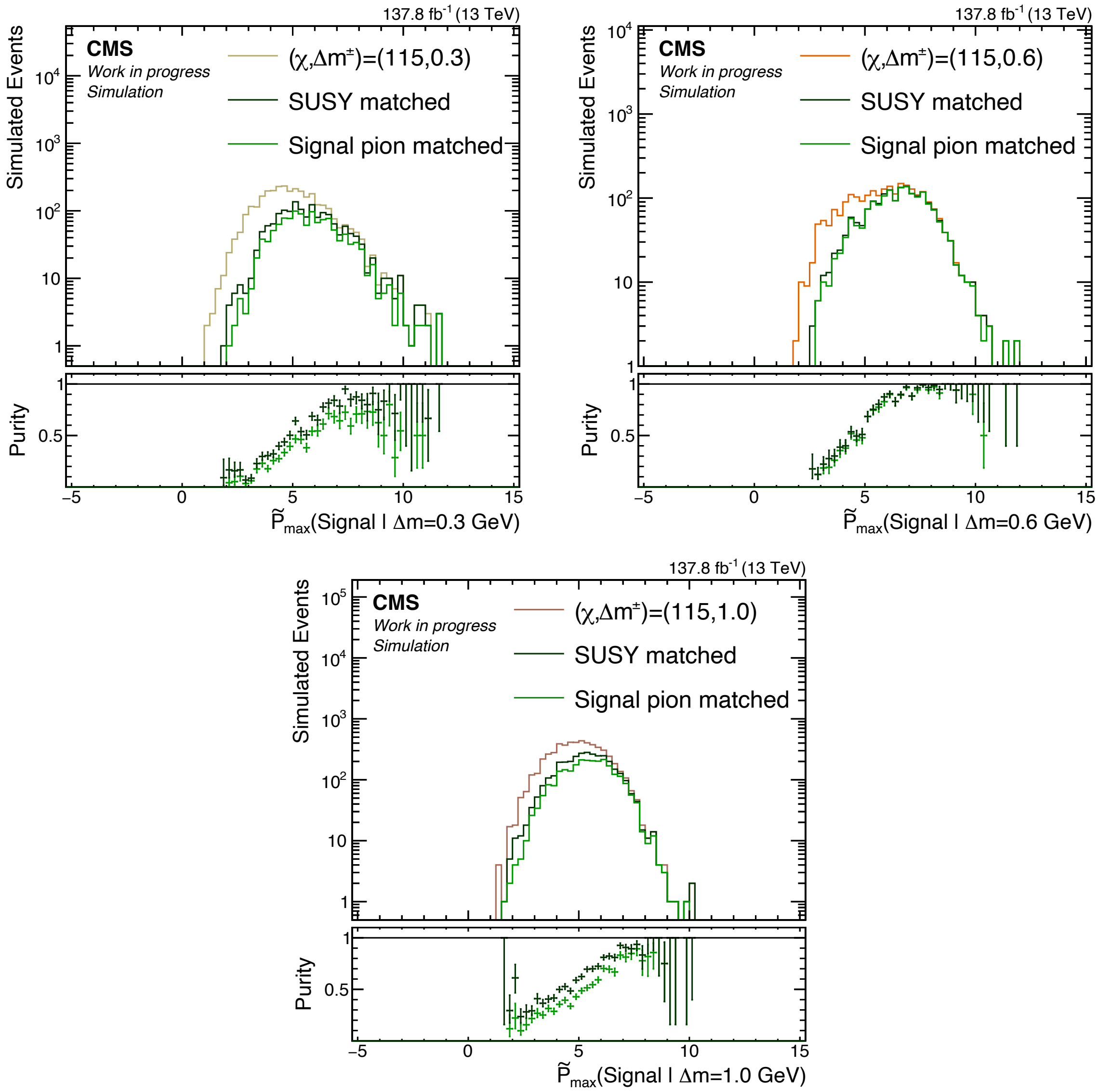

Figure 4.9: Distributions of simulated signal events along with the subsets of events in which the highest scoring track is matched to the signal pion (light green) or any particle associated to the decay of a higgsino (dark green) versus $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal} \,|\, \Delta m = 0.3\,\mathrm{GeV})$ (top left), $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal} \,|\, \Delta m = 0.6\,\mathrm{GeV})$ (top right), and $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal} \,|\, \Delta m = 1.0\,\mathrm{GeV})$ (bottom) with the corresponding fractions shown in the lower panels.

5 Event Selection

5.1 Event Filters

As recommended by the CMS JetMET physics object group[1], events are required to pass the following noise filters to exclude badly reconstructed events:

- `goodVertices`
- `globalSuperTightHalo2016Filter`
- `HBHENoiseFilter`
- `HBHENoiseIsoFilter`
- `EcalDeadCellTriggerPrimitiveFilter`
- `BadPFMuonFilter`
- `BadPFMuonDzFilter` (not for 2016)
- `eeBadScFilter`
- `ecalBadCalibFilter` (not for 2016).

Furthermore, to improve data/MC agreement, events with abnormally large energy deposits in the forward region are discarded by requiring:

- $H_\mathrm{T}^5/H_\mathrm{T} < 2\,.$

To mitigate the influence of the outage of two sectors of the hadronic endcap calorimeter in the 2018 data taking period ("HEM 15/16 failure"), events in the 2018 era are vetoed if there is a jet present that fulfills:

- $p_\mathrm{T} > 30\,\mathrm{GeV}, \Delta\varphi(\mathrm{jet}, \vec{H}_\mathrm{T}^\mathrm{miss}) < 0.5, -3.2 < \eta < -1.2, -1.77 < \varphi < -0.67\,.$

---

[1] `https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2`

5.2 Baseline Selection

The cuts that form the baseline event selection are listed in Table 5.1. They follow closely from the selection in the Run 2 CMS "Monojet" analysis [72], requiring the presence of a high-p_{T} jet and, importantly, large missing transverse energy and vetoing on the presence of isolated photons or leptons, taus, and b-tagged jets.

Tables 5.2 and 5.3 list the efficiencies for each of the cuts for various backgrounds and for signal. Figures 5.1 and 5.2 show data versus MC comparisons for various event-level and track-level observables after the baseline event selection is applied. The disagreement in H_{T} and $H_{\mathrm{T}}^{\mathrm{miss}}$ is addressed by a $p_{\mathrm{T}}^{\mathrm{miss}}$-dependent correction, as explained in Section 6.2.3. The discrepancy in the tail of the N_{PV} (number of primary vertices) distribution is found not to have an impact in the signal regions.

Table 5.1: The event-level cuts defining the baseline event selection. For detailed object definitions see Section 3.

Observable	Criterion
$p_{\mathrm{T}}^{\mathrm{miss}}$	$> 300\,\mathrm{GeV}$
$N_{\mathrm{jet}}(p_{\mathrm{T}} > 100\,\mathrm{GeV})$	≥ 1
$H_{\mathrm{T}}^{\mathrm{miss}}$	$> 300\,\mathrm{GeV}$
$\Delta\varphi_{\mathrm{min}}(\vec{p}_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{jet}_{1,2,3,4})$	> 0.5
$N_{\mathrm{jet}}(p_{\mathrm{T}} > 30\,\mathrm{GeV})$	< 5
$N_{\mathrm{jet}}(p_{\mathrm{T}} > 30\,\mathrm{GeV}, \text{b-tagged loose WP})$	$= 0$
$N_{\text{isolated photon}}$	$= 0$
$N_{\text{isolated electron}}$	$= 0$
$N_{\text{isolated muon}}$	$= 0$
$N_{\text{hadronic tau}}$	$= 0$

Table 5.2: Cut flow tables for background processes indicating the efficiency of each cut of the baseline event selection and the cumulative efficiency with respect to the skim selection in parentheses. The skim selection refers to the selection of events out of all simulated events that are considered for analysis, including noise filters and a cut of $p_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$. A detailed reference of the triggers is given in Section 2.1.

Cut	Z $\rightarrow \nu\nu$ $H_{\mathrm{T}} = 200 - 400\,\mathrm{GeV}$	Z $\rightarrow \nu\nu$ $H_{\mathrm{T}} > 2500\,\mathrm{GeV}$	$\mathrm{t\bar{t}}$ (di-leptonic)
Skim selection	12.85% (100.0%)	50.05% (100.0%)	2.203% (100.0%)
$p_{\mathrm{T}}^{\mathrm{miss}} > 300\,\mathrm{GeV}$	13.18% (13.18%)	62.26% (62.26%)	15.17% (15.17%)
$N_{\mathrm{jet}}(p_{\mathrm{T}} > 100\,\mathrm{GeV}) \geq 1$	99.54% (13.12%)	100.0% (62.26%)	98.74% (14.98%)
$H_{\mathrm{T}}^{\mathrm{miss}} > 300\,\mathrm{GeV}$	69.86% (9.17%)	94.18% (58.63%)	75.56% (11.31%)
$\Delta\varphi_{\mathrm{min}}(\vec{p}_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{jet}_{1,2,3,4}) > 0.5$	97.41% (8.93%)	36.27% (21.27%)	47.80% (5.41%)
$N_{\mathrm{jet}}(p_{\mathrm{T}} > 30\,\mathrm{GeV}) < 5$	98.72% (8.81%)	62.00% (13.18%)	58.63% (3.17%)
$N_{\mathrm{jet}}(p_{\mathrm{T}} > 30\,\mathrm{GeV}, \text{b-tag}) = 0$	75.88% (6.69%)	40.35% (5.32%)	13.02% (0.41%)
$N_{\text{isolated photon}} = 0$	98.42% (6.58%)	97.67% (5.20%)	54.10% (0.22%)
$N_{\text{isolated electron/muon}} = 0$	99.76% (6.57%)	99.64% (5.18%)	29.90% (0.07%)
$N_{\mathrm{tau}} = 0$	99.86% (6.56%)	99.94% (5.17%)	89.81% (0.06%)
Trigger fired	99.30% (6.51%)	99.92% (5.17%)	98.76% (0.06%)

Cut	W $\rightarrow$ lν $H_{\mathrm{T}} = 200 - 400\,\mathrm{GeV}$	W $\rightarrow$ lν $H_{\mathrm{T}} > 2500\,\mathrm{GeV}$	WW (inclusive)
Skim selection	3.951% (100.0%)	31.46% (100.0%)	0.395% (100.0%)
$p_{\mathrm{T}}^{\mathrm{miss}} > 300\,\mathrm{GeV}$	8.910% (8.91%)	52.92% (52.92%)	24.67% (24.67%)
$N_{\mathrm{jet}}(p_{\mathrm{T}} > 100\,\mathrm{GeV}) \geq 1$	99.47% (8.86%)	100.0% (52.92%)	97.63% (24.09%)
$H_{\mathrm{T}}^{\mathrm{miss}} > 300\,\mathrm{GeV}$	73.32% (6.50%)	94.48% (50.00%)	86.20% (20.76%)
$\Delta\varphi_{\mathrm{min}}(\vec{p}_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{jet}_{1,2,3,4}) > 0.5$	91.52% (5.95%)	28.03% (14.01%)	81.96% (17.02%)
$N_{\mathrm{jet}}(p_{\mathrm{T}} > 30\,\mathrm{GeV}) < 5$	98.40% (5.85%)	59.09% (8.28%)	92.00% (15.66%)
$N_{\mathrm{jet}}(p_{\mathrm{T}} > 30\,\mathrm{GeV}, \text{b-tag}) = 0$	75.80% (4.44%)	24.49% (2.03%)	61.08% (9.56%)
$N_{\text{isolated photon}} = 0$	80.73% (3.58%)	63.33% (1.28%)	86.31% (8.25%)
$N_{\text{isolated electron/muon}} = 0$	67.52% (2.42%)	40.71% (0.52%)	82.37% (6.80%)
$N_{\mathrm{tau}} = 0$	97.64% (2.36%)	95.10% (0.50%)	98.45% (6.69%)
Trigger fired	99.28% (2.34%)	99.92% (0.50%)	99.50% (6.66%)

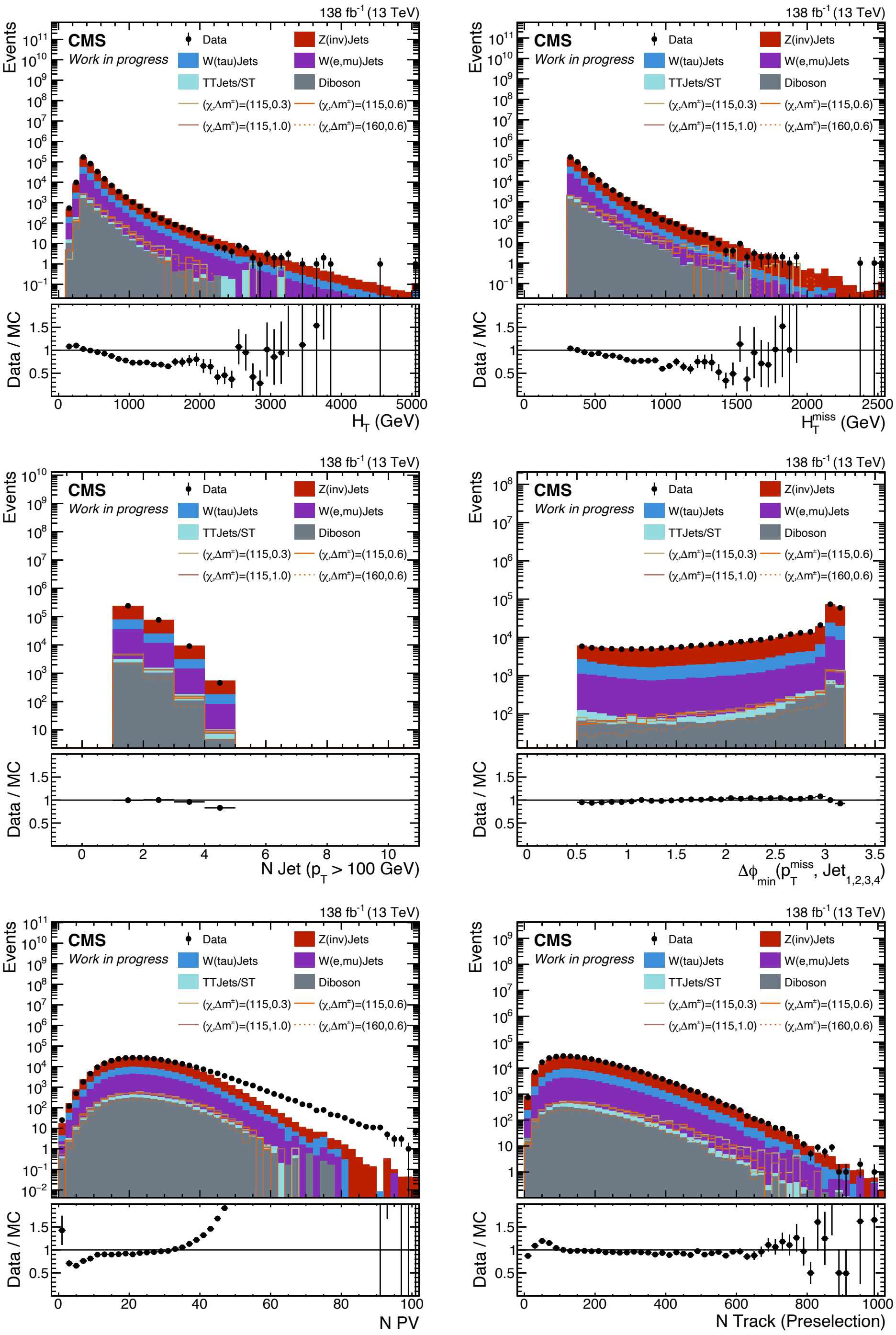

Figure 5.1: Comparison of event-level observables for data and simulated SM background after the baseline event selection is applied. The distributions are shown before the application of the MC corrections introduced in Section 6.2 and without a cut on the soft track classifier output.

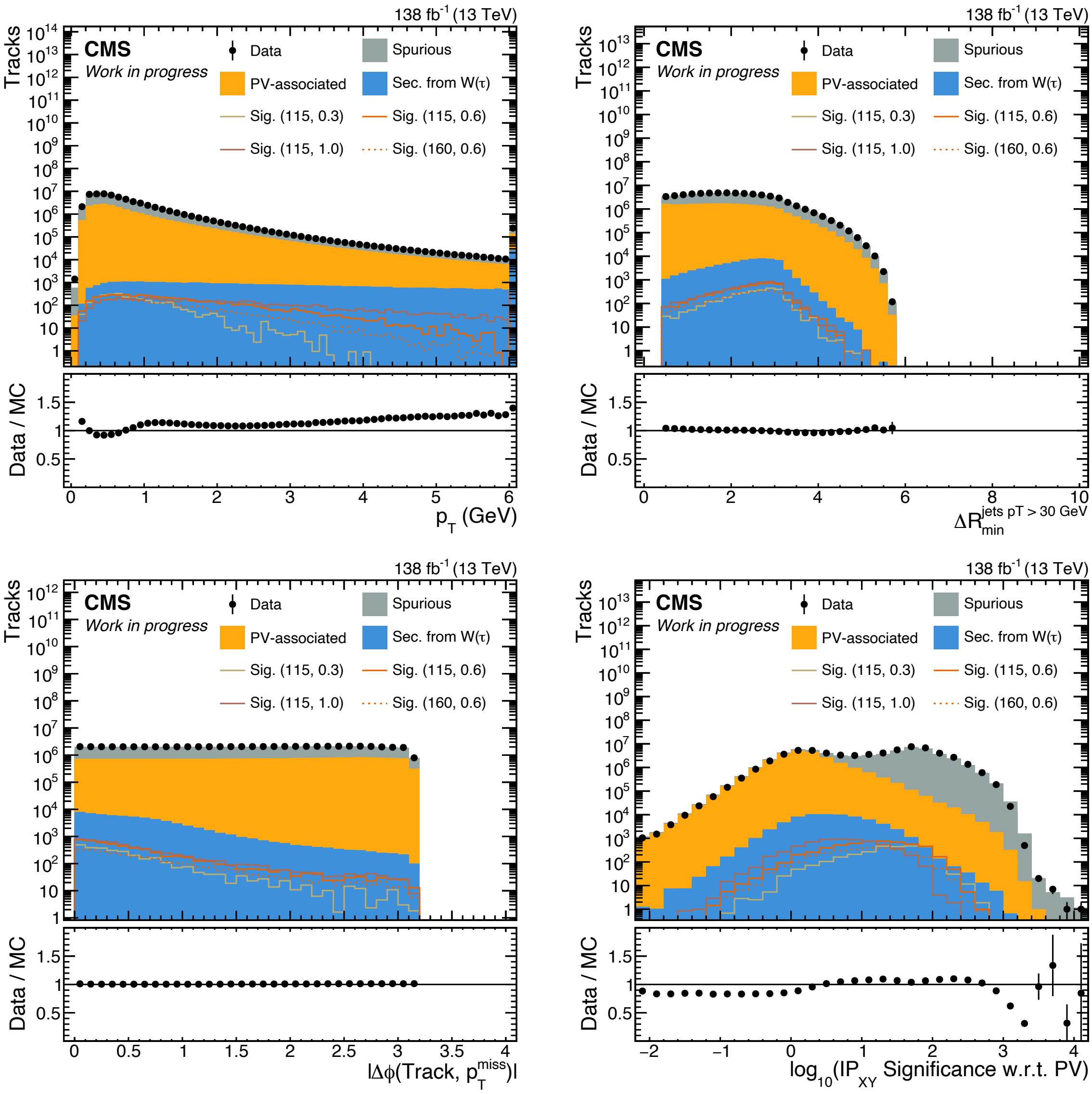

Figure 5.2: Comparison of track-level observables for data and simulated SM background after the baseline event selection is applied. The distributions are shown before the application of the MC corrections introduced in Section 6.2 and without a cut on the soft track classifier output.

Table 5.3: Cut flow tables for specific signal model points indicating the efficiency of each cut of the baseline event selection and the cumulative efficiency with respect to the skim selection in parentheses. The skim selection refers to the selection of events out of all simulated events that are considered for analysis, including noise filters and a cut of $p_\mathrm{T}^\mathrm{miss} > 200\,\mathrm{GeV}$. For signal, the trigger efficiency is not available due to the usage of FastSim.

Cut	$m(\widetilde{\chi}_1^\pm) = 115\,\mathrm{GeV}$, $\Delta m^\pm = 0.3\,\mathrm{GeV}$	$m(\widetilde{\chi}_1^\pm) = 115\,\mathrm{GeV}$, $\Delta m^\pm = 1.0\,\mathrm{GeV}$	$m(\widetilde{\chi}_1^\pm) = 160\,\mathrm{GeV}$, $\Delta m^\pm = 0.6\,\mathrm{GeV}$
Skim selection	2.001% (100.0%)	1.956% (100.0%)	3.011% (100.0%)
$p_\mathrm{T}^\mathrm{miss} > 300\,\mathrm{GeV}$	33.81% (33.81%)	33.50% (33.50%)	37.25% (37.25%)
$N_\mathrm{jet}(p_\mathrm{T} > 100\,\mathrm{GeV}) \geq 1$	99.74% (33.72%)	99.71% (33.41%)	99.79% (37.18%)
$H_\mathrm{T}^\mathrm{miss} > 300\,\mathrm{GeV}$	85.48% (28.82%)	87.26% (29.15%)	86.56% (32.18%)
$\Delta\varphi_\mathrm{min}(\vec{p}_\mathrm{T}^\mathrm{miss}, \mathrm{jet}_{1,2,3,4}) > 0.5$	93.32% (26.90%)	93.40% (27.22%)	93.57% (30.11%)
$N_\mathrm{jet}(p_\mathrm{T} > 30\,\mathrm{GeV}) < 5$	94.63% (25.45%)	94.68% (25.78%)	94.28% (28.39%)
$N_\mathrm{jet}(p_\mathrm{T} > 30\,\mathrm{GeV}, \mathrm{b\text{-}tag}) = 0$	84.19% (21.43%)	84.83% (21.87%)	83.35% (23.66%)
$N_\mathrm{isolated\ photon} = 0$	98.78% (21.17%)	98.50% (21.54%)	98.77% (23.37%)
$N_\mathrm{isolated\ electron/muon} = 0$	99.88% (21.14%)	99.42% (21.41%)	99.84% (23.33%)
$N_\mathrm{tau} = 0$	99.90% (21.12%)	99.87% (21.38%)	99.85% (23.30%)
Trigger fired	—	—	—

5.3 Signal Regions

Phase space regions with enhanced sensitivity to signal events ("signal regions") are constructed and optimized in the following way:

1. Compute the event-level max score observables $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV})$, $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$, and $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 1.0\,\mathrm{GeV})$ (defined in Section 4.3) by evaluating the neural network soft track classifier for all tracks in the event using each of the three mass splittings.
2. Categorize each event into one of three signal distributions (SR 0.3, SR 0.6, SR 1.0) according to the *highest* of the three $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = ...\,\mathrm{GeV})$ values. This ensures orthogonality and avoids double counting of events.
3. In each signal distribution, scan for the highest cut significance in $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = ...\,\mathrm{GeV})$ observable, see Fig. 5.3. The cut significance is computed as the Asimov significance assuming a 20% systematic error on the background yield. For the evaluation of the cut significance, the signal model point with mass splitting corresponding to the mass splitting for which the classifier is evaluated for, is used.
4. In each signal distribution, define three signal regions using the cut obtained in the previous step and two looser cuts which lead to the same number of expected signal events in each signal region ("flat-S binning").
5. To further enhance the sensitivity to the signal, split the tightest signal region in each signal distribution in two, giving four signal regions per classifier evaluation.

This procedure leads to 12 signal regions with different background compositions that are sensitive to different parts of the signal model space (see Tab. 5.4). The efficiencies for each of the mutually exclusive signal regions are given in Tab. 5.5 and 5.6 for a set of background processes and for three signal model points.

Fig. 5.4 shows the efficiency of the signal region cuts for signal events versus different variables. It can be seen how the different signal regions are sensitive to different regions of the signal phase space and that mainly signal events with large $p_{\mathrm{T}}^{\mathrm{miss}} > 500\,\mathrm{GeV}$ contribute to the signal regions. The impact of the high $p_{\mathrm{T}}^{\mathrm{miss}}$ region can also be seen in Fig. 5.5, which shows the significance in SR 0.6 in bins of $p_{\mathrm{T}}^{\mathrm{miss}}$.

