Explaining Data Anomalies over the NMSSM Parameter Space with Deep Learning Techniques

A. Hammad[1], Raymundo Ramos[2],
Amit Chakraborty[3], Pyungwon Ko[4], and Stefano Moretti[5,6]

[1]Theory Center, IPNS, KEK, 1-1 Oho, Tsukuba, Ibaraki 305-0801, Japan
[2]Quantum Universe Center, KIAS, Seoul 02455, Korea
[3]Department of Physics, SRM University-AP, Amaravati, Mangalagiri 522240, India
[4]School of Physics, KIAS, Seoul 02455, Korea
[5]School of Physics & Astronomy, University of Southampton, Southampton SO17 1BJ, UK
[6]Department of Physics & Astronomy, Uppsala University, Box 516, 75120 Uppsala, Sweden

Abstract

Motivated by recent results from particle physics analyses, we investigate the Next-to-Minimal Supersymmetric Standard Model (NMSSM) as a framework capable of accommodating a range of current data anomalies across low- and high-energy experiments. These include the so-called 95 GeV and 650 GeV excesses from Higgs studies, the Electro-Weakino excess from Supersymmetry searches, the latest $(g - 2)_\mu$ measurements as well as potential deviations from Standard Model (SM) predictions that would appear as a consequence in mono-H (where $H = h_{\mathrm{SM}}$) and -Z signatures of Dark Matter. Our analysis demonstrates that viable NMSSM parameter regions exist where all these features can be accommodated at the 2σ level while remaining consistent with the most up-to-date theoretical and experimental constraints. To identify such regions, we employ an efficient numerical scanning strategy assisted by Deep Learning techniques. We further present several Benchmark Points that realize these scenarios, offering promising directions for future phenomenological studies.

1

Contents

1 Introduction

Leaving aside much experimental evidence pointing to the fact that the Standard Model (SM) cannot be a *complete* description of Nature (e.g., the existence of neutrino masses, Dark Matter (DM) and the matter-antimatter asymmetry of the Universe), from a mere theoretical point of view, it is also clear that the SM cannot be a realization of the *ultimate* theory that is stable to high energies either (e.g., up to the Planck scale, $M_{\mathrm{Planck}} \sim 1.2 \times 10^{19}$ GeV, where gravity is expected to become strong and possibly unify with the other interactions). In fact, there are theoretical inconsistencies leading to instabilities of the Higgs boson self-interactions (or, equivalently, of the associate mass) that systematically grow with the energy at which the SM is being probed, ultimately leading to either its breakdown or a solution perceived as highly ad-hoc, hence unnatural. This conceptual problem is known as the 'hierarchy problem'. It is associated with the absence of a symmetry

protecting the Higgs mass, which is necessarily placed at the Electro-Weak (EW) scale (as now confirmed by experiment), when the interactions of the Higgs field are instead generated where the natural cutoff scale is or indeed above it, i.e., the energy at which a Grand Unification Theory (GUT) (also including gravity) will inevitably have to be formulated, $M_{\mathrm{GUT}} \simeq 3 \times 10^{16}$ GeV.

Thus, neither experimentally, nor theoretically, the SM can be considered a *fundamental* theory of Nature and we need to search for new physics constructions beyond it, which can nonetheless reproduce the SM in the phenomenological situations where it has been probed to a phenomenal degree of accuracy. One of the best candidates for Beyond the SM (BSM) physics is Supersymmetry (SUSY), which is nothing but the most general symmetry of the Lorentz-Einstein theory which relates bosons and fermions, hence implying a new kind of unification between particles of different spin.

Despite the absence of experimental verifications of such a theory, relevant theoretical arguments can be given in favour of it, i.e., unlike the SM: (i) the SUSY Higgs potential is generated dynamically; (ii) SUSY predicts the existence of a Higgs boson at the EW scale, like the one discovered in 2012; (iii) SUSY facilitates the convergence of the three gauge coupling 'constants' at high scale; (iv) the Lightest Supersymmetric Particle (LSP) can be a natural candidate for DM, typically the lightest so-called neutralino.