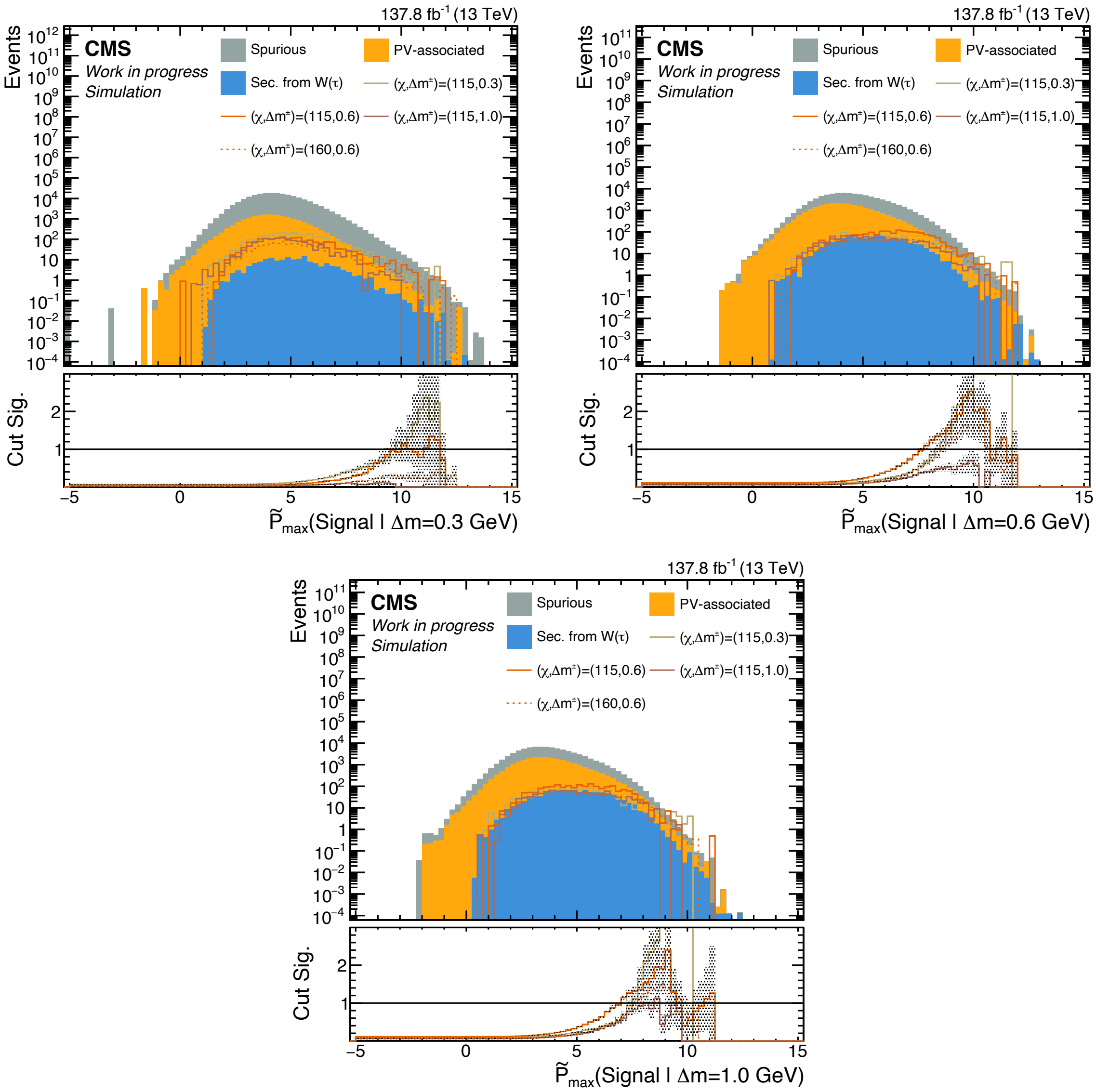

Figure 5.3: Cut significance scans in $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV})$ (top left), $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$ (top right), and $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 1.0\,\mathrm{GeV})$ (bottom) as explained in step 3 of the procedure used in constructing signal regions. The cut significance is computed as the Asimov significance assuming a 20% systematic error on the background yield. In the lower panels, the uncertainties shown around the solid lines indicate the statistical uncertainty of the MC samples.

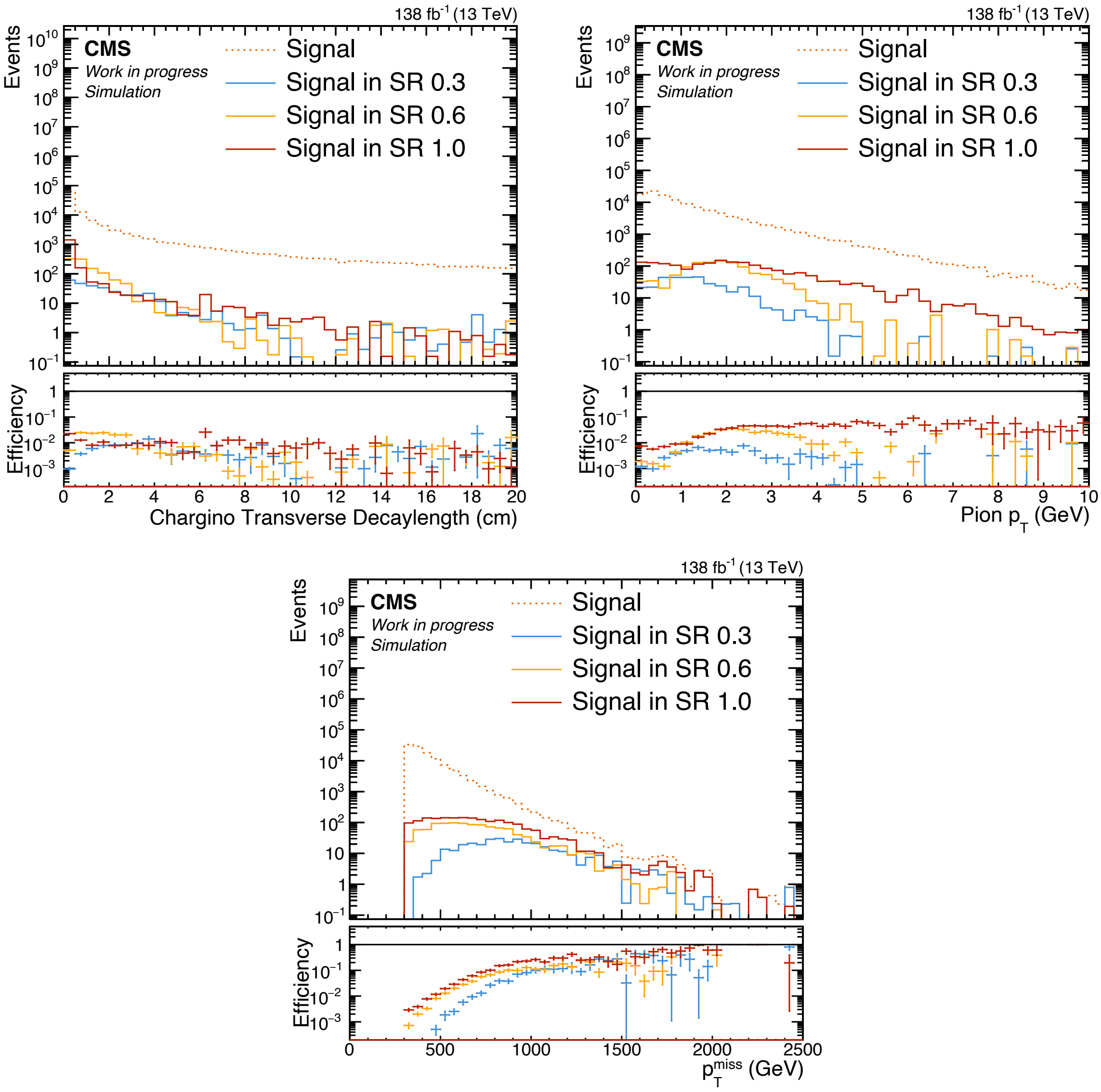

Figure 5.4: Signal efficiency per bin in the three signal regions versus generator-level chargino transverse decaylength (top left), pion p_{T} (top right), and $p_{\mathrm{T}}^{\mathrm{miss}}$ (bottom). The signal events for these plots are pooled from all signal model points with $m(\widetilde{\chi}_1^{\pm}) < 200\,\mathrm{GeV}$ and $\Delta m^{\pm} < 1\,\mathrm{GeV}$.

Table 5.4: Cut values on $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV})$ (for SR 0.3), $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$ (for SR 0.6), and $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 1.0\,\mathrm{GeV})$ (for SR 1.0) defining the signal regions.

Signal region		Lower cut	Upper cut
SR 0.3	I	9.5	10.0
	II	10.0	11.0
	III	11.0	11.5
	IV	11.5	—
SR 0.6	I	8.75	9.0
	II	9.0	9.5
	III	9.5	10.5
	IV	10.5	—
SR 1.0	I	7.25	7.5
	II	7.5	8.0
	III	8.0	8.75
	IV	8.75	—

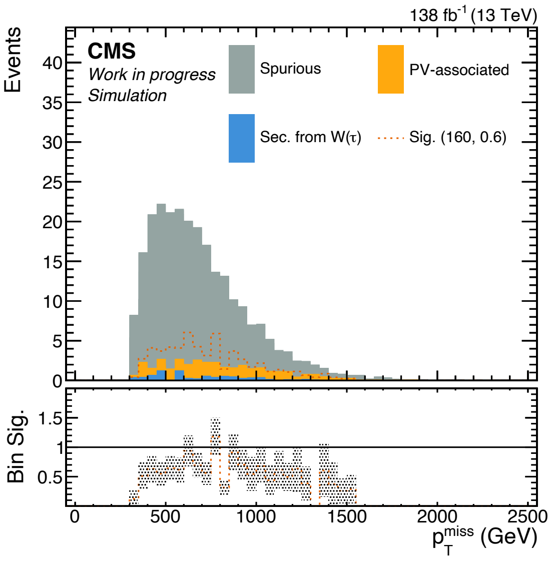

Figure 5.5: Distributions of $p_{\mathrm{T}}^{\mathrm{miss}}$ for signal and background events passing the SR 0.6 selection. The lower panel shows the significance in each bin.

Table 5.5: Efficiencies in the signal regions for various background processes. Efficiencies are stated with respect to the baseline event selection, see Section 5.2.

Signal region		$\mathrm{Z} \to \nu\nu$ $H_\mathrm{T} = 200 - 400\,\mathrm{GeV}$	$\mathrm{Z} \to \nu\nu$ $H_\mathrm{T} > 2500\,\mathrm{GeV}$	$\mathrm{t}\bar{\mathrm{t}}$ (di-leptonic)
SR 0.3	I	0.0000%	0.3632%	0.0000%
	II	0.0000%	0.4269%	0.0000%
	III	0.0000%	0.1274%	0.0000%
	IV	0.0000%	0.1211%	0.0000%
SR 0.6	I	0.0000%	0.0765%	0.0000%
	II	0.0031%	0.1274%	0.0000%
	III	0.0000%	0.1211%	0.0000%
	IV	0.0000%	0.0701%	0.0000%
SR 1.0	I	0.0000%	0.2357%	0.0000%
	II	0.0016%	0.3568%	0.0000%
	III	0.0000%	0.5097%	0.0000%
	IV	0.0000%	0.2740%	0.0000%

Signal region		$\mathrm{W} \to \mathrm{l}\nu$ $H_\mathrm{T} = 200 - 400\,\mathrm{GeV}$	$\mathrm{W} \to \mathrm{l}\nu$ $H_\mathrm{T} > 2500\,\mathrm{GeV}$	WW (inclusive)
SR 0.3	I	0.0000%	0.0805%	0.0228%
	II	0.0000%	0.2213%	0.0228%
	III	0.0000%	0.0604%	0.0000%
	IV	0.0000%	0.0000%	0.0076%
SR 0.6	I	0.0000%	0.0402%	0.0456%
	II	0.0000%	0.0805%	0.0076%
	III	0.0000%	0.0805%	0.0456%
	IV	0.0000%	0.0000%	0.0076%
SR 1.0	I	0.0043%	0.1811%	0.0532%
	II	0.0087%	0.1609%	0.0228%
	III	0.0000%	0.3219%	0.0609%
	IV	0.0000%	0.1207%	0.0000%

Table 5.6: Efficiencies of the signal region cuts for a set of signal model points. The efficiencies are stated with respect to the baseline event selection, see Section 5.2.

Signal region		$m(\widetilde{\chi}_1^\pm) = 115\,\text{GeV}$, $\Delta m^\pm = 0.3\,\text{GeV}$	$m(\widetilde{\chi}_1^\pm) = 115\,\text{GeV}$, $\Delta m^\pm = 1.0\,\text{GeV}$	$m(\widetilde{\chi}_1^\pm) = 160\,\text{GeV}$, $\Delta m^\pm = 0.6\,\text{GeV}$
SR 0.3	I	0.1720%	0.0129%	0.1505%
	II	0.1178%	0.0129%	0.1728%
	III	0.0223%	0.0032%	0.0130%
	IV	0.0223%	0.0097%	0.0409%
SR 0.6	I	0.1306%	0.0453%	0.4236%
	II	0.1879%	0.0615%	0.4551%
	III	0.1306%	0.0388%	0.3920%
	IV	0.0255%	0.0194%	0.1245%
SR 1.0	I	0.1975%	0.6504%	0.5387%
	II	0.2357%	0.6633%	0.6669%
	III	0.1338%	0.4109%	0.4236%
	IV	0.0605%	0.1165%	0.1375%

6 Background Estimation and MC Corrections

The expected SM background event yields in the signal regions of this analysis are estimated using MC simulation corrected by data. The following sections describe the background modeling and the MC corrections used in this analysis. An important component is the usage of a *cleaned* Drell-Yan (DY) control region as described in Section 6.1.

6.1 Cleaned Drell-Yan Control Region

To prepare a well-controlled data sample that mimics the main SM background process $\mathrm{Z} \to \nu\nu$, the Drell-Yan cleaning method is used. The idea is that selecting $\mathrm{Z} \to \mu\mu$ events and "removing" the muons from the events gives a good proxy for $\mathrm{Z} \to \nu\nu$ events since the overall kinematics of the Drell-Yan process are independent of the lepton flavor. Given that the soft component of the events, namely the soft tracks, is largely independent from the hard scatter process, this control region provides a faithful template of both the PV-associated (non-tau) and Spurious tracks across all of the SM background processes.

The cleaned Drell-Yan sample is constructed by first selecting events with two oppositely charged, isolated muons with an invariant mass close to the Z boson mass. The detailed selection criteria are

$$\begin{gathered} p_{\mathrm{T}}(\mu) > 30\,\mathrm{GeV}\,, \\ \mathrm{Iso}_{\mu}/p_{\mathrm{T}}(\mu) < 0.2\,, \\ 75\,\mathrm{GeV} < m(\mathrm{Z}) < 105\,\mathrm{GeV}\,, \\ p_{\mathrm{T}}(\mathrm{Z}) > 200\,\mathrm{GeV}\,, \end{gathered}$$

where Z denotes the dimuon system and the isolation measure Iso_{μ} is the sum of the deposited energy associated with the primary vertex within $\Delta R < 0.3$ around the muon, as discussed in Section 3.

As a second step, the muons are removed from the event record, rendering them neutrino-like. First, the transverse momenta of the selected muons are vectorially added to the missing transverse momentum:

$$\vec{p}_{\mathrm{T}}^{\mathrm{miss}} \rightarrow \vec{p}_{\mathrm{T}}^{\mathrm{miss}} + \sum_{i=1}^{2} \vec{p}_{\mathrm{T}}(\mu_i)\,. \tag{6.1}$$

The muons and associated matched tracks, PF candidates, photons, and jets are removed from the respective collections and are not considered for further analysis of the event. The matching is done using an angular ΔR criterion with thresholds of 0.4 for the jet matching and 0.05 for the other objects.

It is found that the p_{T} distribution of the generator-level Z boson in the cleaned DY MC sample is biased with respect to the distribution found in simulated $\mathrm{Z} \rightarrow \nu\nu$ events. Therefore, the simulated cleaned DY events are reweighted according to the p_{T} of the generator-level Z boson. For the derivation of the weights, the event selection is loosened to the skim selection (see Section 5.2) with $p_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$ and no cut on the p_{T} of the reconstructed dimuon system is applied. The weights are determined individually for each data taking era. The corresponding histograms can be found in Fig. A.1 in the Appendix. As can be seen in Fig. 6.1, this reweighting improves the agreement of the cleaned DY MC $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution with both MC $\mathrm{Z} \rightarrow \nu\nu$ and cleaned DY data.

Figure 6.2 shows the distributions of the invariant mass and transverse momentum of the muon pair for both data and MC. The lower panel shows that there is good agreement of the simulated cleaned DY sample with the data sample. Furthermore, Figs. 6.4 and 6.5 show that the observable shapes found in the cleaned DY sample very well approximate the real $p_{\mathrm{T}}^{\mathrm{miss}}$ background, $\mathrm{Z} \rightarrow \nu\nu$, even in the neural network max scores observables.

Throughout this section, data distributions in the cleaned DY control region are displayed with green dots as opposed to black dots used for the real $p_{\mathrm{T}}^{\mathrm{miss}}$ data.

Track- vs. Event-level Signal Node Output Score Observables

The cleaned DY data set is also used to study the connection between the inclusive track-level signal node scores $\widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = ...\,\mathrm{GeV})$ and the event-level max scores $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = ...\,\mathrm{GeV})$. As shown in Fig. 6.6, the two observables become congruent for values of $\widetilde{\mathrm{P}}_{(\mathrm{max})}(\mathrm{Signal}\,|\,\Delta m = ...\,\mathrm{GeV}) \gtrapprox 5$, signifying that the number of tracks above this threshold per event is (close to) one.

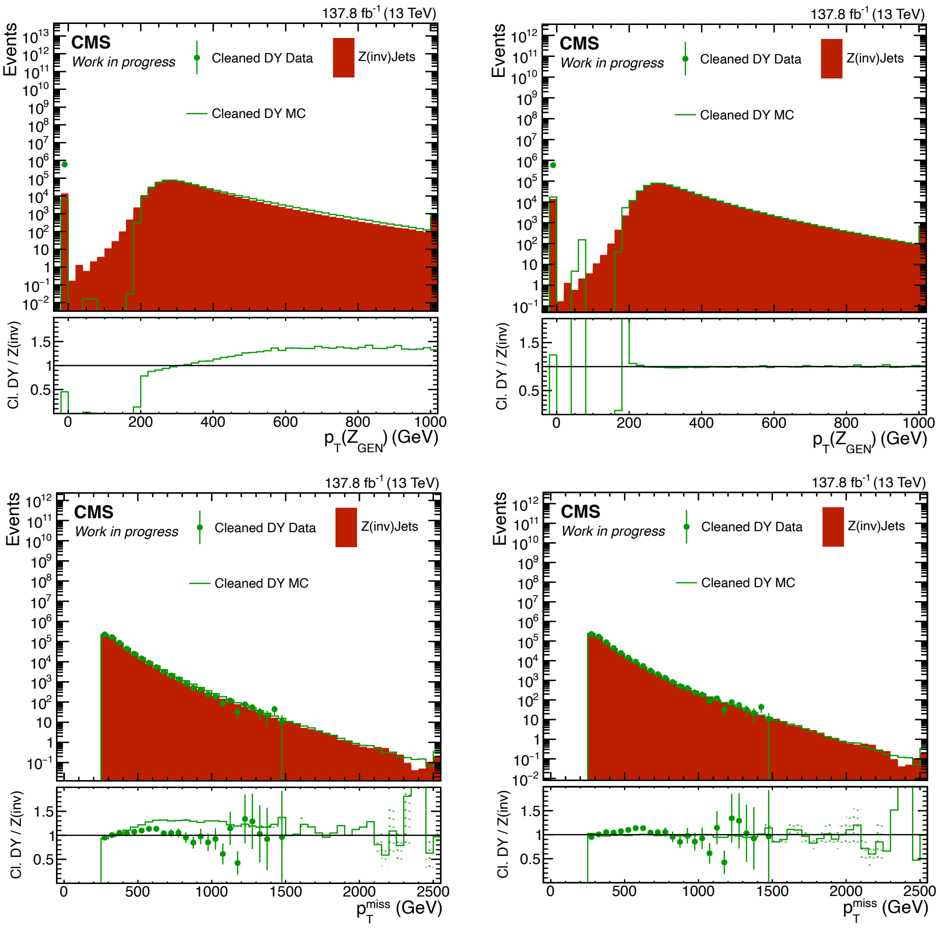

Figure 6.1: Distributions of $p_{\mathrm{T}}(\mathrm{Z}_{\mathrm{GEN}})$ (top row) and $p_{\mathrm{T}}^{\mathrm{miss}}$ (bottom row) for cleaned DY and $\mathrm{Z} \rightarrow \nu\nu$ before (left column) and after (right column) the cleaned DY MC events are reweighted. For these distributions, the $p_{\mathrm{T}}(\mathrm{Z}) > 200\,\mathrm{GeV}$ cut is not applied.

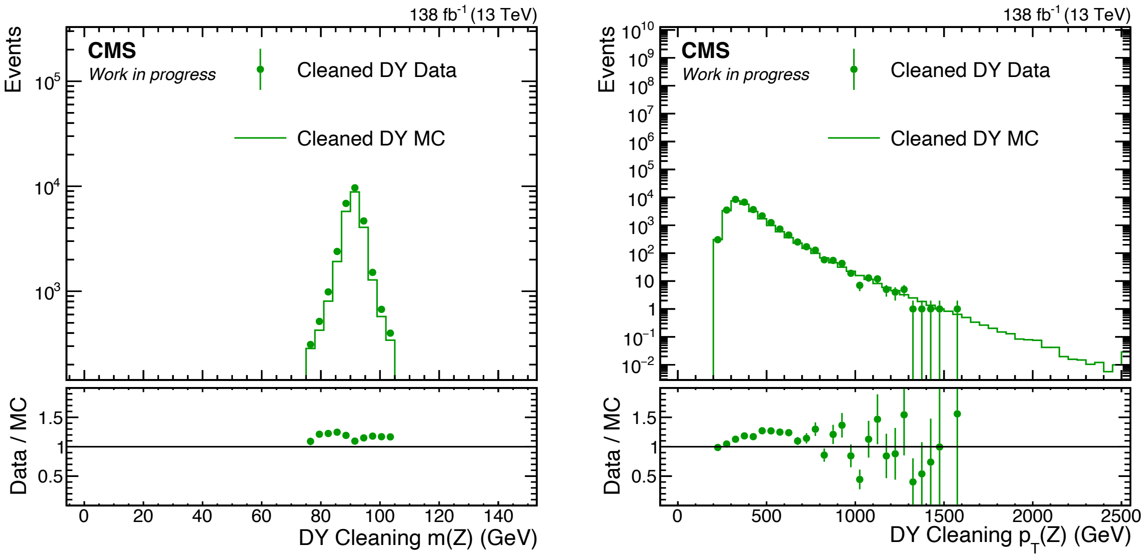

Figure 6.2: Invariant mass and p_T of the dimuon system used in the Drell-Yan cleaning.

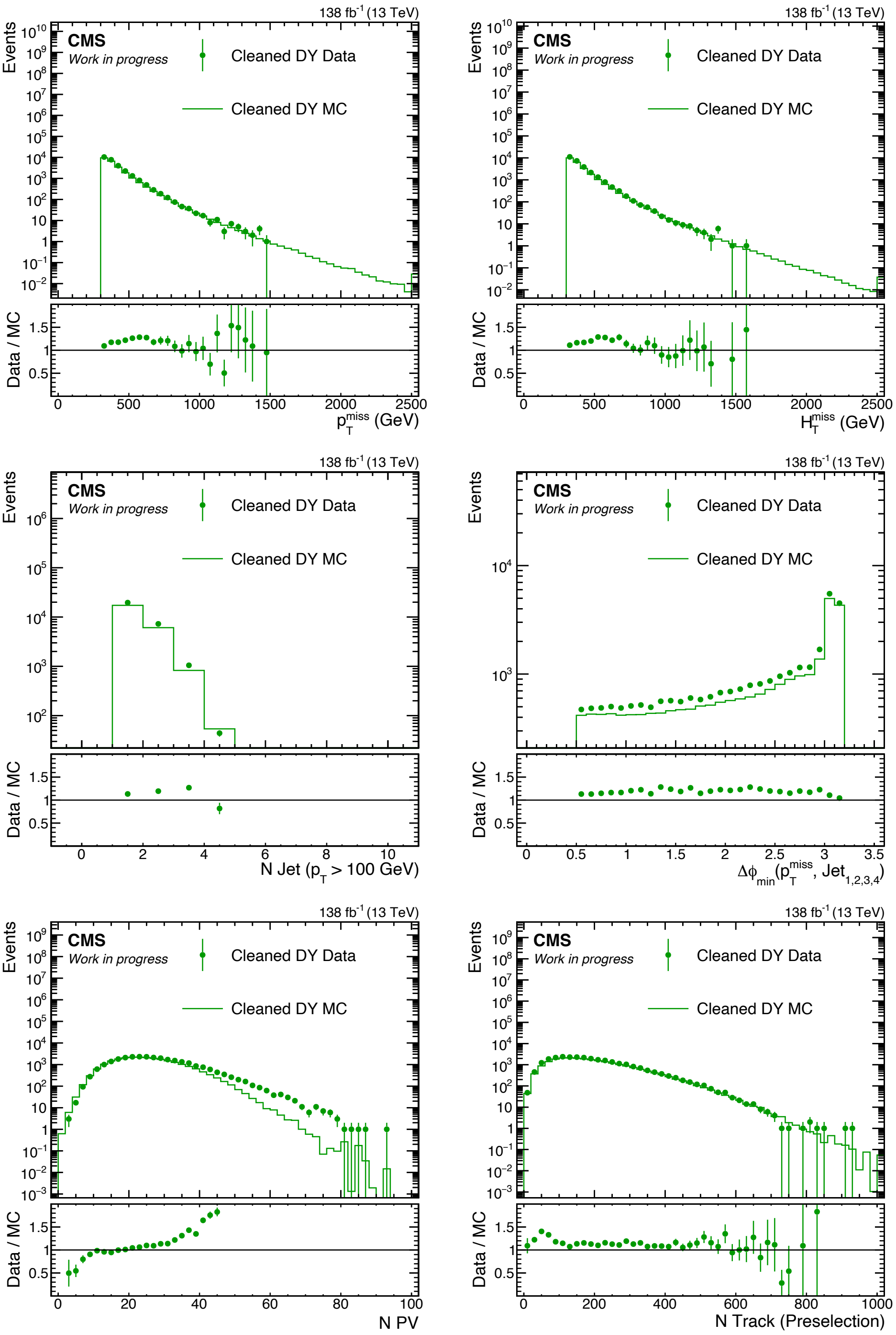

Figure 6.3: Comparison of event-level observables for the cleaned DY sample in data and MC.

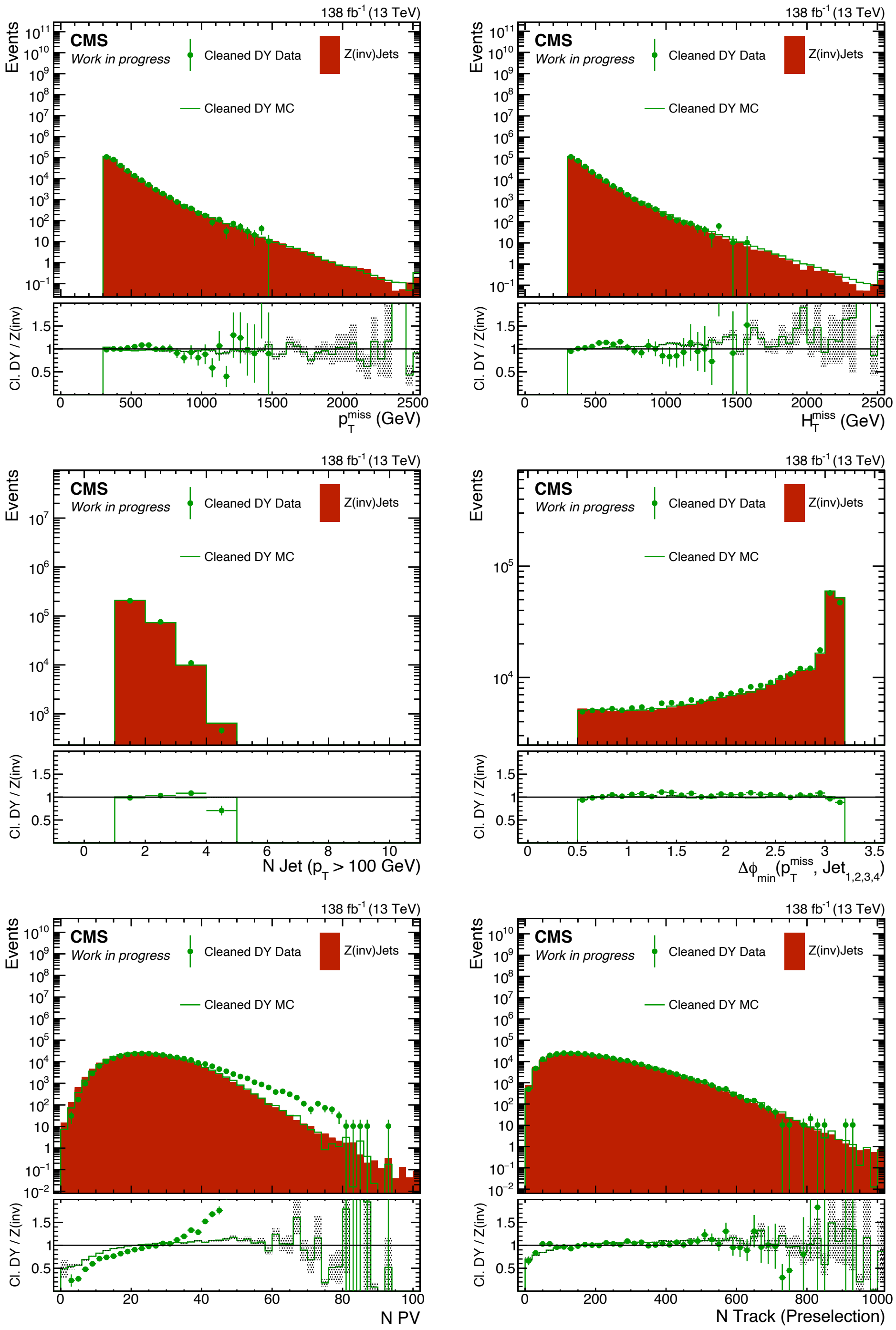

Figure 6.4: Distributions of event-level observables for Z → $\nu\nu$ events and cleaned DY (data and MC) scaled to Z → $\nu\nu$.

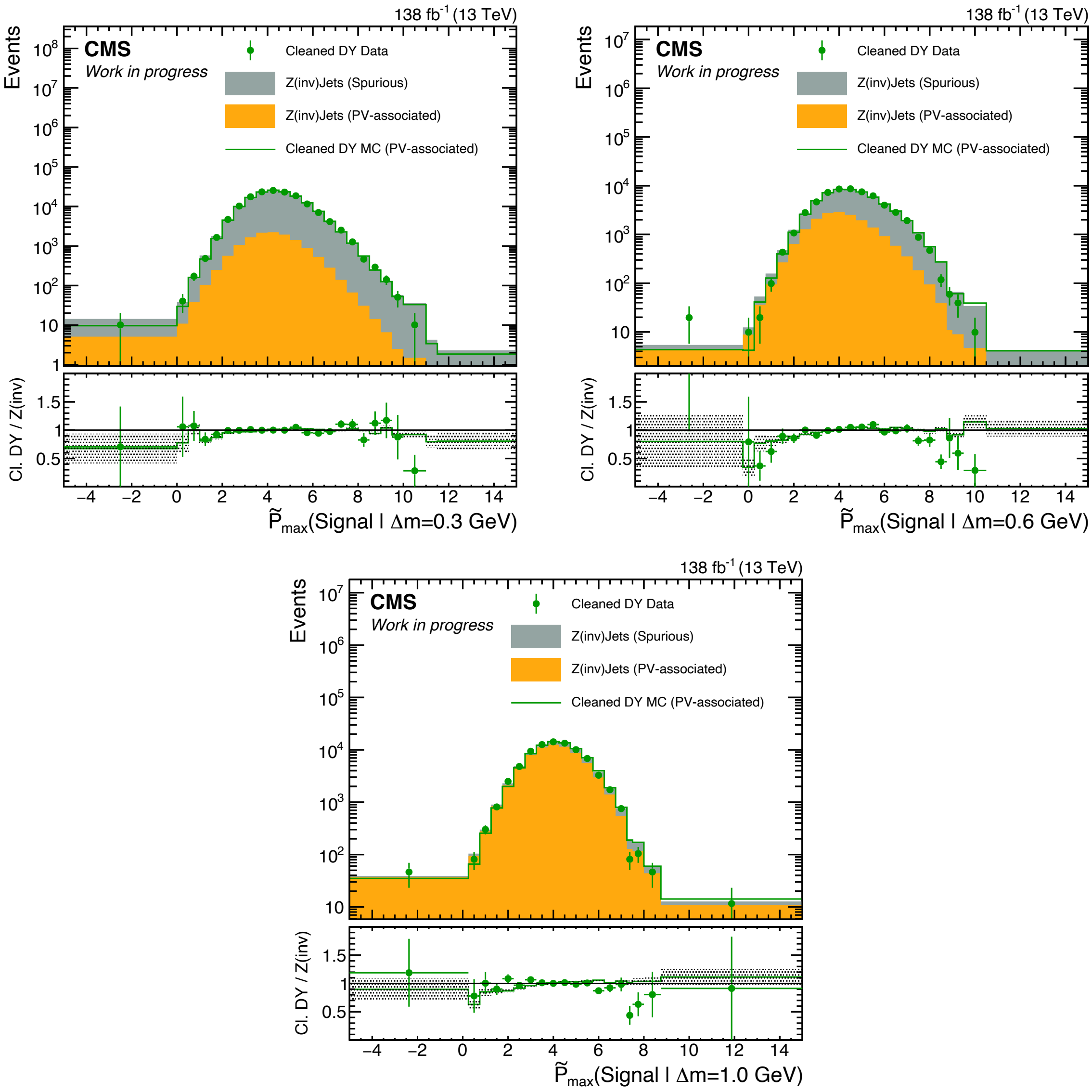

Figure 6.5: Distributions of $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = \ldots\,\mathrm{GeV})$ for $Z \to \nu\nu$ events and cleaned DY (data and MC) scaled to $Z \to \nu\nu$. The distribution for $Z \to \nu\nu$ is divided into whether the highest scoring track is associated to the Spurious or PV-associated track class.

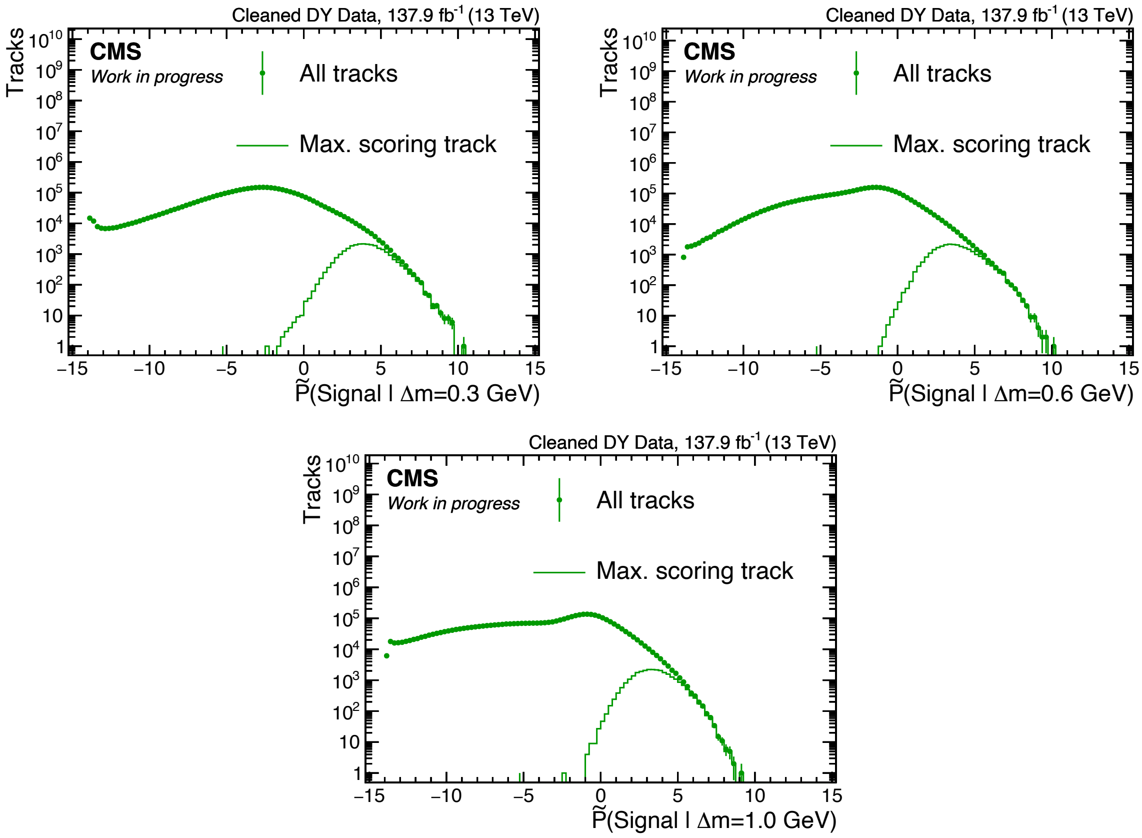

Figure 6.6: Distributions of $\widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV})$, $\widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$, and $\widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = 1.0\,\mathrm{GeV})$ for all tracks in the cleaned DY data set as well as only for the max scoring track per event.

6.2 MC Corrections

To get an accurate prediction of the expected event yields in the signal regions, the MC samples are corrected using data. The correction procedure applied to the SM background samples consists of normalization corrections and shape corrections, uses the cleaned DY control sample and is validated in a single-lepton validation region. Before an extended explanation is given, the following short summary provides the essential details, as well as a roadmap for the section below.

- **Background Normalization Correction:** Scale factors are derived for the PV-associated and Spurious background classes both in the cleaned DY and real $p_{\mathrm{T}}^{\mathrm{miss}}$ regions, as discussed in Section 6.2.1. The scale factor for the background associated to secondaries from tau lepton decays is derived in the context of the signal efficiency study in Section 7. Additional scale factors specific to the $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = ...\,\mathrm{GeV})$ observables are derived in sidebands of the respective observables after the background class-specific scale factors are applied, as discussed below in Section 6.2.1.
- **MC Shape Correction:** Three track variables are simultaneously corrected using a refinement neural network approach. The network is trained in the cleaned DY region after the application of the normalization scale factors and is applied to all simulated tracks (background and signal) in the real $p_{\mathrm{T}}^{\mathrm{miss}}$ region. Additional correction factors for each SR are taken from the residual cleaned DY data/MC ratios using a line fit, as described below in Section 6.2.2.
- **Background $\boldsymbol{p}_{\mathbf{T}}^{\mathbf{miss}}$-based Correction:** Background MC events are reweighted to better model the observed $p_{\mathrm{T}}^{\mathrm{miss}}$ spectrum in data, as shown in Section 6.2.3. This event-level correction is independent of the previous two track-level corrections and thus treated in a factorized way.
- **Signal FastSim Correction:** Simulated signal events are corrected with SR-specific correction factors to compensate for the usage of FastSim, as given in Section 6.2.4. These factors are derived after the track-level refinement is applied.

6.2.1 Background Normalization Correction

Normalization correction factors (scale factors) are derived for the PV-associated and Spurious track classes by comparing inclusive MC and data distributions. To distinguish tracks from the two classes, the longitudinal impact parameter is used; by requiring

$$\log_{10}(\mathrm{dz\ w.r.t.\ PV}) < 10^{-3.0}\,\mathrm{cm} = 0.001\,\mathrm{cm} \tag{6.2}$$

an effectively pure sample of PV-associated tracks can be obtained, whereas requiring

$$\log_{10}(\text{dz w.r.t. PV}) > 10^{-0.2}\,\text{cm} \approx 0.631\,\text{cm} \tag{6.3}$$

isolates Spurious tracks.

Figure 6.7 shows the corresponding distributions used to derive the scale factors for the set of inclusive tracks passing the preselection. The regions mentioned above are the first and third (last) bin in the rebinned histogram. The scale factors (with statistical errors) in the cleaned DY control region are determined to be

- 0.9116 ± 0.0036 for PV-associated and
- 1.1767 ± 0.0012 for Spurious tracks.

A selection of track-level variables is shown in Figs. 6.8 and 6.9 after applying the scale factors.

The corresponding scale factors for the two background classes in the real $p_{\mathrm{T}}^{\mathrm{miss}}$ data are found to be

- 0.82417 ± 0.00086 for PV-associated and
- 1.05799 ± 0.00030 for Spurious tracks.

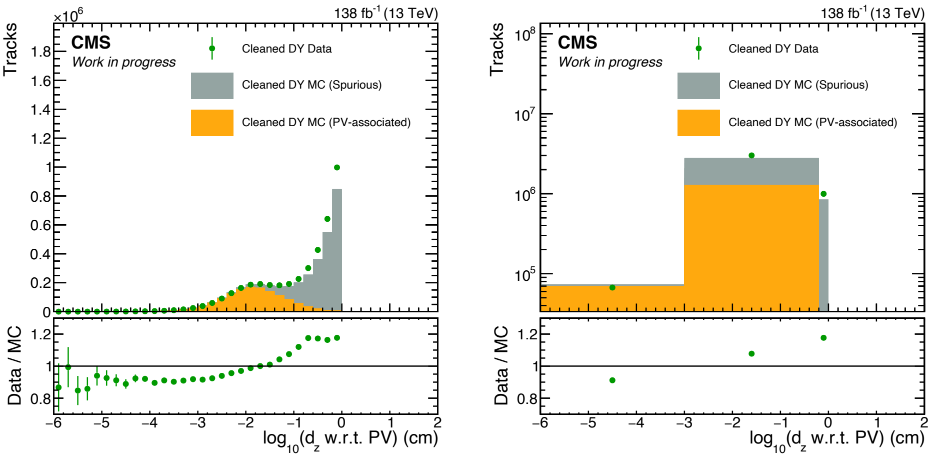

Figure 6.7: Distributions of $\log_{10}$(dz w.r.t. PV) for tracks in (simulated) cleaned DY events. Left: linear y-scale, fine-binning; right: logarithmic y-scale, coarse binning used to derive the scale factors.

Additionally to the global normalization corrections for each track class, signal region-specific scale factors are derived in sidebands of the signal-sensitive observables: $5 < \widetilde{\mathrm{P}}_{\max}(\text{Signal} \,|\, \Delta m = \{0.3,\, 0.6,\, 1.0\}\,\text{GeV}) < 7$. They are compiled in Table 6.1.

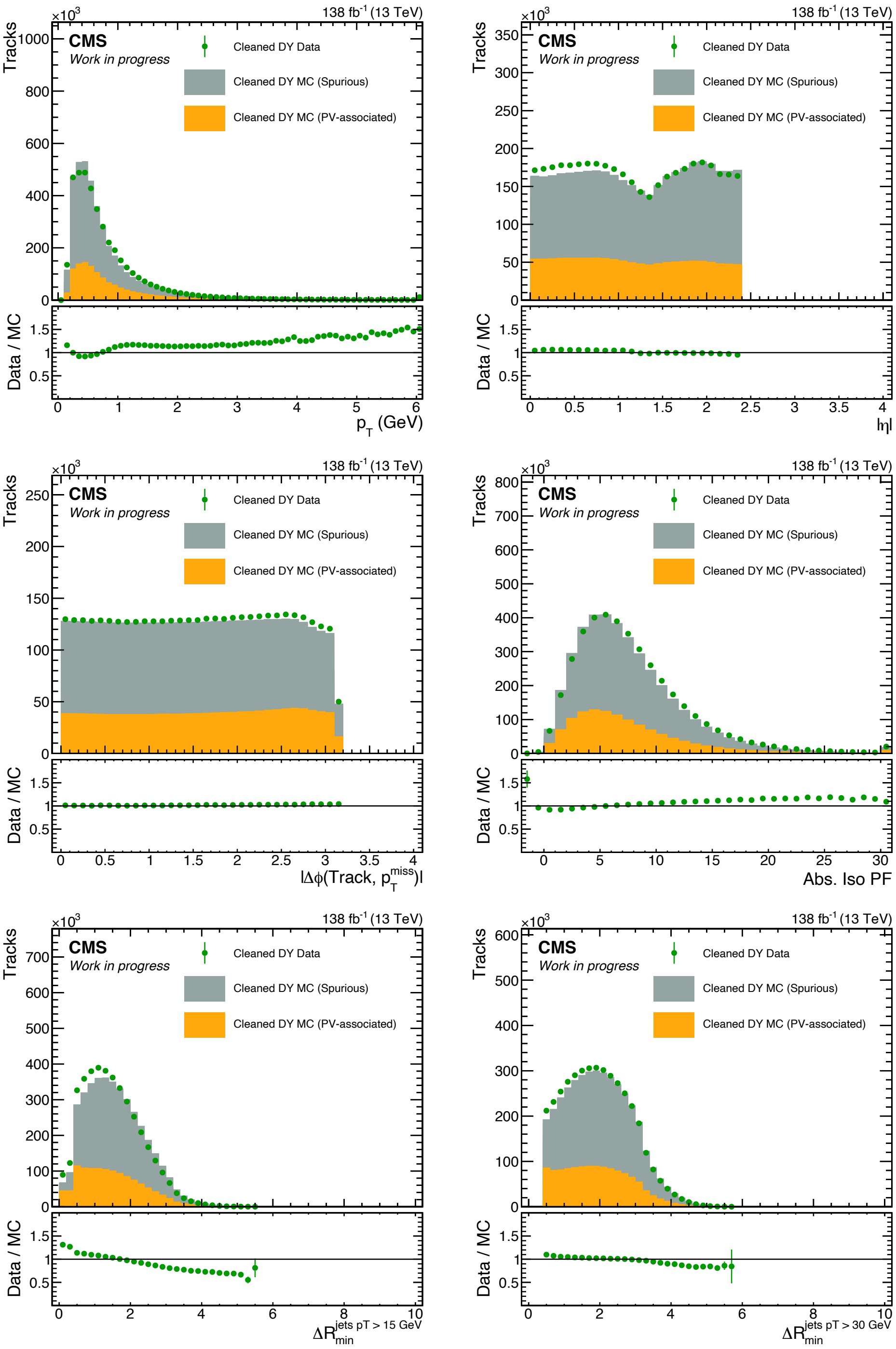

Figure 6.8: Track-level observables for cleaned DY data and MC after application of the scale factors. The overlap of the barrel and forward parts of the tracker detector is visible as a dip in the pseudorapidity distribution at $|\eta| \approx 1.4$.

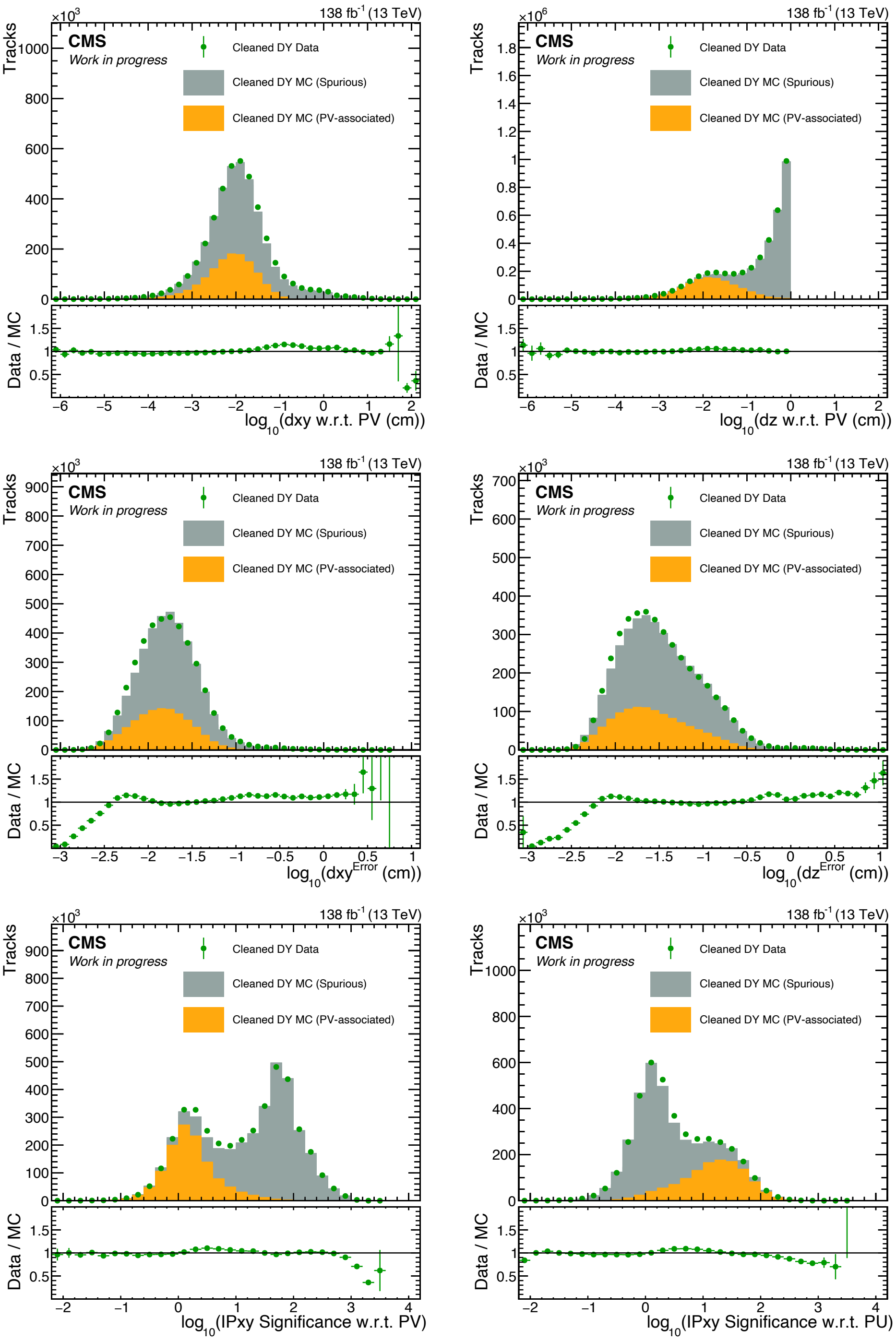

Figure 6.9: Displacement-related track-level observables for cleaned DY data and MC after application of the scale factors.

Table 6.1: Sideband scale factors for the $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = ...\,\mathrm{GeV})$ distributions.

Observable	Scale factor
$\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV})$	1.0103 ± 0.0064
$\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$	1.0618 ± 0.0102
$\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 1.0\,\mathrm{GeV})$	0.9869 ± 0.0096

6.2.2 MC Shape Correction

Figures 6.8 and 6.9 show generally good agreement between data and simulation after applying the normalization correction. Still, to correct for remaining shape differences between the distributions of simulated and observed track variables, a machine learning-based approach is employed. It is realized by applying a regression neural network to the simulated MC samples, which is trained to improve the agreement to data. To obtain a training sample, inclusive sets of tracks passing the preselection are taken from the cleaned DY control region in data and MC. Scale factors are applied to the PV-associated and Spurious tracks in MC, correcting for global efficiency biases; after these corrections, residual discrepancies can be assumed to be due to shape mismodeling.