Thus SUSY is an elegant theory and, like the SM, needs not be minimal, neither in its gauge and Higgs content nor in its matter spectrum [1]. Indeed, one of the most attractive non-minimal SUSY models is the Next-to-Minimal Supersymmetric SM (NMSSM) (see Refs. [2, 3] for reviews), which extends the time-honoured MSSM by the introduction of just one singlet Superfield, $\hat{S}$, in addition to the two doublets of the MSSM, $\hat{H}_u$ and $\hat{H}_d$. When the scalar component of $\hat{S}$ acquires a TeV scale Vacuum Expectation Value (VEV), the Superpotential term $\hat{S}\hat{H}_u\hat{H}_d$ generates an 'effective' so-called μ term, i.e., $\lambda <S> \hat{H}_u\hat{H}_d$, in the form of an interaction for the Higgs doublet fields. Such a term is essential for acceptable phenomenology and is put in by hand in the MSSM (in turn giving rise to the μ problem of the latter). Furthermore, the presence of the additional Higgs state and its VEV also means that the tree-level mass of the aforementioned discovered Higgs boson is (in the NMSSM) much larger than in the MSSM, so that it is more natural to achieve a value for it in the vicinity of 125 GeV, without fine-tuning the loop contributions to it due (primarily) to the heavy top squarks (or stops), the notorious little (or small) hierarchy problem of the MSSM. Thus, the phenomenological implications of the NMSSM ought to be considered very seriously.

One aspect of all this is the fact that the Higgs sector of the MSSM is extended, so that in the NMSSM there are three CP-even Higgs bosons ($H_{1,2,3}$, $m_{H_1} < m_{H_2} < m_{H_3}$), two CP-odd Higgs bosons ($A_{1,2}$, $m_{A_1} < m_{A_2}$) and a charged Higgs pair ($H^\pm$)[1]. Another is that the SUSY counterpart of the additional singlet is an additional neutralino, altering the DM phenomenology of the model with respect to the MSSM. Then, there is the vast number and the sheer variety of sparticles entering at one-loop virtually any physics process. An important question is then whether such a model could be behind a long list of data anomalies currently present in various experimental datasets, especially considering that they have emerged in the context of searches for additional Higgs states (beyond the 125 GeV one), sparticle searches (in particular, non-colored ones) as well as precision measurements of the anomalous magnetic moment of the muon ($(g-2)_\mu$). It is entirely possible to make this assessment with a high level of precision, given the sophisticated tools available to make predictions in the context of the NMSSM, ranging from programs computing production cross sections and decay rates [4–6], to protocols interfacing generated NMSSM events to

---

[1]We assume that CP is not violated in the Higgs sector.

parton shower, hadronization, heavy flavor decay as well as detector software [7]. Finally, structured collaborations across theory and experiment exist to define NMSSM benchmark points (BPs) which can be tested at present and future facilities, starting from Ref. [8] and ending with the most recent efforts within, e.g., the NMSSM Subgroup of the Large Hadron Collider Higgs Cross-Section Working Group 3 (LHC-HXSWG3)[2].

Specifically, the data anomalies for which we seek an NMSSM explanation are as follows. CMS has observed mild excesses in the diphoton channel near 95.4 GeV with about 2σ significance, while ATLAS reported a smaller 1.7σ excess at 95 GeV [9–12]. Another resonant excess decaying into SM like Higgs and a scalar of mass 95 GeV has been observed in $\gamma\gamma\bar{b}b$ [13], $\tau^+\tau^-\bar{b}b$ [14] and $\gamma\gamma\tau^+\tau^-$ [15] final states. The phenomenology of these scalars has received some attention recently [16–19]. Another notable anomaly arises from a mild excess of events reported by both ATLAS [20, 21] and CMS [22, 23] over the SM expectations: these events feature final states with multiple leptons and large Missing Transverse Energy (MET or $\not{E}_T$). In particular, they are consistent with the so-called Electro-Weakino (EWino) excess, i.e., chargino–neutralino pair production, where the subsequent decays of the chargino and neutralino proceed via off-shell $W^\pm$ and Z bosons, resulting in soft leptons. Lastly, we take into account the recent measurement highlighting the discrepancy between the experimentally measured muon anomalous magnetic moment [24] and the SM prediction. We aim at explaining all such observations in the context of the NMSSM, in turn predicting anomalies in DM signals (specifically, mono-H and -Z ones, with H and Z being the Higgs and neutral weak boson of the SM, respectively), that will therefore act as collateral predictions of our analysis.