A neural network equipped with skip connections is now trained to *refine* a set of track observables in MC which are found to be biased and important for the soft track classifier. Refinement (or morphing) means that the values of the observables are changed by the method. This approach contrasts with reweighting approaches where the values are unchanged but object- or event-level weights are determined to make up for discrepancies in distributions. The refiner neural network is implemented as a feedforward regression network (see Fig. 6.11), taking the MC values as inputs and trained to output refined values such that their distributions better resemble the target distributions observed in data. This is achieved by training the network with the biased estimator of the maximum mean discrepancy (MMD) [45] as the loss function. The MMD is a multidimensional two-sample test measuring the similarity between two sets of data points of batch size m and is calculated as:

$$\mathrm{MMD}(\theta)_\mathrm{b} = \frac{1}{m^2}\sum_{i,j=1}^{m} k(\mathbf{x}_i,\mathbf{x}_j) + \frac{1}{m^2}\sum_{i,j=1}^{m} k(\hat{\mathbf{x}}_i(\theta),\hat{\mathbf{x}}_j(\theta)) - \frac{2}{m^2}\sum_{i,j=1}^{m} k(\mathbf{x}_i,\hat{\mathbf{x}}_j(\theta))\,. \tag{6.4}$$

Here, $\mathbf{x}_i$ are vectors taken from the data distribution whereas $\hat{\mathbf{x}}_i$ are the outputs of the refiner which depend on the trainable network parameters θ.

A Gaussian kernel function is used,

$$k(\mathbf{a},\mathbf{b}) = \exp\left(-\sum_{d=1}^{D}\frac{(b^d-a^d)^2}{\sigma_d^2}\right)\,, \tag{6.5}$$

with individual bandwidths σ_d for each dimension d. The notation a^d refers to the d-th component of a vector $\mathbf{a}$. The bandwidths are set to the median distance between all pairs of input and target samples in the corresponding dimension.

Further details of the method can be found in Part I of this thesis.

For the shape correction, three variables are refined: the track p_T, $\log_{10}(\mathrm{dxy}^\mathrm{Error})$, and $\log_{10}(\mathrm{dz}^\mathrm{Error})$, whose distributions are shown in Fig. 6.12. For the training, p_T is transformed by $p_\mathrm{T} \mapsto \log_{10}(p_\mathrm{T})$. In principle, more variables could be added but in order to keep the loss function sensitive also to the tails of the distributions and not “wash out” the training set over many dimensions, the correction is only applied to these three dimensions as they are (a) found to be important to the soft track classifier (Figs. 4.7 and 4.8) and (b) exhibit mismodeling in the data to MC comparisons (see Figs. 6.8 and 6.9). Table 6.2 shows the hyperparameters of the refiner network along with the training parameters. Figure 6.10 displays the evolution of the loss during training. To study and minimize the influence of the random initialization of weights and the random sampling of batches during training, five variations of the refiner network are trained with the same settings. For evaluation, the mean output of the five networks is taken to be the nominal refined value.

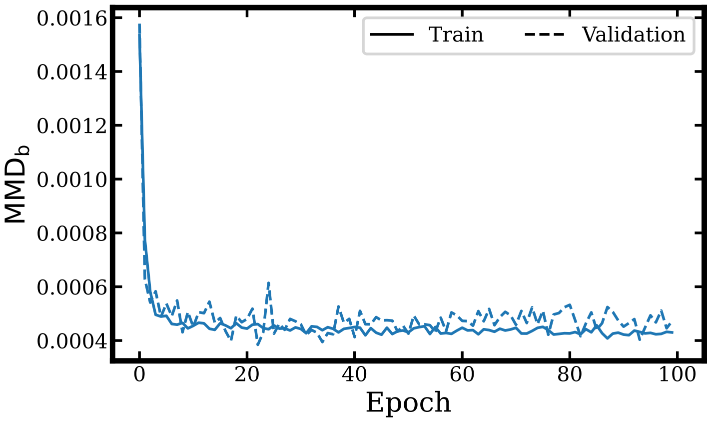

Figure 6.10: Learning curve for the refinement network used for the MC shape correction.

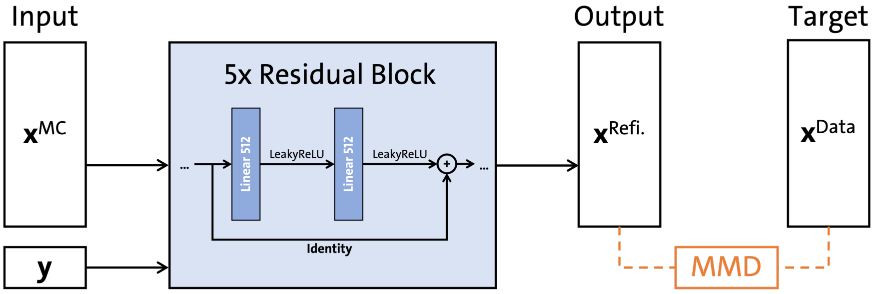

Figure 6.11: Sketch of the training setup for the MC shape correction strategy using a refinement neural network.

The results of the refinement applied to the three variables are shown in Fig. 6.12. It is evident that the correction leads to improved data/MC agreement over the full range of p_T as well as in the signal-like regime of small to medium values of $\log_{10}(\mathrm{dxy}^{\mathrm{Error}})$ and $\log_{10}(\mathrm{dz}^{\mathrm{Error}})$. A worsened agreement is observed for very large values of $\log_{10}(\mathrm{dxy}^{\mathrm{Error}})$ and $\log_{10}(\mathrm{dz}^{\mathrm{Error}})$, however, this exotic phase space region has negligible impact on the analysis.

The MC shape correction is subsequently propagated to the signal node scores $\widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = ...\,\mathrm{GeV})$ which are used to define the signal regions. This is done by re-evaluating the soft track classifier taking as input the refined track variables. The results are shown in Fig. 6.13. Also in the signal-like tails of these distributions, an improved data/MC agreement is observed despite the fact that those observables are not directly refined but only change via the modified input variables to the soft track classifier. Furthermore, Fig. 6.14 shows the corresponding event-level $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = ...\,\mathrm{GeV})$ distributions, that are used for the signal regions.

To account for small remaining discrepancies observed between the data and MC distributions of the max score observables, two line fits are performed to the data/MC ratio in the cleaned DY region using the fit ranges $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = ...\,\mathrm{GeV}) > 5$ and $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = ...\,\mathrm{GeV}) > 7$, see Fig. 6.15 and Tables 6.3 and 6.4. Background events in the signal regions are reweighted with the function value of the line fit in the looser region evaluated at the low edge of the signal region. A systematic uncertainty is applied that is quantified by the difference between this function value and the corresponding function value using the fit in the tighter region, see Section 7.

Table 6.2: Hyperparameters corresponding to the architecture and training of the MC shape correction refiner neural network.

Parameter	Value
Number of residual blocks	5
Hidden layers per residual block	2
Nodes per hidden layer	512
Activation function	LeakyReLU ($\alpha = 0.01$)
Loss function	MMD
Kernel bandwidths σ_d	0.083, 0.109, 0.193 for $\log_{10}(p_\mathrm{T})$, $\log_{10}(\mathrm{dxy}^\mathrm{Error})$, $\log_{10}(\mathrm{dz}^\mathrm{Error})$, resp.
Learning rate	$1{\times}10^{-5}$
Optimizer	Adam
Number of epochs	100
Batch size	4096
Number of batches for training, testing, validation	350, 25, 25
Total data set size	1638400

Table 6.3: Line slopes and intercepts fitted to the cleaned DY data/MC ratios in Fig. 6.15 in the signal-sensitive regions $\widetilde{\mathrm{P}}_\mathrm{max}(\mathrm{Signal}\,|\,\Delta m = \{0.3,\ 0.6,\ 1.0\}\,\mathrm{GeV}) > 5$.

Observable	Fitted slope	Fitted intercept
$\widetilde{\mathrm{P}}_\mathrm{max}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV})$	-0.033 ± 0.021	1.27 ± 0.12
$\widetilde{\mathrm{P}}_\mathrm{max}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$	-0.046 ± 0.024	1.38 ± 0.14
$\widetilde{\mathrm{P}}_\mathrm{max}(\mathrm{Signal}\,|\,\Delta m = 1.0\,\mathrm{GeV})$	-0.003 ± 0.037	1.07 ± 0.20

Table 6.4: Line slopes and intercepts fitted to the cleaned DY data/MC ratios in Fig. 6.15 in the signal-sensitive regions $\widetilde{\mathrm{P}}_\mathrm{max}(\mathrm{Signal}\,|\,\Delta m = \{0.3,\ 0.6,\ 1.0\}\,\mathrm{GeV}) > 7$.

Observable	Fitted slope	Fitted intercept
$\widetilde{\mathrm{P}}_\mathrm{max}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV})$	-0.146 ± 0.073	2.23 ± 0.57
$\widetilde{\mathrm{P}}_\mathrm{max}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$	-0.192 ± 0.062	2.54 ± 0.48
$\widetilde{\mathrm{P}}_\mathrm{max}(\mathrm{Signal}\,|\,\Delta m = 1.0\,\mathrm{GeV})$	-0.042 ± 0.109	1.32 ± 0.83

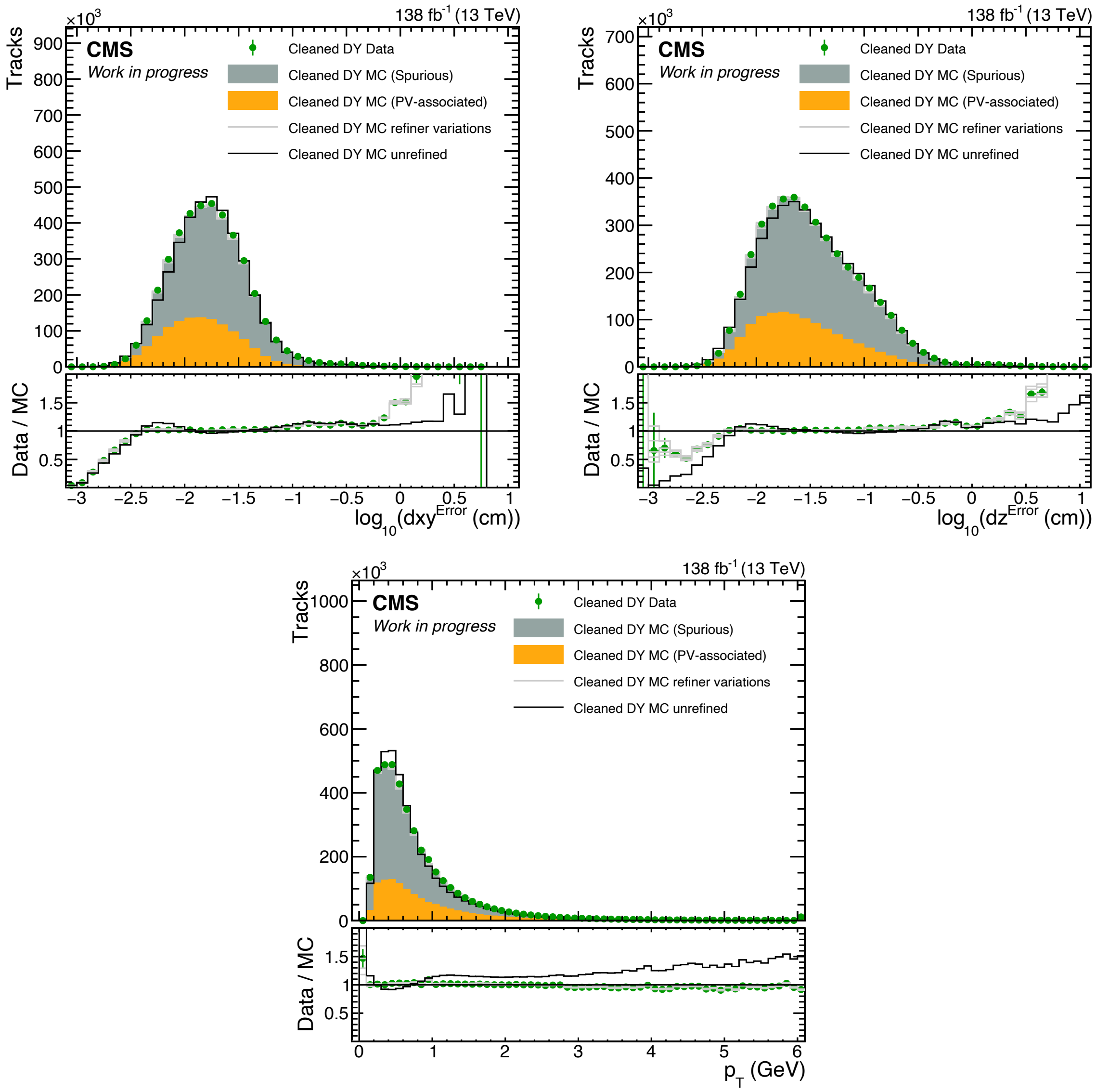

Figure 6.12: Distributions of the three track variables to which the MC shape correction (refinement) is applied. The distributions for data are shown alongside those for refined and unrefined MC in the cleaned DY region. In the lower panel, the green dots indicate the ratio of data to refined MC, the black line indicates the ratio of data to unrefined MC, and the gray lines show the ratio of data to each of the five refinement variations.

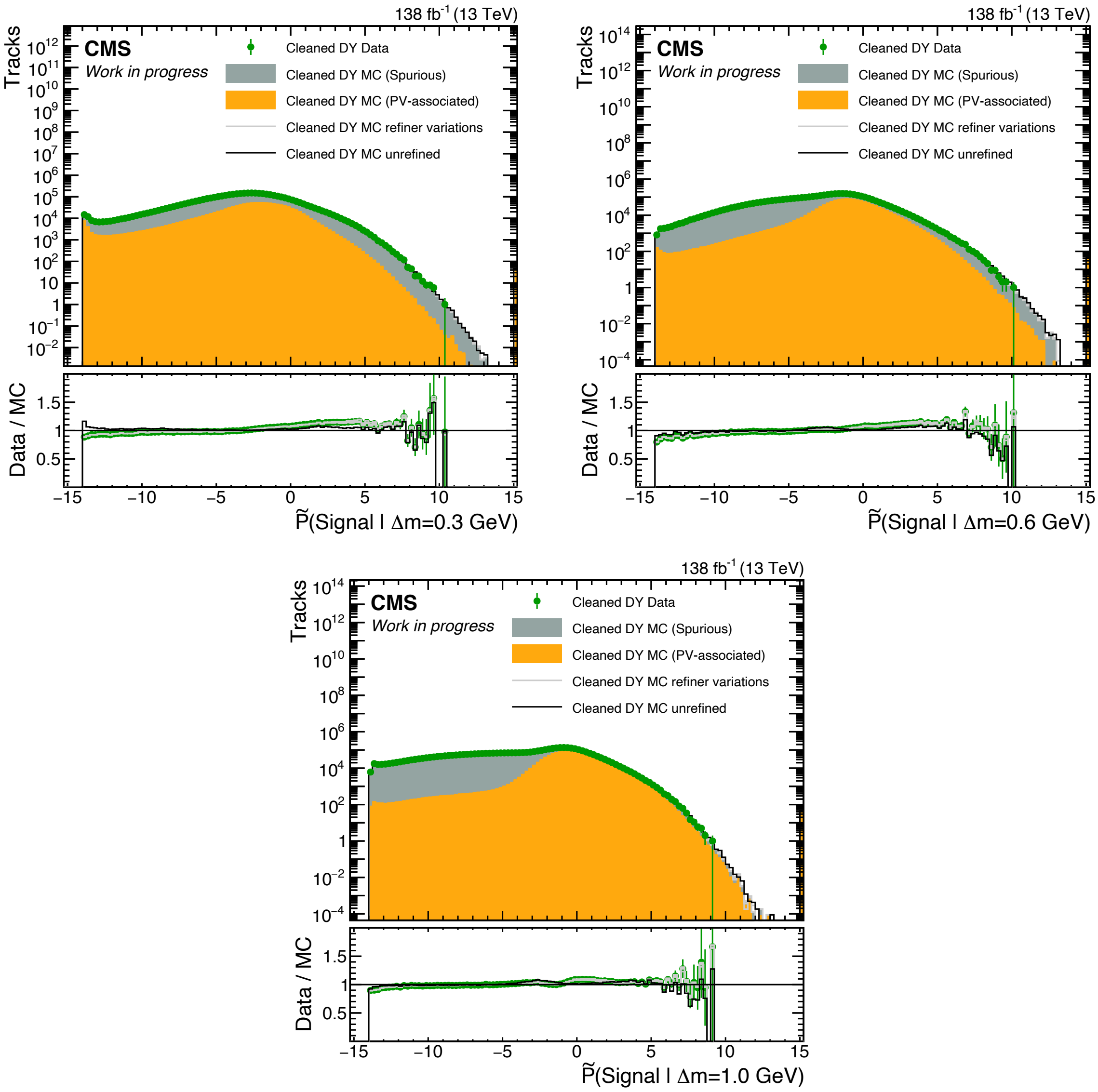

Figure 6.13: Track-level distributions of $\widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV})$, $\widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$, and $\widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = 1.0\,\mathrm{GeV})$ for tracks in data as well as in MC with and without refinement applied to the input variables in the cleaned DY region. In the lower panel, the green dots indicate the ratio of data to refined MC, the black line indicates the ratio of data to unrefined MC, and the gray lines show the ratio of data to each of the five refinement variations.

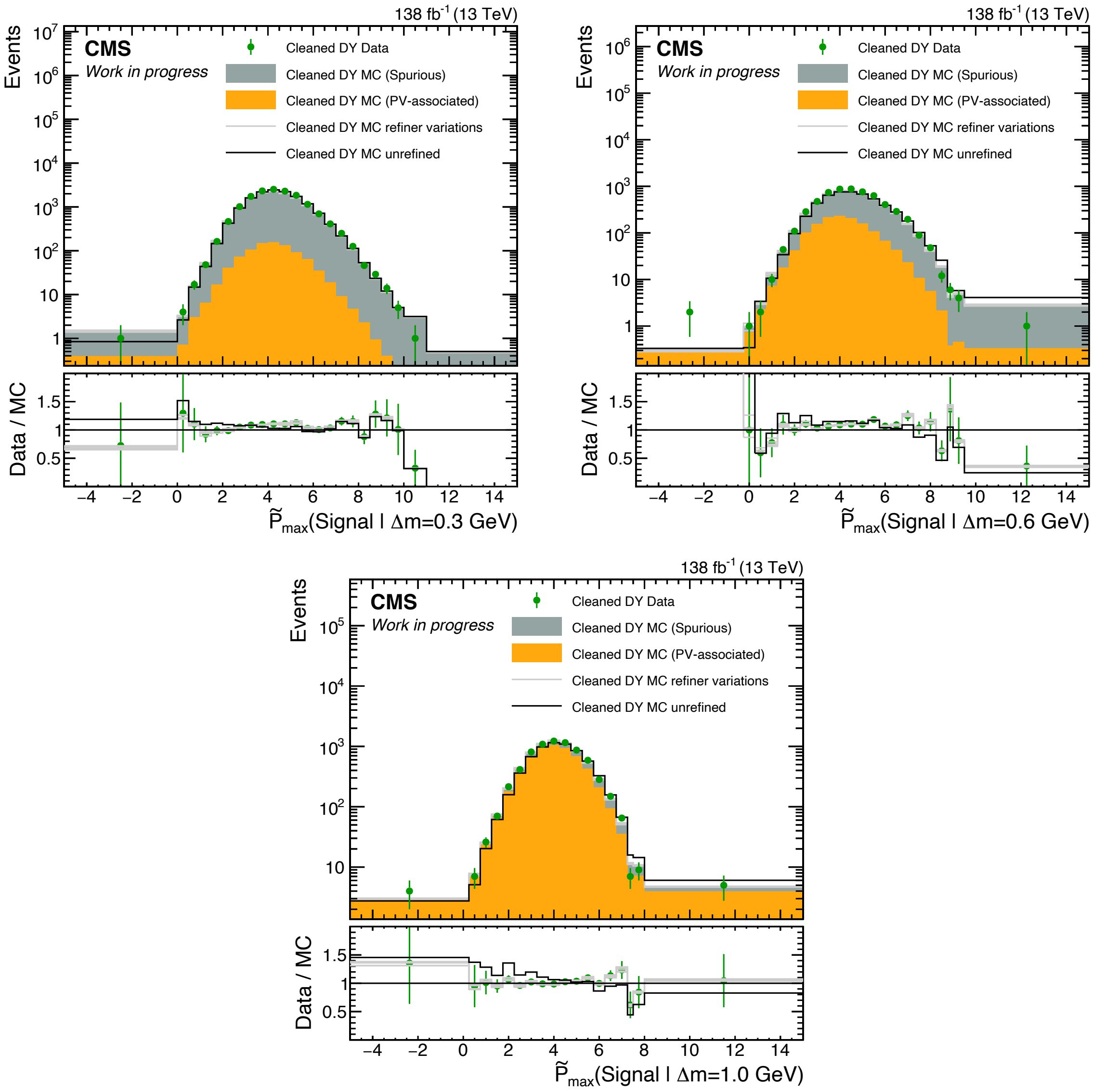

Figure 6.14: Event-level distributions of $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV})$, $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$, and $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 1.0\,\mathrm{GeV})$ for events in data as well as in MC with and without refinement applied to the input variables in the cleaned DY region. In the lower panel, the green dots indicate the ratio of data to refined MC, the black line indicates the ratio of data to unrefined MC, and the gray lines show the ratio of data to each of the five refinement variations. The three tightest bins in each distribution correspond to the signal regions defined in Section 5.3 (the two tightest signal regions are merged into one bin to increase the statistics).

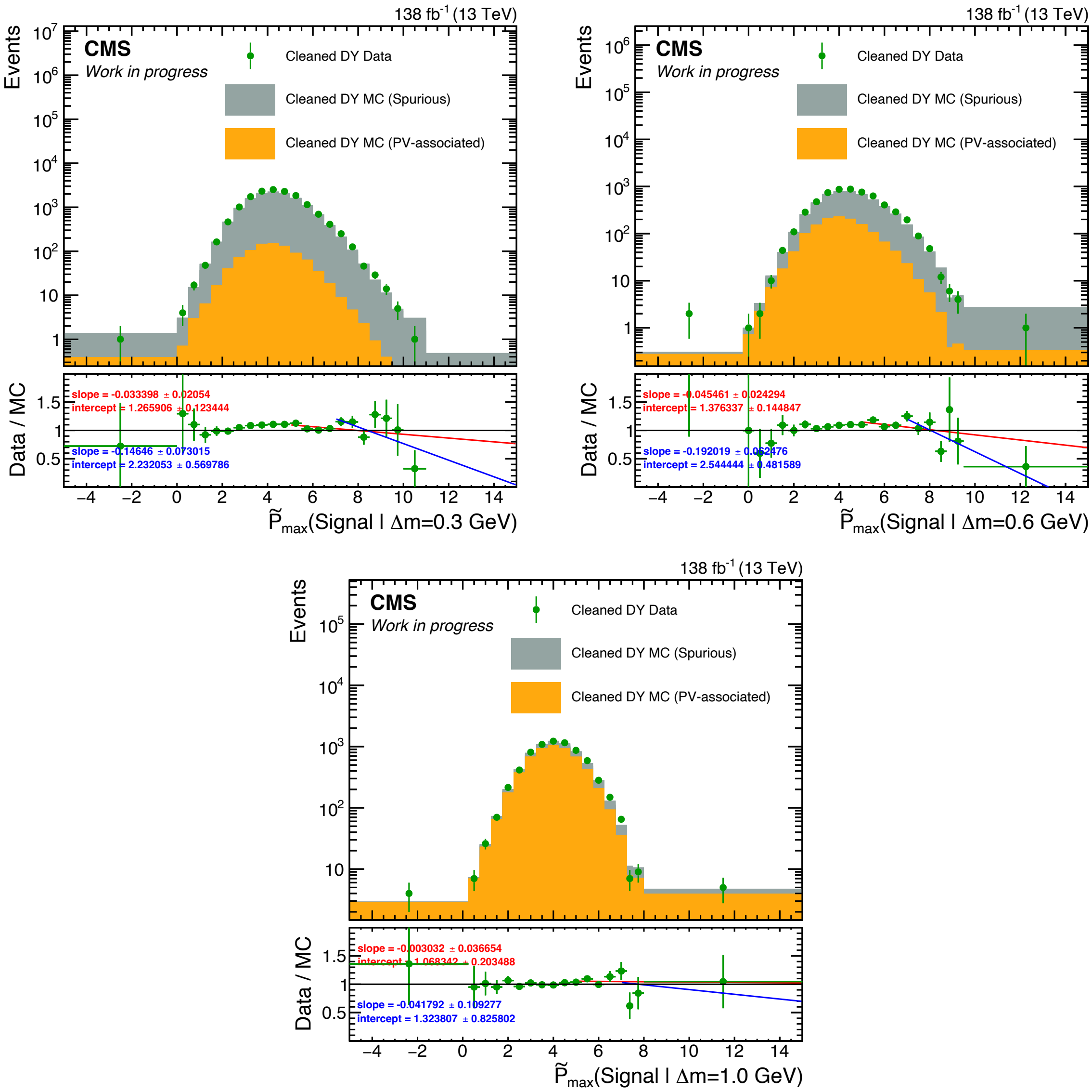

Figure 6.15: Event-level distributions of $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV})$, $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$, and $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 1.0\,\mathrm{GeV})$ for events in data as well as in refined MC in the cleaned DY region, as also shown in Fig. 6.14. Linear fits to the data/MC ratio are shown for two ranges of the observables, once for $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = \{0.3, 0.6, 1.0\}\,\mathrm{GeV}) > 5$ (red) and once for $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = \{0.3, 0.6, 1.0\}\,\mathrm{GeV}) > 7$ (blue).

6.2.3 Background $p_\mathrm{T}^\mathrm{miss}$-based Correction

The background MC samples are further corrected to improve the agreement with the $p_\mathrm{T}^\mathrm{miss}$ distribution observed in data. For this, a line fit is performed to events passing the baseline event selection (but no cut on the signal node scores) in the region $300\,\mathrm{GeV} < p_\mathrm{T}^\mathrm{miss} < 1000\,\mathrm{GeV}$ and simulated events are reweighted according to the corresponding function value. Events with $p_\mathrm{T}^\mathrm{miss} > 1000\,\mathrm{GeV}$ are weighted with the value at $p_\mathrm{T}^\mathrm{miss} = 1000\,\mathrm{GeV}$. This procedure is safe with respect to sensitivity to signal events since the signal regions only make up a very small fraction of all high-$p_\mathrm{T}^\mathrm{miss}$ events.

The distributions and outcome of the fit can be seen in Fig. 6.16. Figure 6.17 shows the distributions of $p_\mathrm{T}^\mathrm{miss}$ and $\widetilde{\mathrm{P}}_\mathrm{max}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$ for events categorized into the SR 0.6 distribution before and after applying the weights.

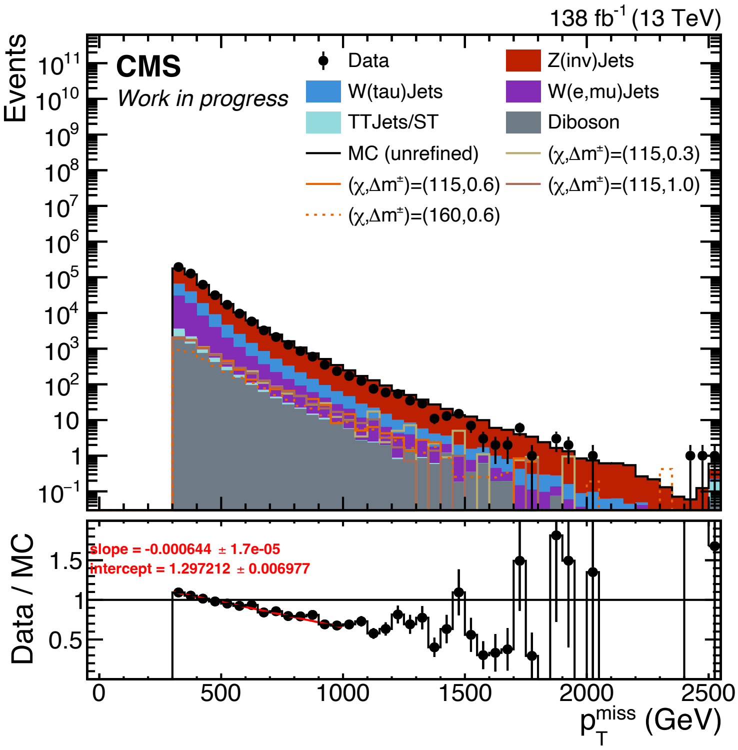

Figure 6.16: Distributions of $p_\mathrm{T}^\mathrm{miss}$ used to derive the correction weights.

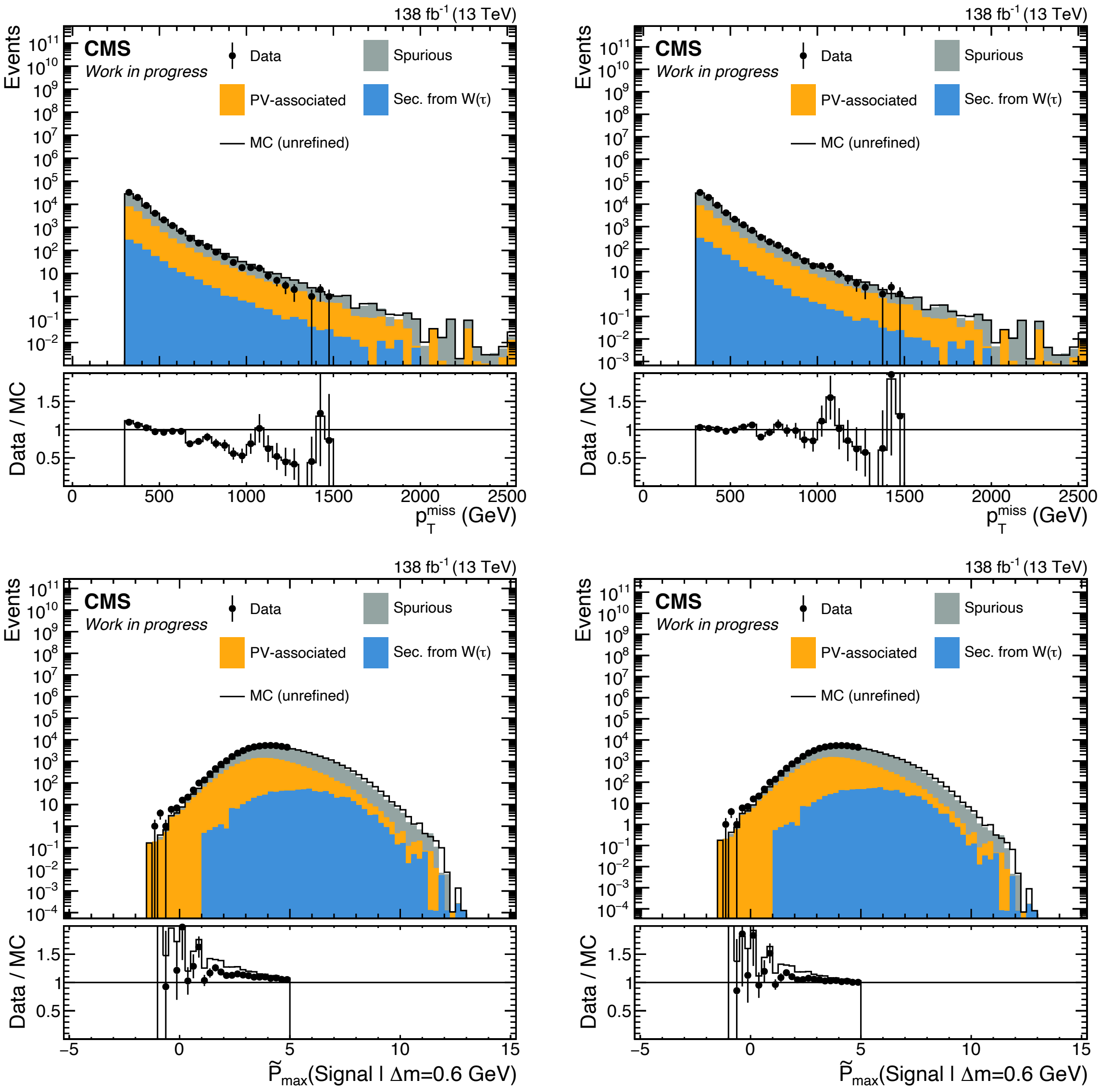

Figure 6.17: Distributions of $p_{\mathrm{T}}^{\mathrm{miss}}$ (top) and $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$ (bottom) for background MC and data for events categorized into the SR 0.6 distribution before (left) and after (right) applying the $p_{\mathrm{T}}^{\mathrm{miss}}$-based correction weights. The data histogram for $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$ is blinded beyond 5. The lower panels show the ratio of data to unrefined MC as a black line and data to refined MC as black dots.

6.2.4 Signal FastSim Correction

To derive corrections for the signal samples compensating for mismodeling in FastSim, the signal is also simulated with FullSim for the 2018 era of data taking. Distributions of observables for three groups of model points corresponding to three different mass splittings ($\Delta m^{\pm} = 0.3\,\text{GeV}$, $\Delta m^{\pm} = 0.6\,\text{GeV}$, and $\Delta m^{\pm} = 1.0\,\text{GeV}$) are compared. For each of these mass splittings, the model points for all chargino masses are pooled: $m(\widetilde{\chi}_1^{\pm}) = \{100, 115, 140, 140, 160, 180, 200, 250, 300, 500\}$ GeV.

Firstly, Fig. 6.18 shows good agreement in the simulation of standard event-level quantities. Still, in Fig. 6.19 discrepancies in the $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = \ldots\,\mathrm{GeV})$ observables used to define the signal regions can be observed. From these, as a first order correction, normalization correction factors are derived by comparing the total yields for each group of signal model points and then taking the average correction factor within all groups for each observable, see Table 6.5. To further correct for different signal region efficiencies in FastSim and FullSim, a second set of correction factors is derived for each signal region individually. For each signal region, the group of signal model points corresponding to the mass splitting for which the classifier is evaluated for, is used. The resulting correction factors are compiled in Table 6.6.

The distributions with both correction factors applied are shown in Fig. 6.20. An improved agreement between the corrected FastSim samples and the FullSim samples can be observed, especially in the diagonal (top left to bottom right) which shows the classifier evaluated for the $\Delta m^{\pm}$ equal to the $\Delta m^{\pm}$ of the probed signal model point.

Table 6.5: FastSim/FullSim normalization correction factors taken from the distributions in Fig. 6.19. The error on the average is calculated as the standard deviation of the three individual estimates divided by $\sqrt{3}$.

	$\Delta m^{\pm}_{\mathrm{true}} = 0.3\,\mathrm{GeV}$	$\Delta m^{\pm}_{\mathrm{true}} = 0.6\,\mathrm{GeV}$	$\Delta m^{\pm}_{\mathrm{true}} = 1.0\,\mathrm{GeV}$	Average
SR 0.3	1.208	1.376	1.597	1.394 ± 0.113
SR 0.6	1.023	0.952	1.282	1.085 ± 0.100
SR 1.0	0.605	0.595	0.604	0.601 ± 0.003

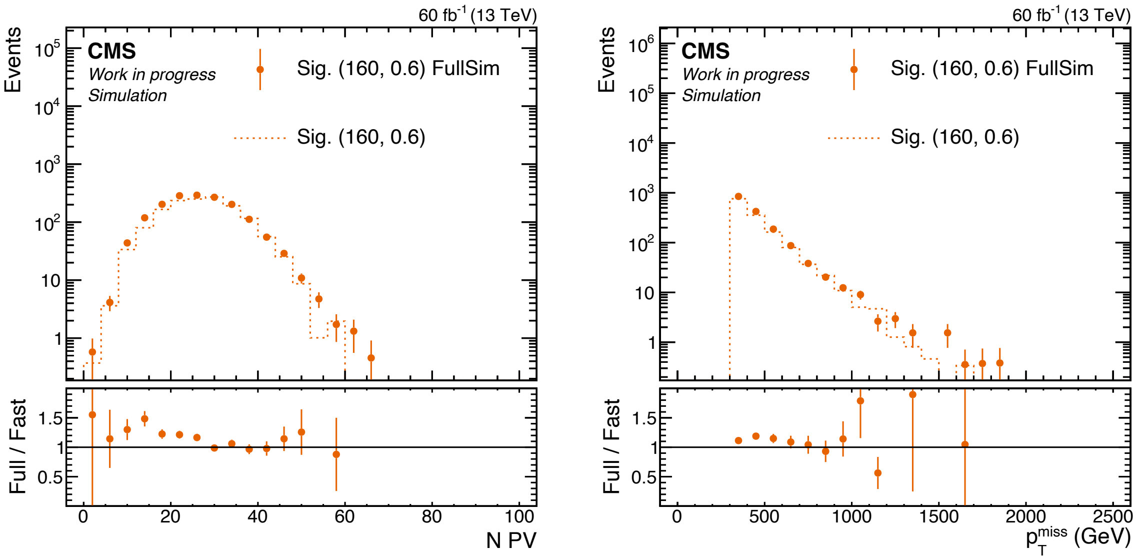

Figure 6.18: Distributions of the number of primary vertices and $p_{\mathrm{T}}^{\mathrm{miss}}$ for a benchmark signal model point simulated with FastSim and with FullSim.

Table 6.6: FastSim/FullSim signal region correction factors with statistical uncertainties taken from the distributions in Fig. 6.19.

		$\Delta m^{\pm}_{\mathrm{true}} = 0.3\,\mathrm{GeV}$	$\Delta m^{\pm}_{\mathrm{true}} = 0.6\,\mathrm{GeV}$	$\Delta m^{\pm}_{\mathrm{true}} = 1.0\,\mathrm{GeV}$
SR 0.3	I	0.358 ± 0.112		
	II	0.514 ± 0.135		
	III	0.155 ± 0.093		
	IV	0.047 ± 0.033		
SR 0.6	I		0.544 ± 0.091	
	II		0.425 ± 0.067	
	III		0.476 ± 0.087	
	IV		0.710 ± 0.218	
SR 1.0	I			0.763 ± 0.122
	II			0.649 ± 0.092
	III			0.711 ± 0.133
	IV			0.418 ± 0.130

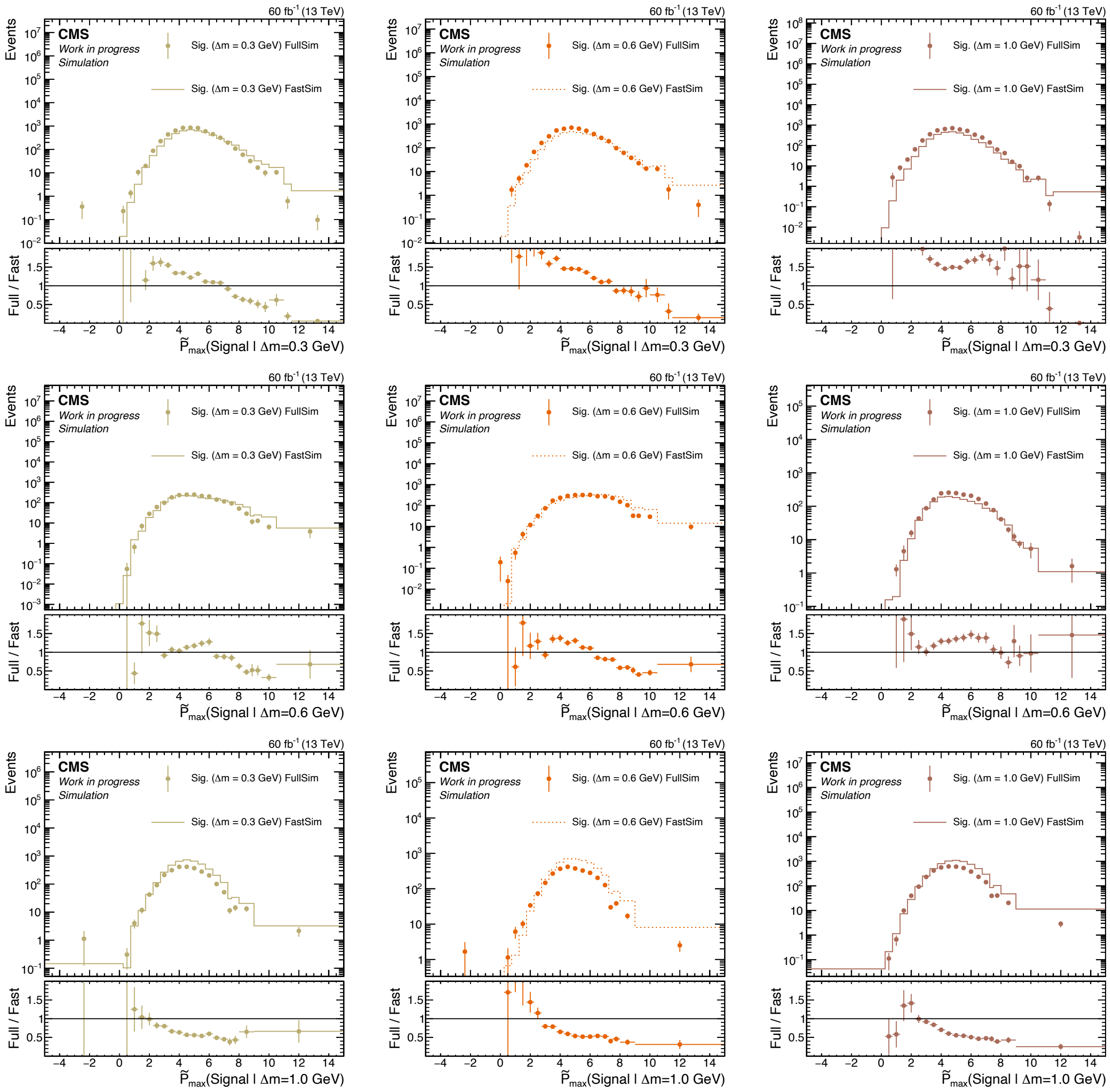

Figure 6.19: Distributions of $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = \{0.3,\, 0.6,\, 1.0\}\,\mathrm{GeV})$ for three signal model mass splittings simulated with FastSim and with FullSim **before** FastSim corrections. For each value of signal model $\Delta m^{\pm}$ (each column), multiple values of $m(\widetilde{\chi}_1^{\pm})$ are pooled: $m(\widetilde{\chi}_1^{\pm}) = \{100, 115, 140, 140, 160, 180, 200, 250, 300, 500\}$ GeV. The last four bins in each histogram correspond to the signal regions for the respective observable, as given in Table 5.4, the remaining bins are not used for the statistical interpretation.

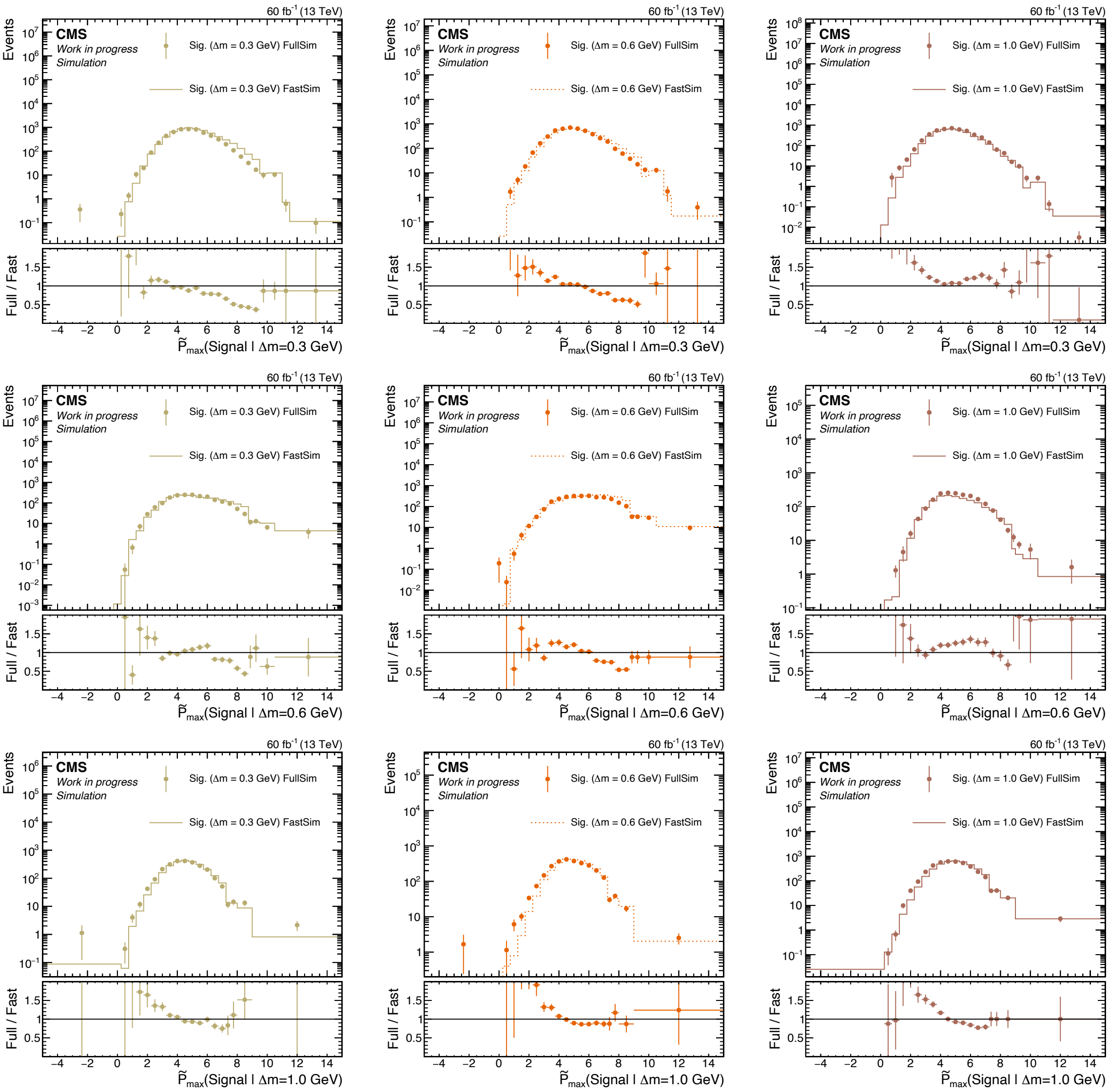

Figure 6.20: Distributions of $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = \{0.3,\,0.6,\,1.0\}\,\mathrm{GeV})$ for three signal model mass splittings simulated with FastSim and with FullSim **after** FastSim corrections are applied. For each value of signal model $\Delta m^{\pm}$ (each column), multiple values of $m(\widetilde{\chi}_1^{\pm})$ are pooled: $m(\widetilde{\chi}_1^{\pm}) = \{100, 115, 140, 140, 160, 180, 200, 250, 300, 500\}$ GeV. The last four bins in each histogram correspond to the signal regions for the respective observable, as given in Table 5.4, the remaining bins are not used for the statistical interpretation.

7 Systematic Uncertainties

The following uncertainties are relevant for the estimates of the yields of the signal and background processes in the signal regions:

- **Background scale/normalization:** The uncertainty in the scale factors for the PV-associated and Spurious backgrounds is determined by deriving alternative scale factors within different sets of tracks, namely the highest scoring tracks for $\Delta m^{\pm} = 0.3\,\text{GeV}$ (for Spurious) and $\Delta m^{\pm} = 1.0\,\text{GeV}$ (for PV-associated) in a sideband region of $5 < \widetilde{\mathrm{P}}_{\mathrm{max}}(\text{Signal}\,|\,\Delta m = 0.3\ (1.0)\,\text{GeV}) < 7$. The nominal and alternative scale factors are compiled in Table 7.1, leading to a variation of 9% (5%) for the Spurious (PV-associated) class; this variation is assigned as a systematic uncertainty for each class. This procedure accounts for a possible bias in the selection of the highest scoring track due to under- or overrepresentation in the MC. However, based on Fig 6.6, it is expected that this effect is minimal in the signal regions. For the subdominant background arising from τ secondaries in $\mathrm{W} \to \tau\nu$ events, the normalization uncertainty is taken to be 10%, determined from the difference of the corresponding scale factor to unity, as seen in Fig. 7.4 and Table 7.2.
- **Background shape:** The uncertainty in the background shape is twofold. First, the influence of the random initialization of weights and random iteration through the training dataset is determined by training the refinement network with the same settings five times and comparing the signal region yields for background. It can be seen in Figs. 6.12, 6.13, and 6.14 that this effect is negligible. Secondly, the residual disagreement between cleaned DY data and MC in the signal regions is assessed as the difference of the two line fits to the ratios in Fig. 6.15. The differences of the corresponding function values at the low edges of the signal region bins are taken as variations of the nominal prediction, yielding values of O(10%), as seen in Table 7.3.
- **Jet energy scale:** Corrections to the jet energy scale simulation are varied using p_T and η-dependent uncertainties derived centrally. The variations are propagated to higher level variables such as jet multiplicity, $p_{\mathrm{T}}^{\mathrm{miss}}$, H_{T} and $H_{\mathrm{T}}^{\mathrm{miss}}$. By re-evaluating the classifier with the varied $p_{\mathrm{T}}^{\mathrm{miss}}$, the influence of the variation on the final signal regions is assessed. Figure 7.1 shows an approximately flat effect of 5% for both signal and background.

- **Pileup modeling:** The effect of potential pileup mismodeling is checked by determining the expected event yields in the $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV})$ signal regions — which are most dominated by tracks from pileup interactions — without applying the weights. Based on the distributions in Fig. 7.2, an uncertainty of 3% is applied.
- **L1 prefiring:** In the 2016 and 2017 data taking periods, an issue with the L1 trigger resulted in possible losses of signal events due to the wrongful association of trigger bits to the previous bunch crossing. The impact of this prefiring on the analysis is evaluated by comparing the nominal $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV})$ distributions for signal and background to the respective distributions after removing events with jets with $p_{\mathrm{T}} > 100\,\mathrm{GeV}$ and $2.25 < |\eta| < 3.0$. Figure 7.3 shows that, towards the signal regions with $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.6\,\mathrm{GeV}) > 8.75$, the effect is $\approx 1\%$ and a systematic uncertainty of 2% is applied.
- **Luminosity:** An uncertainty of 1.6% is applied corresponding to the accuracy of the luminosity measurement as per central recommendation by the CMS Collaboration [73].
- **Signal efficiency:** To get an estimate of the signal efficiency in data compared to MC, soft tracks from tau decays in $\mathrm{W} \rightarrow \tau\nu$ are used as a proxy since the underlying physics process is similar to the signal. Those tau tracks are isolated in the tail of the tau node output distribution $\widetilde{\mathrm{P}}(\mathrm{W}(\tau)\,|\,\Delta m = 0.3\,\mathrm{GeV})$. The discrepancy between data and prediction yields is measured in six regions with a successively tighter selection applied on $\widetilde{\mathrm{P}}(\mathrm{W}(\tau)\,|\,\Delta m = 0.3\,\mathrm{GeV})$. For this procedure, all tracks passing the preselection and a loose cut on $\widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV})$ of -5 are used. The observable is shown in Fig. 7.4 and the results are compiled in Table 7.2. Derived from the deviation of the tau track scale factor from unity, a signal uncertainty of 10% is applied.