Moreover, our numerical study starts with a scan of the parameter space of the NMSSM that employs the recently released package `DLScanner` [25], a tool designed to perform scans of parameter spaces in high energy physics models. It uses deep learning (DL) techniques to identify regions in the parameter space where a target (conditions, constraints, extrema or other) is likely reached. It is an appropriate tool for this case as it can be tuned to find target regions that are complicated or small. One characteristic of `DLScanner` that helps with small targets is the use of `VEGAS` [26] maps that can be trained to generate random distributions that are more weighted towards the target. Considering that `DLScanner` is relatively new, this study serves as both a test and a demonstration of its capabilities.

Our paper is organized as follows. In Sect. 2 we introduce the NMSSM. In the two following ones, we list the datasets that will be used to test such a theoretical model, covering signatures emerging in both the visible (Higgs, gauge and matter) and invisible (SM) sectors of the NMSSM. Then, we will present and discuss our results in Sect. 6. Finally, we will conclude.

2 The NMSSM

The NMSSM [27–31] is an extension of the Minimal Supersymmetric Standard Model (MSSM) that introduces an additional gauge singlet Superfield. (See also Refs. [32, 33] for alternative SUSY realizations with the same particle content.) This addition provides a solution to the μ-problem of the MSSM and can naturally explain the mass of the 125 GeV Higgs boson without the need for excessively heavy stop quarks. Primarily, the NMSSM offers a broader range of possibilities for DM candidates and collider signals due to its extended neutralino and Higgs sectors, respectively, as discussed below.

The NMSSM includes a single-gauge chiral Superfield $\hat{S}$ in addition to the usual two MSSM-type Higgs doublets $\hat{H}_u$ and $\hat{H}_d$, giving rise to additional Higgs and neutralino states

---

[2]See `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHWGNMSSM`.

compared to their MSSM counterparts. The Z_3 invariant NMSSM Superpotential takes the following form:

$$W_{\text{NMSSM}} = W_{\text{MSSM}} + \lambda \hat{H}_u \hat{H}_d \hat{S} + \frac{1}{3}\kappa \hat{S}^3, \tag{1}$$

where W_{MSSM} is the MSSM Superpotential with the μ parameter set to zero. Once the scalar component of the Superfield $\hat{S}$ develops a VEV, $s \equiv \langle S \rangle$, the second term in the above equation provides the effective μ term: $\mu_{\text{eff}} = \lambda s$. Also notice that, without the κ term, the Superpotential above would have a $U(1)'$ symmetry, so-called Peccei–Quinn symmetry [34, 35]. The role of the κ term is to break this $U(1)'$ symmetry. Further, this term is introduced such that κ is dimensionless. However, there remains a discrete Z_3 symmetry, which can finally be broken spontaneously without cosmological consequences (in terms of domain walls) [36, 37].

The soft SUSY breaking part of the Lagrangian is given by

$$V_{\text{soft}} = m_{H_u}^2 |H_u|^2 + m_{H_d}^2 |H_d|^2 + m_S^2 |S|^2 + \lambda A_\lambda \hat{H}_u \hat{H}_d \hat{S} + \frac{1}{3}\kappa A_\kappa \hat{S}^3, \tag{2}$$

where A_λ and A_κ are free parameters that take positive and negative values. After EW Symmetry Breaking (EWSB) in a CP-conserving scenario, the Higgs fields acquire non-zero VEVs resulting in three CP-even neutral Higgs bosons $H_i, i = 1, 2, 3$, with $m_{H_1} < m_{H_2} < m_{H_3}$, two CP-odd Higgs bosons $A_i, i = 1, 2$, with $m_{A_1} < m_{A_2}$ and one pair of charged Higgs bosons ($H^\pm$). The presence of additional terms in the NMSSM Superpotential helps in controlling the masses of the Higgses while keeping the free parameters well within the perturbative limits. At tree level, the NMSSM Higgs sector is described by the six independent parameters: λ, κ, $\tan\beta$, A_λ, A_κ and μ_{eff}. Depending on the mixing between the double and singlet Higgs fields, one can easily fit the observed SM-like Higgs boson of mass at 125 GeV with one of the CP-even neutral Higgses of this model.