Table 7.1: Scale factors derived in the nominal and alternative way. Nominal refers to the inclusive set of tracks passing the preselection whereas the alternative sets consist of only the highest-scoring tracks for $\Delta m^{\pm} = 0.3\,\mathrm{GeV}$ (for the Spurious class) and $\Delta m^{\pm} = 1.0\,\mathrm{GeV}$ (for the PV-associated class) in a sideband region of $5 < \widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = 0.3\ (1.0)\,\mathrm{GeV}) < 7$.

Background class	Nominal ± stat.	Alternative ± stat.	Deviation
PV-associated	0.82417 ± 0.00086	0.783 ± 0.036	5.0%
Spurious	1.05799 ± 0.00030	0.968 ± 0.011	8.5%

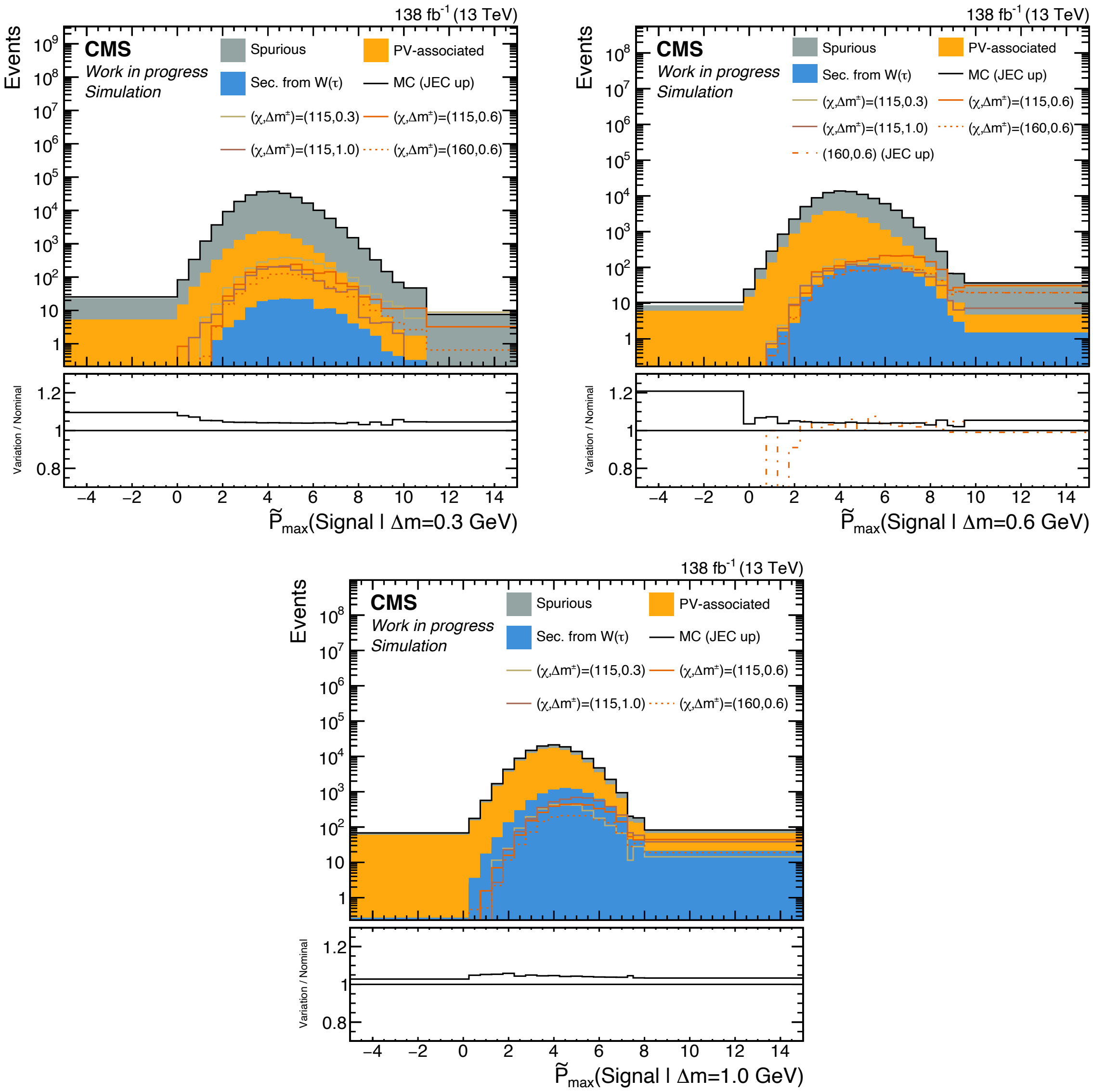

Figure 7.1: Nominal distributions of $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = \{0.3,\, 0.6,\, 1.0\}\,\mathrm{GeV})$ together with the distributions for an up-variation of the jet energy corrections.

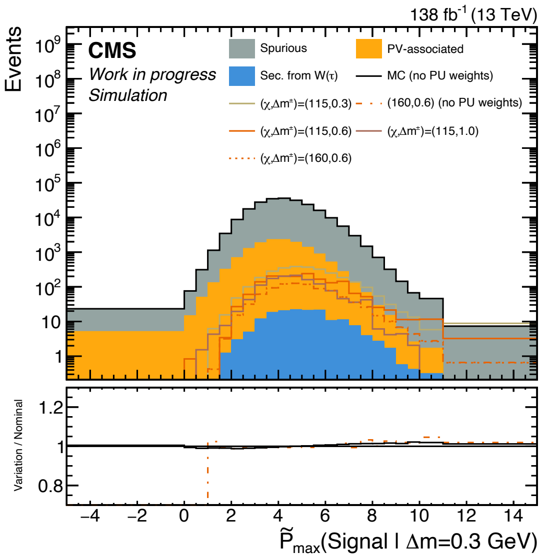

Figure 7.2: Nominal distributions of $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV})$ together with the distributions with no pileup weights applied.

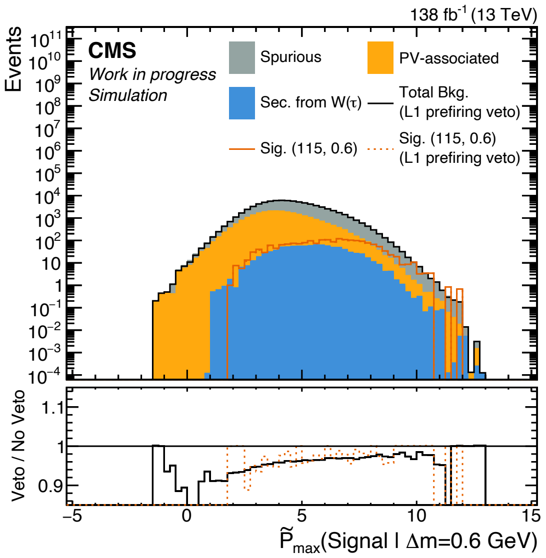

Figure 7.3: Assessment of the impact of the L1 prefiring issue in 2016 and 2017 data taking. The “L1 prefiring veto” refers to the removal of events with jets with $p_{\mathrm{T}} > 100\,\mathrm{GeV}$ and $2.25 < |\eta| < 3.0$.

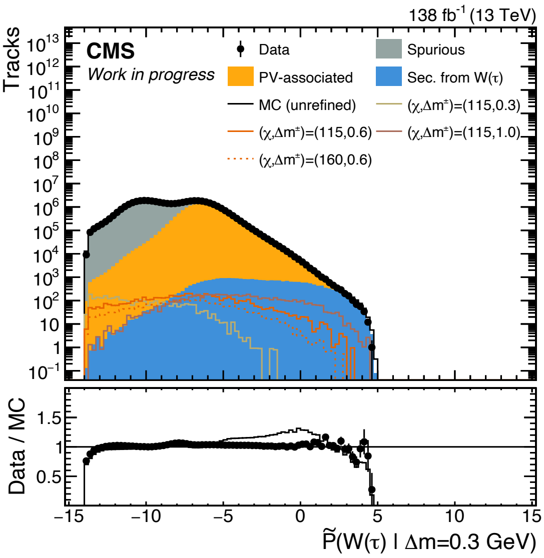

Figure 7.4: Distributions of $\widetilde{\mathrm{P}}(\mathrm{W}(\tau)\,|\,\Delta m = 0.3\,\mathrm{GeV})$ used to determine the tau track scale factor which is also used as an estimate for the uncertainty of the signal efficiency.

Table 7.2: Tau track scale factors computed from the data to MC ratio in the distributions shown in Fig. 7.4.

Cut on $\widetilde{\mathrm{P}}(\mathrm{W}(\tau)\,|\,\Delta m = 0.3\,\mathrm{GeV})$	SF
2	0.975 ± 0.027
2.5	0.957 ± 0.034
3	0.880 ± 0.052
3.5	0.850 ± 0.078
4	0.952 ± 0.166
4.5	0.262 ± 0.279
Weighted mean	0.938 ± 0.112

Table 7.3: Magnitudes of the systematic uncertainties on the background prediction in each signal region. The scale uncertainties for PV-associated and Spurious refer to the difference of the nominal to the alternative scale factors (see Table 7.1), whereas for the tau track background, the scale uncertainty corresponds to the difference of the scale factor to unity (see Table 7.2). The shape uncertainties are derived from two line fits to the data/MC ratio in the cleaned DY region, taking the difference of the function values of the two fitted lines evaluated at the low edge of each signal region as the size of the uncertainty (see Tables 6.3 and 6.4).

Uncertainty		Scale			Shape
Background		PV-associated	Spurious	Tau tracks	All
SR 0.3	I	5%	9%	10%	11%
	II				17%
	III				28%
	IV				34%
SR 0.6	I	5%	9%	10%	12%
	II				15%
	III				23%
	IV				37%
SR 1.0	I	5%	9%	10%	3%
	II				4%
	III				6%
	IV				9%

7.1 Single-lepton Control Region Validation

To validate the background estimation method, a validation region is defined by inverting the lepton veto and requiring exactly one lepton in the event selection. Note that all other baseline selection cuts including the $p_{\mathrm{T}}^{\mathrm{miss}}$ cut of 300 GeV are applied. The single-lepton validation region is expected to be dominated by W $\rightarrow$ lν events, a process very similar to Z $\rightarrow \nu\nu$, the dominant background in the signal regions.

The corresponding distributions of the track-level signal node output scores $\widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = \{0.3,\ 0.6,\ 1.0\}\,\mathrm{GeV})$ are shown in Fig. 7.5 and show good agreement between the data and prediction in the full phase space, including in the signal regions. The impact parameters of tracks in the single-lepton CR are also examined to check for features that may arise from material interactions or other geometric factors. Figure 7.6 shows distributions of the transverse impact parameter (with respect to the leading primary vertex) for all tracks in the single-lepton validation region with three different choices of x-axis limits. The two dimensions are plotted against each other in Fig. 7.7 for a subset of events in data (Run 2018 D) and for simulated W $\rightarrow$ lν events with high H_{T}. For these two-dimensional distributions, a cut of $\widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV}) > 5$ is applied to test the phase space region sensitive to the most displaced signal. None of the figures seems to indicate structure arising from detector material.

7.2 Impacts/Pulls of the Maximum Likelihood Fit

In the maximum likelihood fit, each source of systematic uncertainty is encoded in the likelihood as a log-normal function with a single parameter (nuisance parameter) taken to modify the rate of a given process. The width of the log-normal is estimated as the prefit uncertainty assigned to a given nuisance parameter, as described in Section 7. The robustness of the maximum likelihood fit model is studied using the Asimov data set corresponding to the background-only hypothesis, as well as to the background-plus-signal hypothesis. The results of the maximum likelihood fit are shown in terms of the impact of the fit on each nuisance parameter in Figs. 7.8 and 7.9.

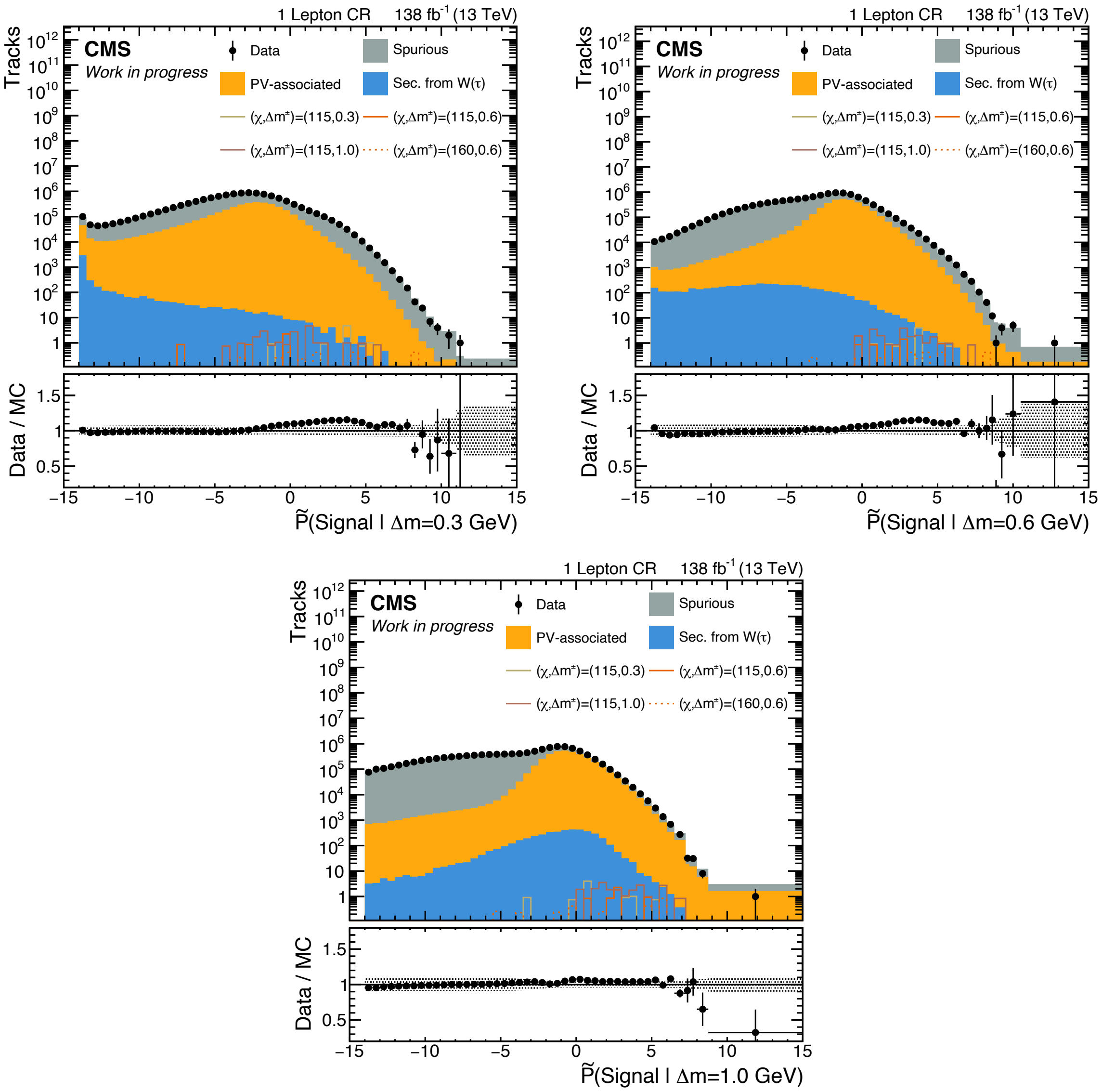

Figure 7.5: $\widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = \{0.3,\, 0.6,\, 1.0\}\,\mathrm{GeV})$ distributions of events in the single-lepton validation region with all MC corrections applied. The last four bins in each distribution correspond to the signal regions. The systematic uncertainties are shown as shaded areas in the lower panels, the black error bars indicate the statistical uncertainty in the data.

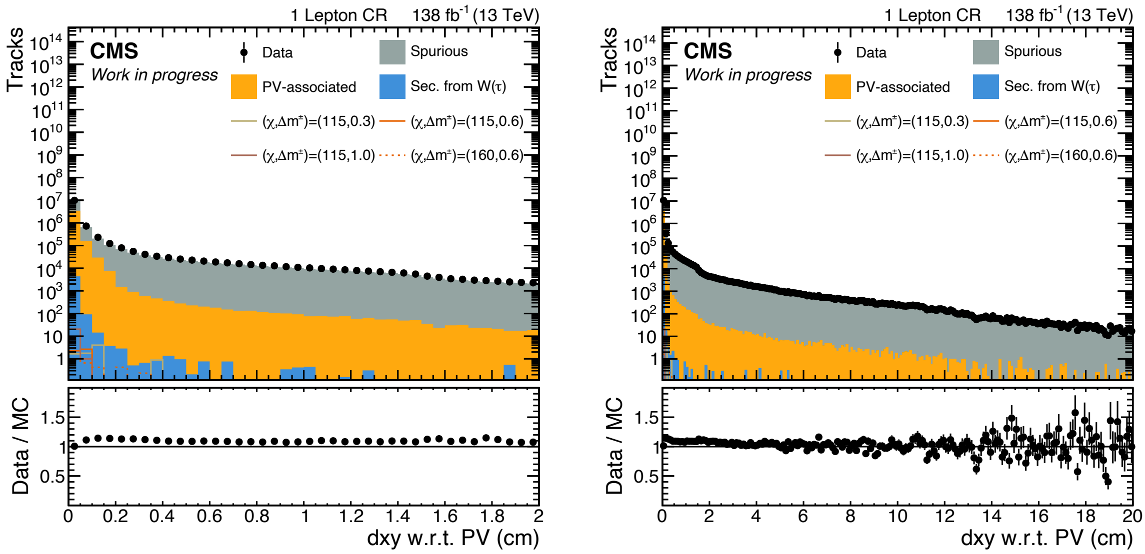

Figure 7.6: Distributions of the transverse impact parameter for tracks in data and MC in the single-lepton validation region. Two different x-axis limits are shown.

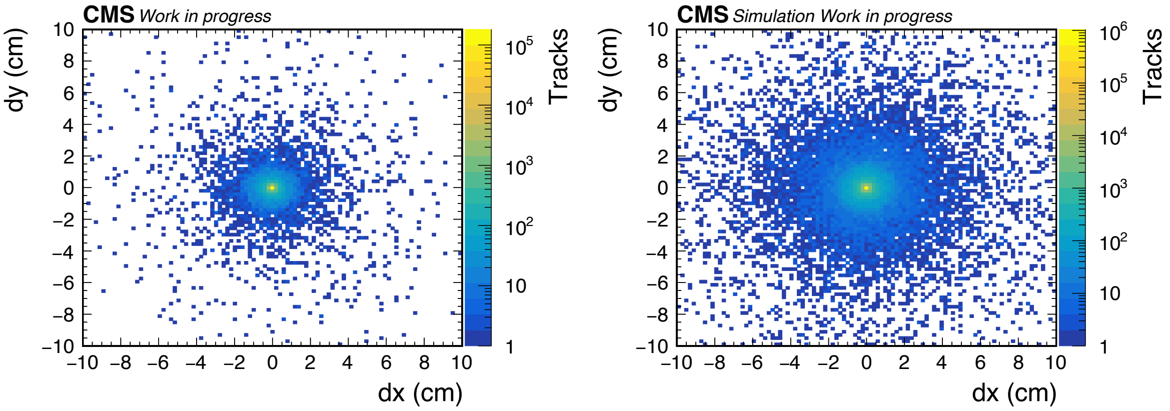

Figure 7.7: Two-dimensional distributions of dx versus dy for tracks with $\widetilde{\mathrm{P}}(\mathrm{Signal}\,|\,\Delta m = 0.3\,\mathrm{GeV}) > 5$ in data (left) and $\mathrm{W} \to \mathrm{l}\nu$ MC (right) in the single-lepton validation region

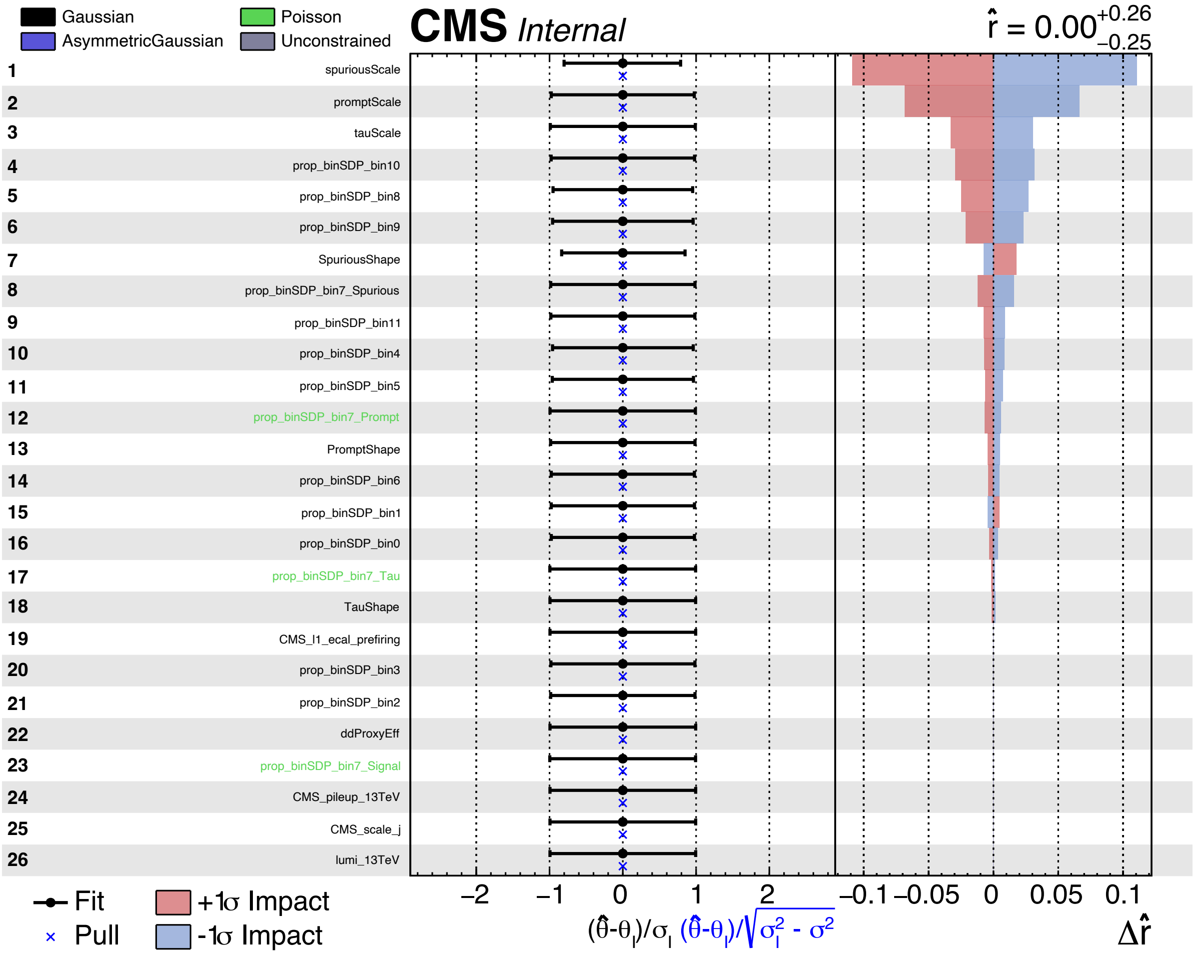

Figure 7.8: Impact on the nuisance parameters of the background-only Asimov data set. The test model corresponds to $m(\tilde{\chi}_1^{\pm}) = 115$ GeV, $\Delta m(\tilde{\chi}_1^{\pm}, \tilde{\chi}_1^0) = 0.768$ GeV.