The neutralino sector of the NMSSM is a much richer and more complex version of its MSSM counterpart. While the MSSM has four neutralino states, the presence of an additional gauge singlet Superfield $\hat{S}$ introduces a new fermionic state called the singlino. This results in a mixing between the spin-1/2 partners of gauge bosons, namely bino ($\tilde{B}$), the charge-neutral wino ($\tilde{W}^0$), the charge-neutral Higgsinos ($\tilde{H}_u^0$ and $\tilde{H}_d^0$) and the singlino ($\tilde{S}^0$), in turn providing five neutralinos $\tilde{\chi}_i^0$, $i = 1...5$, with $m_{\tilde{\chi}_1^0} < m_{\tilde{\chi}_2^0} < ... < m_{\tilde{\chi}_5^0}$. Therefore, the physical neutralino mass eigenstates are linear combinations of the bino, wino, higgsino and singlino states, whose eigen values (at tree level) are determined by six free parameters: M_1, M_2, λ, κ, $\tan\beta$, and μ_{eff}, where $M_{1,2}$ are the bino and wino mass parameters. Note that the charged components of higgsinos and wino also mix among themselves, forming two physical charginos states $\tilde{\chi}_i^\pm$, $i = 1, 2$, with $m_{\tilde{\chi}_1^\pm} < m_{\tilde{\chi}_2^\pm}$.

The presence of a singlino significantly alters the phenomenology of the neutralino sector. For instance, the lightest neutralino can be a prime DM candidate having a dominant singlino component and thereby satisfying experimental constraints on DM relic density and (in)direct detection. The couplings of these singlino-like neutralino to SM particles are often suppressed, allowing it to also evade detection at colliders easily while leading to a wider range of possibile signatures for both collider searches and DM experiments in the years to come.

3 Parameter Space Scan and Constraints

In this section, we begin by presenting a detailed account of the computational tools and frameworks employed to systematically explore the NMSSM parameter space, emphasizing

the methodology adopted in conducting the scans. Following this, we provide a comprehensive discussion of the various constraints applied throughout the analysis, addressing in turn the theoretical requirements that ensure consistency of the model, as well as the experimental bounds derived from current collider data and other relevant measurements.

3.1 Setup for DL Scanning of the NMSSM Parameter Space

We start this subsection by describing the method used to test relevant constraints and obtain suitable regions of parameters. The scan over the parameter space of the NMSSM utilizes the the collection of tools packaged into `NMSSMTools`. This package emcompasses the programs `NMHDECAY` [4, 5], `NMSPEC` [6], `NMGMSB` [38], `NMHDECAY_CPV` [39], `NMSDECAY` [40], `micrOMEGAs` [41] and `SModelS` [42]. For faster calculation of an array of several parameters of the NMSSM, we have created a custom interface in Python that takes care of running different instances of `NMSSMTools` in multiple CPUs and read their outputs into simple lists of parameter vectors. This is particularly useful considering that our scan is performed using the DL parameter scanning tool `DLScanner` [25], which has been implemented as a Python module. This scanning tool uses Deep Neural Networks (DNNs) to learn and guess parameter regions where a series of constraints may be met. It does so by using either regression into a function that quantifies the distance to the best fit point or classification of regions according to whether or not they follow all the constraints. Moreover, it follows an iterative predict-learn loop where after results are predicted for new points, then the corrected results are used to keep training the network. Additionally, it employs `VEGAS` [26] mappings to improve sampling of the parameter space even when trying to identify comparably small regions inside a much larger space.

In this work, we have performed a scan over 12 parameters of the NMSSM, defined at the EW scale, namely, $\tan\beta$, λ, κ, A_λ, A_κ, μ_{eff}, M_1, M_2, M_3, A_t, M_{Q_3} and M_{U_3} (see Ref. [2] for their definition.). We start scanning over a large space in order to identify likely regions before reducing the space. The wide and narrow ranges used in the preliminary and the final scans, respectively, are given in Table 1. The wide ranges of the preliminary scan are consistent with taking maximum upper values and minimum lower values from the ranges scanned in Refs. [17] and [18], except for M_2, M_{Q_3} and M_{U_3}, which where allowed to run up to 10^4 GeV.

	$\tan\beta$	λ	κ	A_λ
wide	[1.97, 10.9]	[0.013, 0.687]	[0.0058, 0.391]	[−5000, 480]
narrow	[3.2, 6.2]	[0.07, 0.42]	[0.05, 0.3]	[351, 834]

	A_κ	μ_{eff}	M_1	M_2
wide	[−621, 362]	[−244, 291]	[178, 3000]	[304, 10000]
narrow	[−300, −150]	[120, 220]	[500, 3000]	[750, 10000]

M_3	A_t	M_{Q_3}	M_{U_3}
[423, 5000]	[−5000, 1288]	[272, 10000]	[570, 10000]

Table 1: Ranges scanned for the 12 parameters considered in this study. The word 'wide' stands for the ranges used in the preliminary scan while the word 'narrow' is used for the scan over the reduced parameter space. Parameters that do not indicate wide or narrow ranges receive the least effects from constraints and are therefore kept in the same range. Dimensionful parameters are given in GeV.

The scanning is done using the classification method in `DLScanner`. By definition,

classification requires a function that returns a binary output that distinguishes between two classes: in our case whether the points meets the conditions or not. However, when testing many constraints, with wide parameter ranges and scanning many parameters, it is possible that the target region is several orders of magnitude smaller than the scanned space. Therefore, instead of an approach purely based on classification, we follow an hybrid approach based on iteratively sampling downwards a penalty function. To be more specific, we define a penalty function based on the constraints: if the constraints are not met for enough points, we pick a large number of points with the smallest penalty function and define them as the "1" class and the rest as the "0" class. This, combined with the `VEGAS` map attempting to sample the "1" class, progressively reduces the minimum value found for the penalty function. The type of penalty function we use is defined for each constraint and is based on having 0 penalty when the set of parameters gives a result within $\pm 2\sigma$ and increase as results move away from this range. The penalty function for each constraint is defined as

$$P_{\text{constraint}}(O^{\text{th}}, O^{\text{exp}}_{\pm 2\sigma}) = \begin{cases} 0, & O^{\text{exp}}_{-2\sigma} < O^{\text{th}} < O^{\text{exp}}_{+2\sigma}, \\ |O^{\text{exp}}_{+2\sigma} - O^{\text{th}}|^2, & O^{\text{th}} > O^{\text{exp}}_{+2\sigma}, \text{ if } O^{\text{exp}}_{+2\sigma} \text{ exists}, \\ |O^{\text{th}} - O^{\text{exp}}_{-2\sigma}|^2, & O^{\text{th}} < O^{\text{exp}}_{-2\sigma}, \text{ if } O^{\text{exp}}_{-2\sigma} \text{ exists}, \end{cases} \quad (3)$$

summed over all constraints. In Eq. (3), O^{th} indicates the prediction obtained in our calculation and O^{exp} stands for experimental values with $\pm 2\sigma$ indicating the 2σ limits reported by the experiment. Note that, for constraints that only have an upper or lower bound, we can only make the comparison against corresponding upper or lower bounds. Additionally, in the case of DM direct detection, we compare against the logarithm of the value to avoid large penalty contributions from the wide range of results that can be obtained for this measurement.

For the running of the scan, we have used 140 steps in each phase of the scans (wide and narrow), testing 10^7 points through the DNN and selecting 10,000 of these that were predicted to be closer to the target space. The network used was a fully connected DNN with two hidden layers with 1000 neurons per layer trained for 1000 epochs. The output layer uses a sigmoid activation function while all the hidden layer nodes use a Rectified Linear Unit (ReLU) activation function. For training of the network we use categorical cross-entropy for the loss function with the Adaptive Moment Estimation (Adam) optimization algorithm.