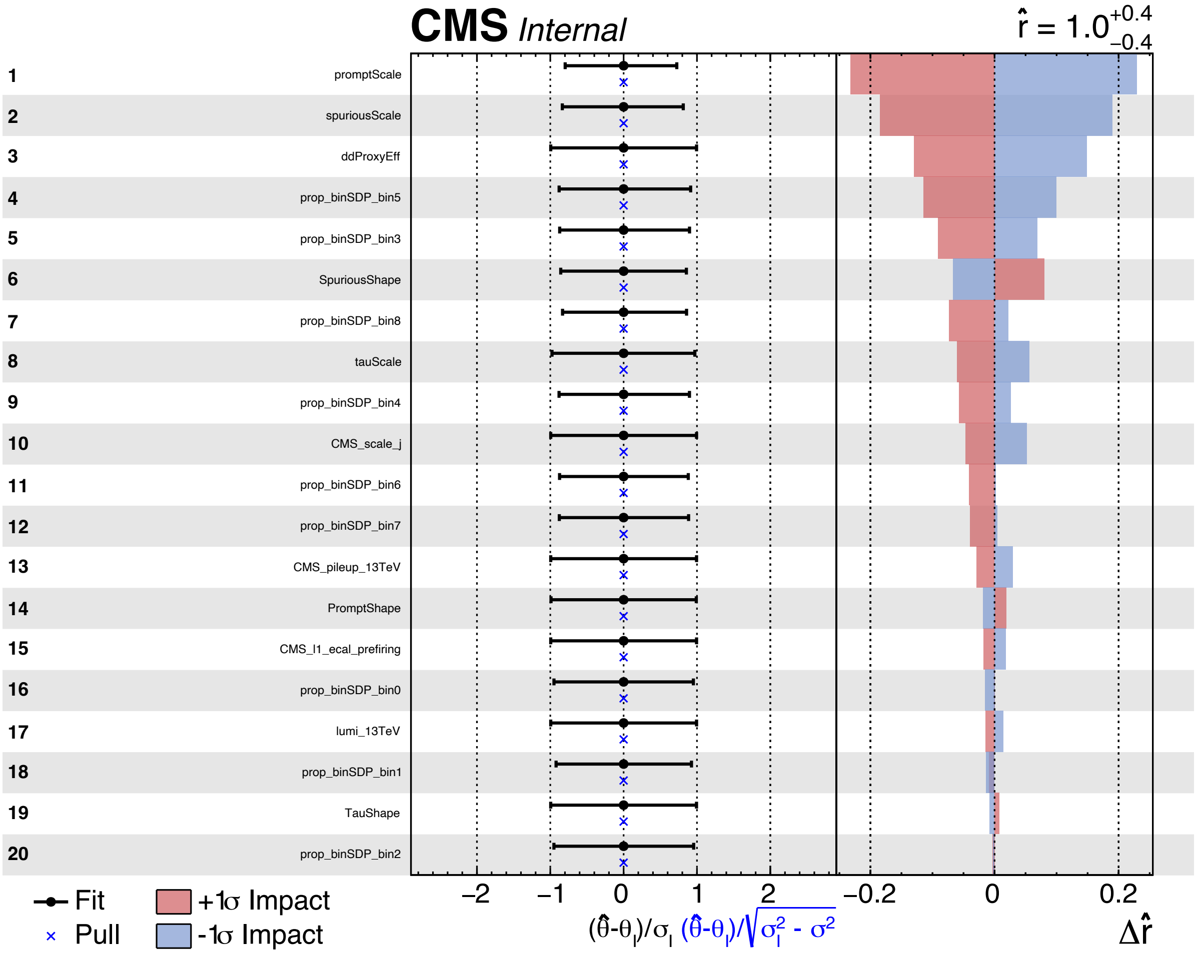

Figure 7.9: Impact on the nuisance parameters of the background+signal Asimov data set. The test model corresponds to $m(\tilde{\chi}_1^{\pm}) = 115$ GeV, $\Delta m(\tilde{\chi}_1^{\pm}, \tilde{\chi}_1^0) = 0.768$ GeV.

8 Results and Interpretation

Figure 8.1 shows the distributions of the $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal}\,|\,\Delta m = \{0.3,\ 0.6,\ 1.0\}\,\mathrm{GeV})$ observables for SM background and various signal model points. Also shown are the observed data histograms. The last four bins in each distribution correspond to the signal regions. The 12 signal regions are additionally compiled in Fig. 8.2. The data are consistent with the background-only model over the full range of the signal-sensitive observables, also in the signal regions.

8.1 Exclusion Limits

In the absence of signs of signal events in the signal regions, exclusion limits are derived based on $\mathrm{CL_s}$ asymptotic limits [74] using the Combine tool [75]. They are shown in Fig. 8.3 in the plane of $\Delta m^{\pm}$ versus chargino mass. The excluded region encompasses $\Delta m^{\pm} \in [0.3, 1.2]\,\mathrm{GeV}$ for $m(\widetilde{\chi}_1^{\pm}) = 110\,\mathrm{GeV}$ and extends to around $m(\widetilde{\chi}_1^{\pm}) = 185\,\mathrm{GeV}$ for $\Delta m^{\pm} = 600\,\mathrm{MeV}$. The impact of the observed data on the nuisance parameters in the maximum likelihood model are shown in Fig. 8.4. The best fit signal strength modifier is consistent with 0 within the uncertainties of the fit. The most significant pull is downward on the Spurious background normalization by 0.9σ.

The maximum likelihood fit leading to the exclusion limits has been carried out by Samuel Bein.

8.2 Comparison to Previous Searches

As expected, the excluded region of this analysis sits between the exclusions of disappearing track and soft lepton searches, see Figs. 1.5 and 1.4. Compared to the limits set by the ATLAS Collaboration using the same signature as this search (Ref. [62] featured in Fig. 1.4), a considerable extension of the ruled out model phase space is observed. The presented search extends the maximum sensitivity to chargino mass by approximately 20 GeV and leads to significantly more extensive sensitivity in the mass splitting $\Delta m^{\pm}$. This improvement can be attributed to the parametrization of the soft track classifier, leading to increased sensitivity to a wide range of $\Delta m^{\pm}$.

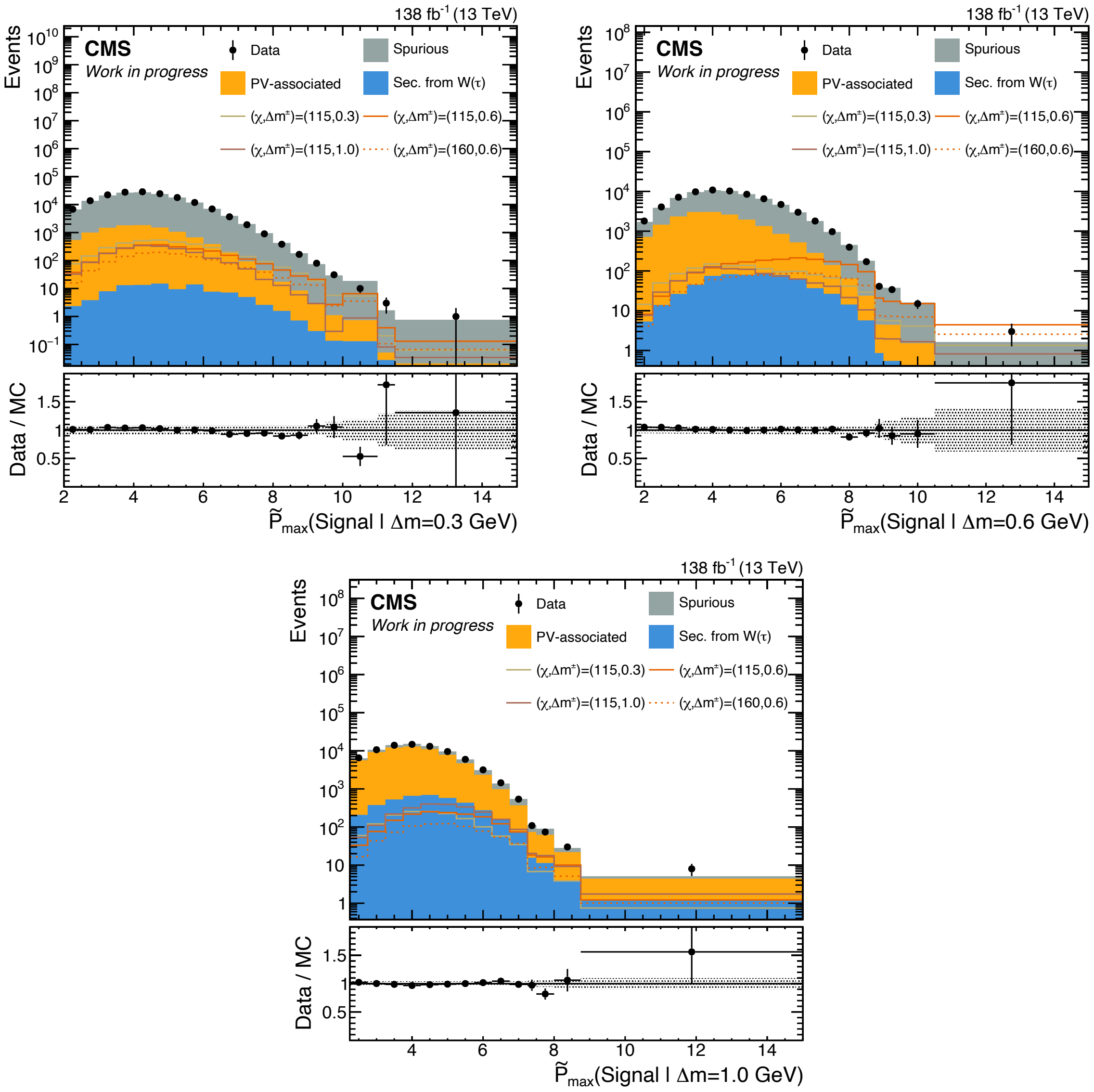

Figure 8.1: Distributions of events in the signal-sensitive observables $\widetilde{\mathrm{P}}_{\mathrm{max}}(\mathrm{Signal} \mid \Delta m = \{0.3,\ 0.6,\ 1.0\}$ GeV) for expected background, signal, and observed data. The systematic uncertainties are shown as shaded areas in the lower panels, the black error bars indicate the statistical uncertainty in the data. The MC distributions are shown pre-fit.

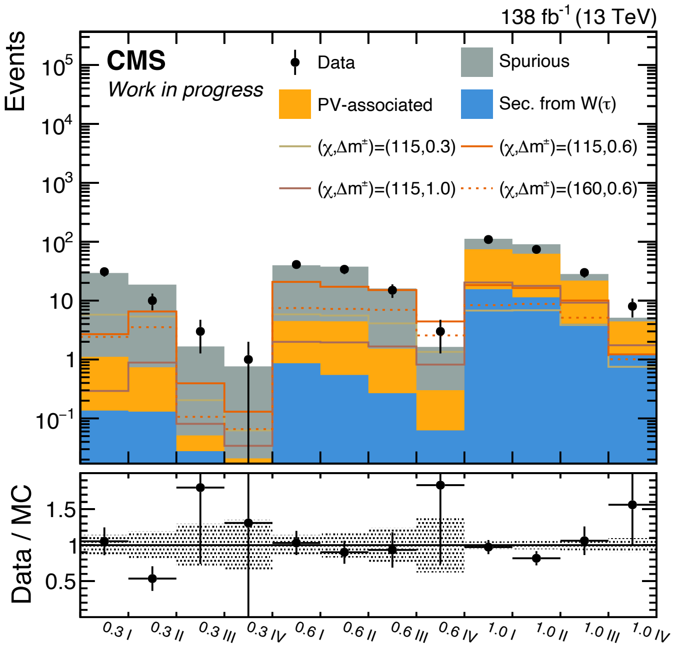

Figure 8.2: Expected yields of background and signal for four benchmark signal model points, as well as observed data yields, in all 12 signal regions. The systematic uncertainties are shown as shaded areas in the lower panels, the black error bars indicate the statistical uncertainty in the data. The MC distributions are shown pre-fit.

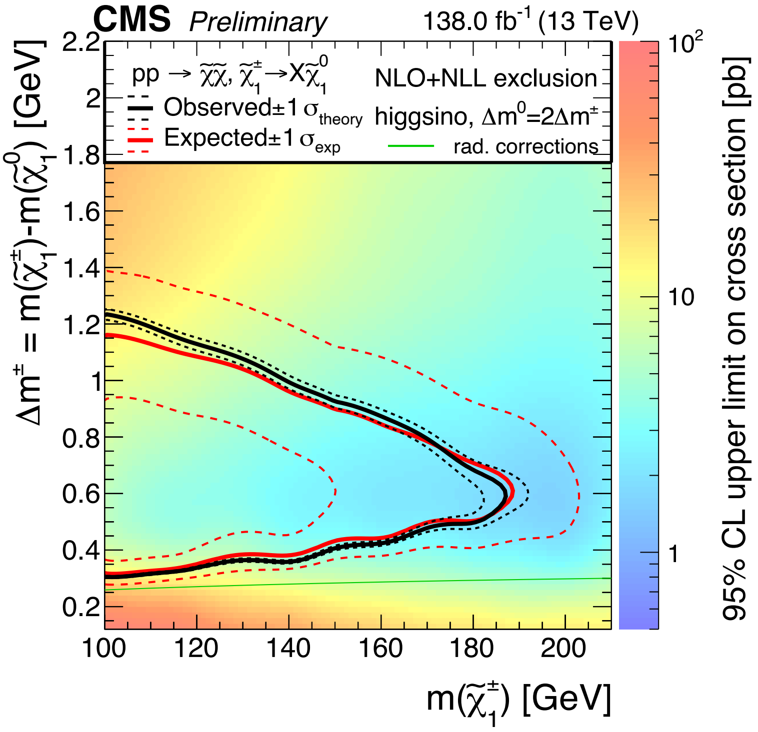

Figure 8.3: Exclusion limits in the plane of $\Delta m^{\pm}$ versus chargino mass $m(\tilde{\chi}_1^{\pm})$. The green line corresponds to the case of pure higgsinos with the mass splitting arising only due to radiative corrections from the SM particles.

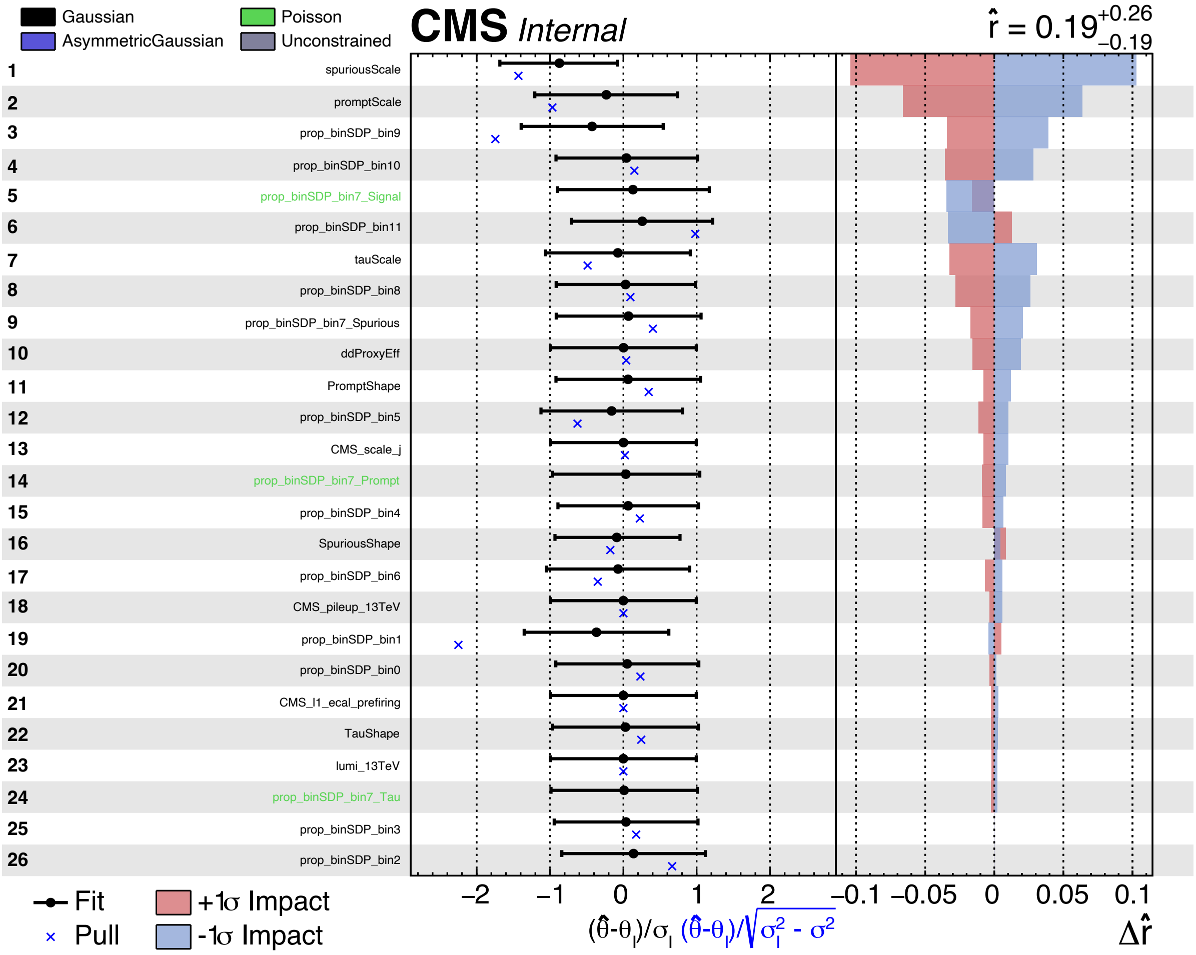

Figure 8.4: Impact on the nuisance parameters of the observed data set. The test model corresponds to $m(\tilde{\chi}_1^{\pm}) = 115$ GeV, $\Delta m(\tilde{\chi}_1^{\pm}, \tilde{\chi}_1^0) = 0.768$ GeV.

9 Summary and Outlook

A search has been presented for new physics manifesting as a minimal supersymmetric extension of the Standard Model featuring electroweakinos with compressed mass spectra. The considered signal models are characterized by the pair-production of electroweakinos, including one or two charginos. For $\Delta m^{\pm} = \Delta m(\widetilde{\chi}_1^{\pm}, \widetilde{\chi}_1^0) \lesssim 1\,\mathrm{GeV}$, the charginos decay to the lightest neutralino and, via an off-shell W boson, most often to a single pion. Due to the small mass difference between the chargino and the lightest neutralino, the chargino has a discernible decay length of up to $O(1\,\mathrm{cm})$. The chargino most often decays within the beam pipe, endowing the pion with a small but measurable impact parameter with respect to the primary interaction vertex.

This signature is experimentally probed by selecting events with large missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$ (owing to the final state neutralinos) and a low-momentum, displaced track (associated with the pion). To isolate signal-like tracks from the vast quantity of background tracks coming from the underlying event, pileup vertices, and erroneous reconstruction, a neural network classifier is trained. It uses various input observables including kinematic track variables, displacement- and isolation-related features, as well as $p_{\mathrm{T}}^{\mathrm{miss}}$ and the angle of the track to $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$. Furthermore, the network is parametrized by the mass splitting $\Delta m^{\pm}$, boosting sensitivity across the considered range of $\Delta m^{\pm}$, which spans significant variations in signal characteristics.

The soft track classifier is evaluated for all tracks in an event using three mass splittings ($\Delta m^{\pm} = 0.3\,\mathrm{GeV}$, $\Delta m^{\pm} = 0.6\,\mathrm{GeV}$, and $\Delta m^{\pm} = 1.0\,\mathrm{GeV}$) and the maximum classifier value considering all tracks and all values of $\Delta m^{\pm}$ is determined. Events are categorized based on which $\Delta m^{\pm}$ value corresponded to the highest scoring track, and signal regions are constructed by placing cuts on the maximum score, leading to four signal regions in each category and 12 overall.

For the background prediction in the signal regions, dedicated corrections are applied to the MC simulation, including a novel shape correction method using a refinement neural network. This regression model is trained in a cleaned Drell-Yan control region to provide a residual correction to the simulated samples such that their distribution-level agreement with the observed data is improved.

The leading systematic uncertainties are associated to the normalization of the background prediction and to the corresponding shape. To assess the validity of the signal simulation, tracks from τ lepton decays are used as a signal proxy.

The search excludes previously unexplored phase space regions of well-motivated natural SUSY scenarios with exclusion limits reaching up to $m(\widetilde{\chi}_1^{\pm}) = 190\,\mathrm{GeV}$ for $\Delta m^{\pm} = 0.5\,\mathrm{GeV}$. The absence of signal in the probed phase space does not rule out SUSY in general. However, it means that, if existent, SUSY must be more fine-tuned, weakening the argument for the theory as the solution to the hierarchy problem.

Remaking this analysis using the data collected during Run 3 of the LHC and during the future High-Luminosity phase will further enhance the sensitivity to the compressed electroweakino model space. Beyond the LHC, Ref. [76] studies the potential reach of a similar search using soft tracks conducted at a 3 TeV muon collider. The estimated exclusion region with chargino masses up to the TeV scale is especially interesting due to the fact that higgsino-like neutralinos in this range can fully explain the dark matter relic density.

A Appendix

$p_{\mathrm{T}}(\mathsf{Z}_{\mathsf{GEN}})$ Weights

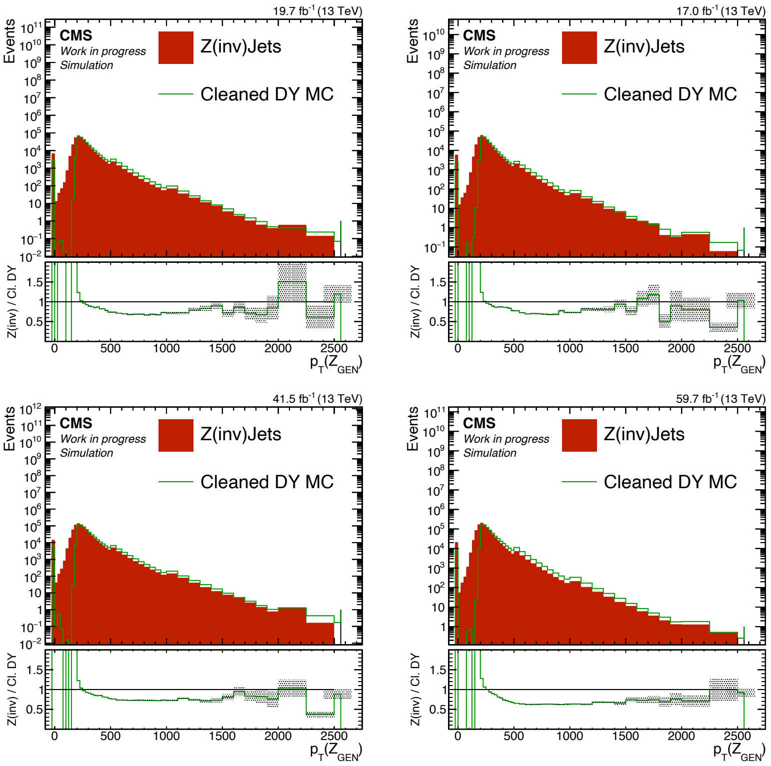

Figure A.1: Histograms used to derive the weights for the $p_{\mathrm{T}}(\mathrm{Z}_{\mathrm{GEN}})$ correction for cleaned DY MC.

Track-level variables used in the soft track classifier

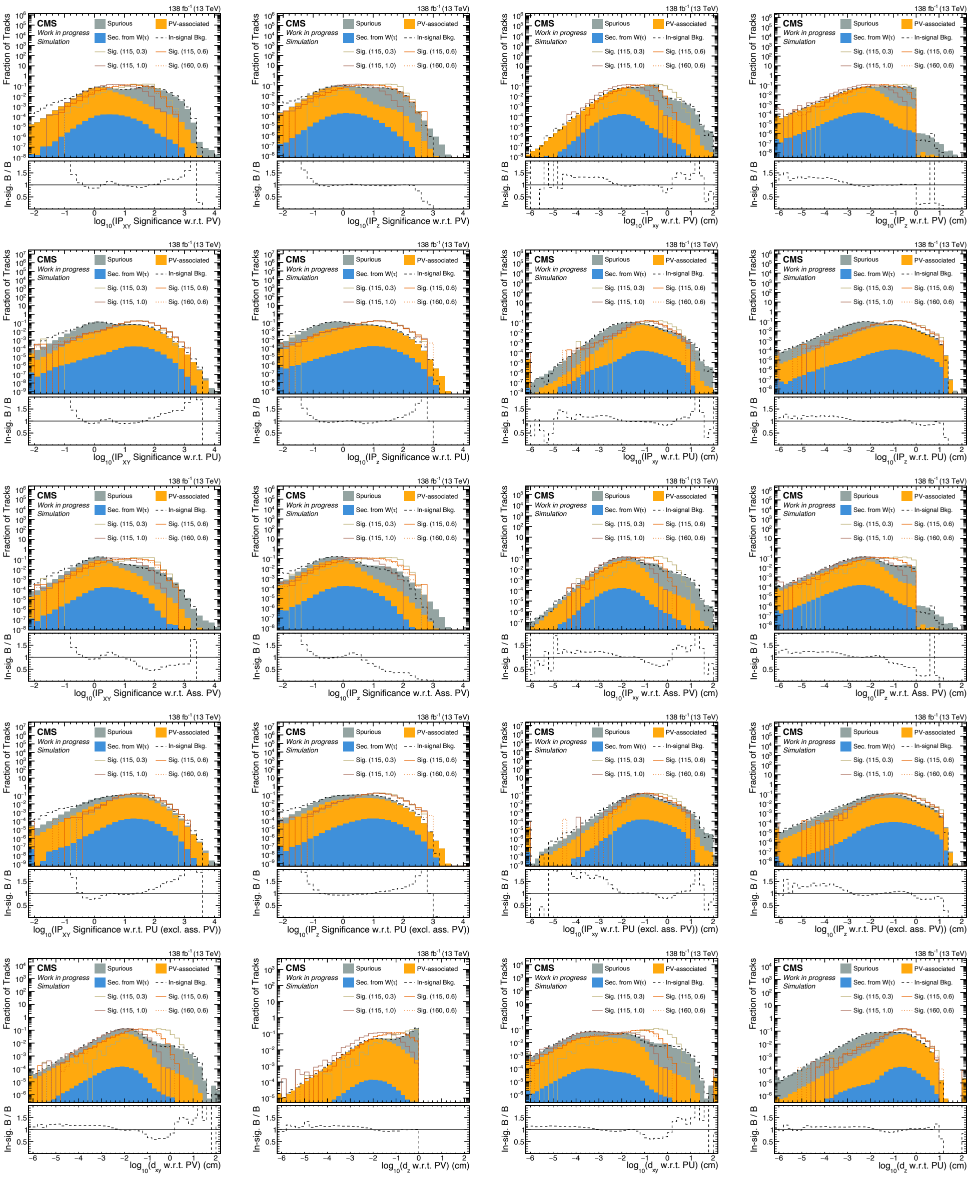

Figure A.2: Distributions of kinematic track observables used in the soft track classifier for background tracks and signal tracks from four exemplary signal model points. The filled histograms show distributions for background tracks in SM background events, whereas the dashed lines correspond to the distributions of background tracks within signal events, i.e. tracks that are not matched to (decay products of) the SUSY particles.