As mentioned before, `DLScanner` uses `VEGAS` to create a map that transforms a random distribution to another one following weights given by a data set. This is done by training a `VEGAS` map using 100 increments with all the available data during each step. For this scan, we set the distribution of the 10^7 test points passed to the network as 75% coming from a `VEGAS` map and 25% from a random uniform distribution. The constraints used during the scan and their results are given in detail in the rest of this section.

3.2 Theoretical Constraints

Theoretical constraints such as perturbative unitarity, vacuum stability and perturbativity of the quartic couplings are essential for the validity of the model. For the NMSSM, these requirements are particularly important. Specifically, the absence of a Landau singularity for the Yukawa couplings below the GUT scale imposes a strict constraint on the NMSSM-specific coupling, λ. This constraint limits the value of λ to less than 0.7. For the minimum of the potential, NMSSMTools checks that the physical minimum is not a false minimum by comparing against the value of the potential when the scalar fields vanish. We apply all of these constraints to the sampled points using the default implementation of `NMSSMTools`.

3.3 Experimental Constraints

The list of experimental constraints that must be taken into account in our analysis is extensive and diverse, reflecting the wide range of collider searches and precision measurements relevant to the NMSSM framework. To present these in a clear and systematic manner, we organize the discussion into several subsections, each devoted to a particular class of constraints, thereby allowing us to highlight their individual roles and implications in shaping the viable parameter space.

3.3.1 Higgs Measurements

The discovery of a SM-like Higgs boson with a mass around 125 GeV by the ATLAS and CMS experiments at the LHC at CERN marks an important milestone in our quest to understand the dynamics of EWSB. Since its discovery, a major focus of research at the LHC has been to precisely measure the properties of this new particle and to verify whether it matches the predictions of the SM. So far no conclusive deviations from the SM expectations are observed, which in turn places strong constraints on any BSM parameter space. Therefore, every BSM set up with an extended scalar sector, either by modifying the couplings of the SM-like Higgs boson or by adding new BSM (pseudo)scalar states to the theory, must take into account the limits from the LHC as well as LEP and Tevatron data. To account for this large number of available searches and measurements, we use the dedicated software suite `HiggsTools` [43] that combines the implementations of `HiggsSignals` [44] and `HiggsBounds` [45], where the former checks against exclusion limits from direct searches for new Higgs particles while the latter evaluates the compatibility of the BSM scenario under consideration with the measurements of the 125 GeV Higgs boson. We scan the NMSSM parameter space and select those points where the second lightest CP-even Higgs boson (H_2) behaves like the SM-like Higgs boson with mass within a 2σ from the measured value, lying between 122 and 128 GeV. The lightest (H_1) and heaviest (H_3) CP-even Higgs boson masses are set around 95 and 650 GeV in order to accommodate the aforementioned observed excess at the LHC around the same mass ranges, we discuss these excesses in detail later in this section. The measurements around the 125 GeV Higgs boson are generally parametrized in terms of the so-called 'signal strengths' defined as follows:

$$\mu_j^i = \frac{\sigma_i \times \mathrm{BR}_j}{(\sigma_i \times \mathrm{BR}_j)_{\mathrm{SM}}} = c_i^2 \cdot \frac{\mathrm{BR}_j}{(\mathrm{BR}_j)_{\mathrm{SM}}}, \tag{4}$$

where the reduced coupling c_i^2 equals the reduced cross section for the production mode i and BR_j denotes the corresponding Branching Ratio (BR) for the decay mode j. At tree level, these reduced couplings depend on $\tan\beta$ and Higgs mixing angle. However, loop-induced reduced couplings, involving Z's, photons or gluons, may receive large contributions from other particles depending on their coupling with the Higgs boson. For all scanned points, we have calculated these reduced couplings, which are then used by `HiggsTools` to verify the consistency with the available data. In fact, a dedicated χ^2 analysis is performed using the available experimental measurements and theoretical predictions (here, NMSSM expectations) that include the correlations between the various measurements. Those points with $\Delta\chi^2$, defined as the difference of the calculated values of χ^2 from the SM and NMSSM predictions, within a 2σ interval off the central value, are considered for the subsequent analysis. This methodology helps us not only to probe the limits on the couplings relevant for the SM Higgs boson with mass close to 125 GeV but also to limit the possible amount of mixing between all three scalars, namely, the 95, 125 and 650 GeV CP-even Higgs bosons present in the BSM setup under consideration. In Figure 1, we show the values of some of these reduced couplings of the SM Higgs boson. The color bar represents the SM Higgs