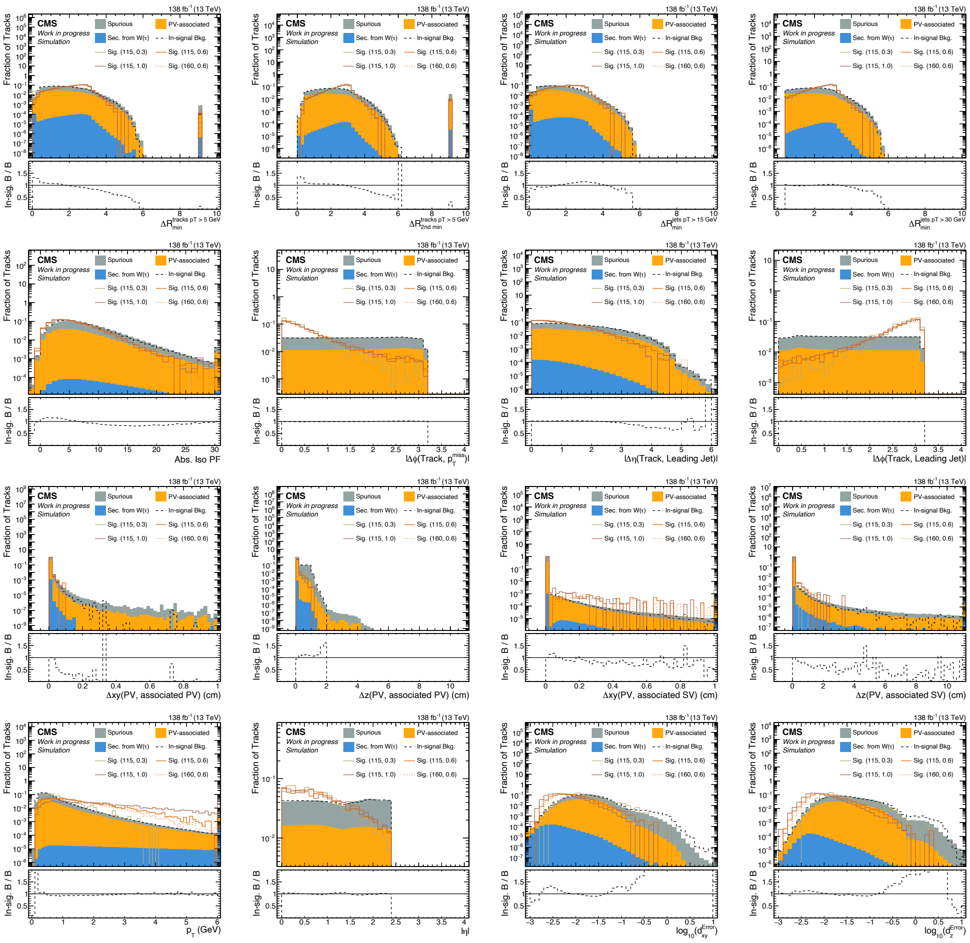

Figure A.3: Distributions of kinematic track observables used in the soft track classifier for background tracks and signal tracks from four exemplary signal model points. The filled histograms show distributions for background tracks in SM background events, whereas the dashed lines correspond to the distributions of background tracks within signal events, i.e. tracks that are not matched to (decay products of) the SUSY particles.

Data sets

Table A.1: Data sets and corresponding integrated luminosities used from the MET data stream. The AOD data tier is used.

Era	Processing string	$\mathcal{L}_{\text{int.}}$ (fb^{-1})
2016 preVFP	Run2016B-21Feb2020_ver1_UL2016_HIPM-v1	36.3
	Run2016B-21Feb2020_ver2_UL2016_HIPM-v1	
	Run2016C-21Feb2020_UL2016_HIPM-v1	
	Run2016D-21Feb2020_UL2016_HIPM-v1	
	Run2016E-21Feb2020_UL2016_HIPM-v1	
	Run2016F-21Feb2020_UL2016_HIPM-v1	
2016 postVFP	Run2016F-21Feb2020_UL2016-v1	
	Run2016G-21Feb2020_UL2016-v1	
	Run2016H-21Feb2020_UL2016-v2	
2017	Run2017B-09Aug2019_UL2017_rsb-v1	41.5
	Run2017C-09Aug2019_UL2017_rsb-v1	
	Run2017D-09Aug2019_UL2017_rsb-v1	
	Run2017E-09Aug2019_UL2017_rsb-v1	
	Run2017F-09Aug2019_UL2017_rsb-v1	
2018	Run2018A-15Feb2022_UL2018-v1	59.8
	Run2018B-15Feb2022_UL2018-v1	
	Run2018C-15Feb2022_UL2018-v1	
	Run2018D-15Feb2022_UL2018-v1	

Table A.2: Data sets and corresponding integrated luminosities used from the SingleMuon data stream. The AOD data tier is used.

Era	Processing string	$\mathcal{L}_{\text{int.}}$ (fb^{-1})
2016 preVFP	Run2016B-21Feb2020_ver1_UL2016_HIPM-v1	36.3
	Run2016B-21Feb2020_ver2_UL2016_HIPM-v1	
	Run2016C-21Feb2020_UL2016_HIPM-v1	
	Run2016D-21Feb2020_UL2016_HIPM-v1	
	Run2016E-21Feb2020_UL2016_HIPM-v1	
	Run2016F-21Feb2020_UL2016_HIPM-v1	
2016 postVFP	Run2016F-21Feb2020_UL2016-v1	
	Run2016G-21Feb2020_UL2016-v1	
	Run2016H-21Feb2020_UL2016-v1	
2017	Run2017B-15Feb2022_UL2017-v1	41.5
	Run2017C-15Feb2022_UL2017-v1	
	Run2017D-15Feb2022_UL2017-v1	
	Run2017E-15Feb2022_UL2017-v1	
	Run2017F-15Feb2022_UL2017-v1	
2018	Run2018A-15Feb2022_UL2018-v1	59.8
	Run2018B-15Feb2022_UL2018-v1	
	Run2018C-15Feb2022_UL2018-v1	
	Run2018D-15Feb2022_UL2018-v1	

Table A.3: Simulated SM background data sets and corresponding cross sections. The AODSIM data tier is used. The cross sections for ZJetsToNuNu* and DYJetsToLL* (WJetsToLNu*) include k-factors of 1.23 (1.21).

Primary data set name	σ (pb)
ZJetsToNuNu_HT-100To200_TuneCP5_13TeV-madgraphMLM-pythia8	325.1
ZJetsToNuNu_HT-200To400_TuneCP5_13TeV-madgraphMLM-pythia8	89.09
ZJetsToNuNu_HT-400To600_TuneCP5_13TeV-madgraphMLM-pythia8	12.21
ZJetsToNuNu_HT-600To800_TuneCP5_13TeV-madgraphMLM-pythia8	2.932
ZJetsToNuNu_HT-800To1200_TuneCP5_13TeV-madgraphMLM-pythia8	1.320
ZJetsToNuNu_HT-1200To2500_TuneCP5_13TeV-madgraphMLM-pythia8	0.3055
ZJetsToNuNu_HT-2500ToInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.0069
WJetsToLNu_HT-100To200_TuneCP5_13TeV-madgraphMLM-pythia8	1523
WJetsToLNu_HT-200To400_TuneCP5_13TeV-madgraphMLM-pythia8	405.2
WJetsToLNu_HT-400To600_TuneCP5_13TeV-madgraphMLM-pythia8	54.23
WJetsToLNu_HT-600To800_TuneCP5_13TeV-madgraphMLM-pythia8	12.92
WJetsToLNu_HT-800To1200_TuneCP5_13TeV-madgraphMLM-pythia8	6.088
WJetsToLNu_HT-1200To2500_TuneCP5_13TeV-madgraphMLM-pythia8	1.398
WJetsToLNu_HT-2500ToInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.0317
TTJets_DiLept_TuneCP5_13TeV-madgraphMLM-pythia8	53.05
TTJets_SingleLeptFromT_TuneCP5_13TeV-madgraphMLM-pythia8	105.9
TTJets_SingleLeptFromTbar_TuneCP5_13TeV-madgraphMLM-pythia8	105.0
ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	67.93
ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	113.4
ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	32.51
ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	32.45
WW_TuneCP5_13TeV-pythia8	76.17
WZ_TuneCP5_13TeV-pythia8	27.44
ZZ_TuneCP5_13TeV-pythia8	12.16
DYJetsToLL_M-50_HT-100to200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	171.2
DYJetsToLL_M-50_HT-200to400_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	47.23
DYJetsToLL_M-50_HT-400to600_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	6.364
DYJetsToLL_M-50_HT-600to800_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	1.547
DYJetsToLL_M-50_HT-800to1200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.6886
DYJetsToLL_M-50_HT-1200to2500_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.1605
DYJetsToLL_M-50_HT-2500toInf_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.0037

Table A.4: Processing strings for the simulated SM background data sets.

Era	Processing string
2016 preVFP	RunIISummer20UL16RECOAPV-106X_mcRun2_asymptotic_preVFP_v8*
2016 postVFP	RunIISummer20UL16RECO-106X_mcRun2_asymptotic_v13*
2017	RunIISummer20UL17RECO-106X_mc2017_realistic_v6*
2018	RunIISummer20UL18RECO-106X_upgrade2018_realistic_v11_L1v1*

Bibliography

[1] M. Wolf et al., “Fast Perfekt: Regression-based refinement of fast simulation”, *SciPost Phys. Core* **8** (2025) 021, `doi:10.21468/SciPostPhysCore.8.1.021`.

[2] CMS Collaboration, “Refining fast simulation using machine learning”, *EPJ Web Conf.* **295** (2024) 09032, `doi:10.1051/epjconf/202429509032`, `arXiv:2309.12919`.

[3] S. L. Glashow, “Partial Symmetries of Weak Interactions”, *Nucl. Phys.* **22** (1961) 579–588, `doi:10.1016/0029-5582(61)90469-2`.

[4] S. Weinberg, “A Model of Leptons”, *Phys. Rev. Lett.* **19** (1967) 1264–1266, `doi:10.1103/PhysRevLett.19.1264`.

[5] S. L. Glashow, J. Iliopoulos, and L. Maiani, “Weak Interactions with Lepton-Hadron Symmetry”, *Phys. Rev. D* **2** (1970) 1285–1292, `doi:10.1103/PhysRevD.2.1285`.

[6] M. Gell-Mann, “A Schematic Model of Baryons and Mesons”, *Phys. Lett.* **8** (1964) 214–215, `doi:10.1016/S0031-9163(64)92001-3`.

[7] G. Zweig, “An SU(3) model for strong interaction symmetry and its breaking. Version 1”, `doi:10.17181/CERN-TH-401`.

[8] 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. B* **716** (2012) 1–29, `doi:10.1016/j.physletb.2012.08.020`, `arXiv:1207.7214`.

[9] CMS Collaboration, “Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC”, *Phys. Lett. B* **716** (2012) 30–61, `doi:10.1016/j.physletb.2012.08.021`, `arXiv:1207.7235`.

[10] I. Neutelings. Licensed under a Creative Commons Attribution-ShareAlike 4.0 International License `https://creativecommons.org/licenses/by-sa/4.0/`. `https://tikz.net/`.

[11] Planck Collaboration, “Planck 2018 results. VI. Cosmological parameters”, *Astron. Astrophys.* **641** (2020) A6, `doi:10.1051/0004-6361/201833910`, `arXiv:1807.06209`. [Erratum: Astron.Astrophys. 652, C4 (2021)].

[12] S. P. Martin, “A Supersymmetry primer”, *Adv. Ser. Direct. High Energy Phys.* **18** (1998) 1–98, `doi:10.1142/9789812839657_0001`, `arXiv:hep-ph/9709356`.

[13] H. Baer et al., “Radiative natural SUSY with a 125 GeV Higgs boson”, *Phys. Rev. Lett.* **109** (2012) 161802, `doi:10.1103/PhysRevLett.109.161802`, `arXiv:1207.3343`.

[14] E. Lopienska, “The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022”,. General Photo.

[15] CMS Collaboration, “Public CMS Luminosity Information”. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`.

[16] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004, `doi:10.1088/1748-0221/3/08/S08004`.

[17] T. Sakuma and T. McCauley, “Detector and Event Visualization with SketchUp at the CMS Experiment”, *J. Phys. Conf. Ser.* **513** (2014) 022032, `doi:10.1088/1742-6596/513/2/022032`, `arXiv:1311.4942`.

[18] CMS Tracker Group Collaboration, “The CMS Phase-1 Pixel Detector Upgrade”, *JINST* **16** (2021), no. 02, P02027, `doi:10.1088/1748-0221/16/02/P02027`, `arXiv:2012.14304`.

[19] CMS Collaboration, “Track impact parameter resolution for the full pseudo rapidity coverage in the 2017 dataset with the CMS Phase-1 Pixel detector”,.

[20] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST* **9** (2014), no. 10, P10009, `doi:10.1088/1748-0221/9/10/P10009`, `arXiv:1405.6569`.

[21] CMS Collaboration, “Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JINST* **15** (2020), no. 10, P10017, `doi:10.1088/1748-0221/15/10/P10017`, `arXiv:2006.10165`.

[22] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector”, *JINST* **12** (2017), no. 10, P10003, `doi:10.1088/1748-0221/12/10/P10003`, `arXiv:1706.04965`.

[23] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-k_t jet clustering algorithm”, *JHEP* **04** (2008) 063, `doi:10.1088/1126-6708/2008/04/063`, `arXiv:0802.1189`.

[24] J. Alwall et al., “MadGraph 5: going beyond”, *JHEP* **06** (2011) 128, `doi:10.1007/JHEP06(2011)128`, `arXiv:1106.0522`.

[25] T. Sjöstrand, S. Mrenna, and P. Skands, "A brief introduction to PYTHIA 8.1", *Comp. Phys. Comm.* **178** (2008) 852, `doi:10.1016/j.cpc.2008.01.036`, `arXiv:0710.3820`.

[26] T. Gleisberg et al., "Event generation with SHERPA 1.1", *JHEP* **02** (2009) 007, `doi:10.1088/1126-6708/2009/02/007`, `arXiv:0811.4622`.

[27] S. Agostinelli et al., "Geant4—a simulation toolkit", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250–303, `doi:https://doi.org/10.1016/S0168-9002(03)01368-8`.

[28] J. Allison et al., "Recent developments in geant4", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **835** (2016) 186–225, `doi:https://doi.org/10.1016/j.nima.2016.06.125`.

[29] J. Allison et al., "Geant4 developments and applications", *IEEE Transactions on Nuclear Science* **53** (2006), no. 1, 270–278, `doi:10.1109/TNS.2006.869826`.

[30] S. Abdullin et al., "The fast simulation of the CMS experiment", *Journal of Physics: Conference Series* **331** (2011) 032049, `doi:10.1088/1742-6596/331/3/032049`.

[31] A. Giammanco, "The Fast Simulation of the CMS Experiment", *Journal of Physics: Conference Series* **513** (2014) 022012, `doi:10.1088/1742-6596/513/2/022012`.

[32] G. Grindhammer, M. Rudowicz, and S. Peters, "The Fast Simulation of Electromagnetic and Hadronic Showers", *Nucl. Instrum. Meth. A* **290** (1990) 469, `doi:10.1016/0168-9002(90)90566-O`.

[33] N. Krammer. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/CMSOfflineComputingResults`.

[34] C. O. Software and Computing, "CMS Phase-2 Computing Model: Update Document", technical report, CERN, Geneva, 2022.

[35] A. Andreassen and B. Nachman, "Neural Networks for Full Phase-space Reweighting and Parameter Tuning", *Phys. Rev. D* **101** (2020), no. 9, 091901, `doi:10.1103/PhysRevD.101.091901`, `arXiv:1907.08209`.

[36] M. Erdmann, L. Geiger, J. Glombitza, and D. Schmidt, "Generating and refining particle detector simulations using the Wasserstein distance in adversarial networks", *Comput. Softw. Big Sci.* **2** (2018), no. 1, 4, `doi:10.1007/s41781-018-0008-x`, `arXiv:1802.03325`.

[37] S. Diefenbacher, V. Mikuni, and B. Nachman, “Refining Fast Calorimeter Simulations with a Schrödinger Bridge”, `arXiv:2308.12339`.

[38] S. Bright-Thonney, P. Harris, P. McCormack, and S. Rothman, “Chained Quantile Morphing with Normalizing Flows”, `arXiv:2309.15912`.

[39] T. Golling et al., “Morphing one dataset into another with maximum likelihood estimation”, *Phys. Rev. D* **108** (2023), no. 9, 096018, `doi:10.1103/PhysRevD.108.096018`, `arXiv:2309.06472`.

[40] O. Amram et al., “CaloChallenge 2022: A Community Challenge for Fast Calorimeter Simulation”, `arXiv:2410.21611`.

[41] H. Zhang, Y. N. Dauphin, and T. Ma, “Residual learning without normalization via better initialization”, in *International Conference on Learning Representations*. 2019.

[42] K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: Surpassing human-level performance on imagenet classification”, in *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1026–1034. 2015. `doi:10.1109/ICCV.2015.123`.

[43] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition”, `doi:10.1109/CVPR.2016.90`, `arXiv:1512.03385`.

[44] K. He, X. Zhang, S. Ren, and J. Sun, “Identity Mappings in Deep Residual Networks”, `arXiv:1603.05027`.

[45] A. Gretton et al., “A kernel two-sample test”, *Journal of Machine Learning Research* **13** (2012), no. 25, 723–773.

[46] E. Pouranbarani, R. W. dos Santos, and A. Nygren, “The basic concept of the MOP solution on a convex Pareto front.”, `doi:10.1371/journal.pone.0225245.g005`.

[47] J. Degrave and I. Korshunova, “Why machine learning algorithms are hard to tune and how to fix it”. `https://www.engraved.blog/why-machine-learning-algorithms-are-hard-to-tune/`. Accessed: February 5, 2025.

[48] J. Platt and A. Barr, “Constrained differential optimization”, in *Neural Information Processing Systems*, D. Anderson, ed., volume 0. American Institute of Physics, 1987.

[49] H. W. Kuhn and A. W. Tucker, “Nonlinear programming”, in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*, pp. 481–492. Univ. California Press, Berkeley-Los Angeles, Calif., 1951.

[50] W. Karush, “Minima of functions of several variables with inequalities as side conditions.”. PhD thesis, Thesis (S.M.)–University of Chicago, Department of Mathematics, December 1939., 1939.

[51] LHC New Physics Working Group Collaboration, “Simplified Models for LHC New Physics Searches”, *J. Phys. G* **39** (2012) 105005, `doi:10.1088/0954-3899/39/10/105005`, `arXiv:1105.2838`.

[52] E. Bols et al., “Jet flavour classification using deepjet”, *JINST* **15** (dec, 2020) P12012, `doi:10.1088/1748-0221/15/12/P12012`.

[53] DELPHES 3 Collaboration, “DELPHES 3, A modular framework for fast simulation of a generic collider experiment”, *JHEP* **02** (2014) 057, `doi:10.1007/JHEP02(2014)057`, `arXiv:1307.6346`.

[54] J. Thaler and K. Van Tilburg, “Identifying boosted objects with n-subjettiness”, *JHEP* **03** (2011) 015, `doi:10.1007/JHEP03(2011)015`, `arXiv:1011.2268`.

[55] I. Tolstikhin, O. Bousquet, S. Gelly, and B. Schoelkopf, “Wasserstein auto-encoders”, in *International Conference on Learning Representations*. 2018.

[56] N. Nagata and S. Shirai, “Higgsino Dark Matter in High-Scale Supersymmetry”, *JHEP* **01** (2015) 029, `doi:10.1007/JHEP01(2015)029`, `arXiv:1410.4549`.

[57] M. Ibe, Y. Nakayama, and S. Shirai, “Precise estimate of charged Higgsino/Wino decay rate”, *JHEP* **03** (2024) 012, `doi:10.1007/JHEP03(2024)012`, `arXiv:2312.08087`.

[58] Y. Nissan, “Search for compressed mass Higgsino production with low-momentum lepton tracks with the CMS experiment”. PhD thesis, University of Hamburg, 2023.

[59] V. G. Kutzner, “Search for exotic long-lived particles using disappearing tracks with the CMS experiment in proton-proton collisions at sqrt(s) = 13 TeV”. PhD thesis, University of Hamburg, 2023.

[60] CMS Collaboration, “Search for supersymmetry in final states with disappearing tracks in proton-proton collisions at s=13 TeV”, *Phys. Rev. D* **109** (2024), no. 7, 072007, `doi:10.1103/PhysRevD.109.072007`, `arXiv:2309.16823`.

[61] Yuval Nissan and Samuel Bein and Peter Schleper and C. Collaboration, "Search for Higgsinos in final states with low-momentum lepton-track pairs", CMS Analysis Note AN-21-130, CMS Collaboration, 2021. CMS Internal Analysis note.

[62] ATLAS Collaboration, "Search for nearly mass-degenerate higgsinos using low-momentum mildly-displaced tracks in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Rev. Lett.* **132** (2024), no. 22, 221801, `doi:10.1103/PhysRevLett.132.221801`, `arXiv:2401.14046`.

[63] H. Fukuda et al., "Cornering Higgsinos Using Soft Displaced Tracks", *Phys. Rev. Lett.* **124** (2020), no. 10, 101801, `doi:10.1103/PhysRevLett.124.101801`, `arXiv:1910.08065`.

[64] Viktor Kutzner and Samuel Bein and Peter Schleper and CMS Collaboration, "Search for long lived SUSY with disappearing tracks in inclusive final states", CMS Analysis Note AN-18-214, CMS Collaboration, 2018. CMS Internal Analysis note.

[65] T. Sjöstrand et al., "An Introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, `doi:10.1016/j.cpc.2015.01.024`, `arXiv:1410.3012`.

[66] W. Beenakker et al., "Production of Charginos, Neutralinos, and Sleptons at Hadron Colliders", *Phys. Rev. Lett.* **83** (1999) 3780, `doi:10.1103/PhysRevLett.83.3780`, `arXiv:hep-ph/9906298`. [Erratum: *P*hys. Rev. Lett. **1**00 (2008) 029901, doi:10.1103/PhysRevLett.100.029901].

[67] 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, `doi:10.1007/JHEP10(2012)081`, `arXiv:1207.2159`.

[68] 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, `doi:10.1140/epjc/s10052-013-2480-0`, `arXiv:1304.0790`.

[69] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018), no. 05, P05011, `doi:10.1088/1748-0221/13/05/P05011`, `arXiv:1712.07158`.

[70] CMS Collaboration, "Performance summary of AK4 jet b tagging with data from proton-proton collisions at 13 TeV with the CMS detector",.

[71] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions", in *Advances in Neural Information Processing Systems*, I. Guyon et al., eds., volume 30. Curran Associates, Inc., 2017.

[72] CMS Collaboration, “Search for new particles in events with energetic jets and large missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JHEP* **11** (2021) 153, `doi:10.1007/JHEP11(2021)153`, `arXiv:2107.13021`.

[73] CMS Collaboration, “Luminosity recommendations for Run 2 analyses”. `https://twiki.cern.ch/twiki/bin/view/CMS/LumiRecommendationsRun2`.

[74] G. J. Feldman and R. D. Cousins, “A Unified approach to the classical statistical analysis of small signals”, *Phys. Rev. D* **57** (1998) 3873–3889, `doi:10.1103/PhysRevD.57.3873`, `arXiv:physics/9711021`.

[75] CMS Collaboration, “The CMS Statistical Analysis and Combination Tool: Combine”, *Comput. Softw. Big Sci.* **8** (2024), no. 1, 19, `doi:10.1007/s41781-024-00121-4`, `arXiv:2404.06614`.

[76] R. Capdevilla, F. Meloni, and J. Zurita, “Discovering Electroweak Interacting Dark Matter at Muon Colliders using Soft Tracks”, `arXiv:2405.08858`.

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben.

Sofern im Zuge der Erstellung der vorliegenden Dissertationsschrift generative Künstliche Intelligenz (gKI) basierte elektronische Hilfsmittel verwendet wurden, versichere ich, dass meine eigene Leistung im Vordergrund stand und dass eine vollständige Dokumentation aller verwendeten Hilfsmittel gemäß der Guten wissenschaftlichen Praxis vorliegt. Ich trage die Verantwortung für eventuell durch die gKI generierte fehlerhafte oder verzerrte Inhalte, fehlerhafte Referenzen, Verstöße gegen das Datenschutz- und Urheberrecht oder Plagiate.

Hamburg, 09.10.2025	M. Wolf
Ort, Datum	Unterschrift des Doktoranden

Verwendete Hilfsmittel Die Arbeit wurde verfasst mithilfe des LaTeX-Editors *Overleaf* und der dort enthaltenen Rechtschreibkorrektur und Synonymsuche (*Writefull* Version 2025.14.0).