boson mass while the shaded areas represent the allowed region within 2σ away from the central value. It is evident from the figure that all the scatter points are well within the allowed limits by the LHC data.
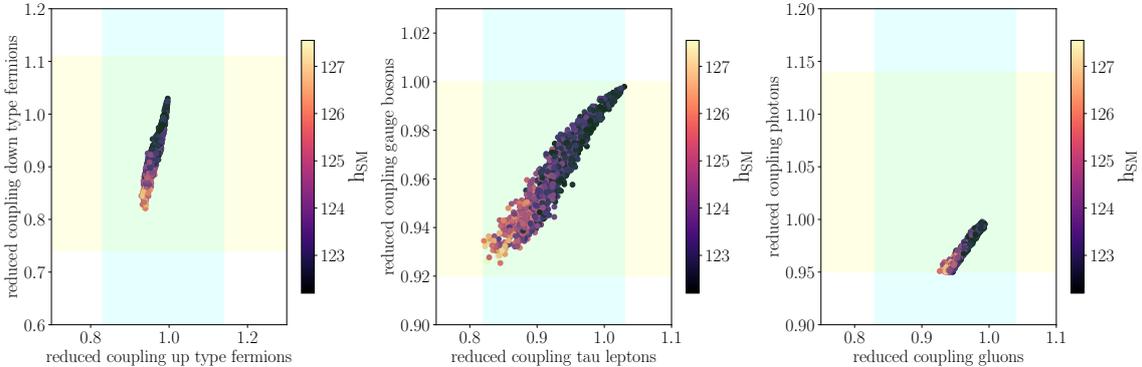

Figure 1: Reduced coupling correlations for the SM Higgs boson. The color bar represents the SM Higgs boson mass while the shaded areas represent the allowed region within 2σ away from the measured central values.

3.3.2 DM Measurements

Unraveling the true nature of DM remains one of the most profound challenges in particle physics. The constraints on the BSM parameter space coming from DM searches are broadly classified into two categories: relic density constraints and limits from the (in)direct searches. The DM relic density must satisfy the following limit as obtained from Planck data [46],

$$\Omega_{\text{CDM}} h^2 = 0.120 \pm 0.001.$$

In our study, we assume that the LSP is the single contributor to the observed total DM relic density. We employ `micrOMEGAs` to calculate the DM relic density for all the scanned data points and demand that all points must satisfy the 2σ upper bound of the observed relic density, i.e.,

$$\Omega_{\text{CDM}} h^2 < 0.122.$$

Also, we take into account the constraints from direct DM searches by calculating the spin-independent (SI) and spin-dependent (SD) scattering cross sections of the dark matter candidate, using `micrOMEGAs`. In Figure 2, we present the SI (left) and SD (right) cross sections as a function of the lightest supersymmetric particle (LSP) mass, with the color bar representing the corresponding relic density for the sampled points. All points considered are consistent with both the observed relic density and the current limits from direct detection experiments, including LZ and XENON [47, 48].

It is noteworthy that the interplay between the LSP composition and the extended Higgs sector in the NMSSM plays a crucial role in determining the scattering cross sections. Points with a significant singlino component typically lead to suppressed SI interactions, while those with more Higgsino admixture can exhibit enhanced scattering rates. Despite this variation, our sampled points remain safely below the current experimental bounds, demonstrating that the NMSSM can accommodate viable DM candidates without conflicting with direct detection data. Moreover, these points are also within the projected sensitivities of upcoming experiments such as XENON-nT and the LZ upgrade [49, 50],
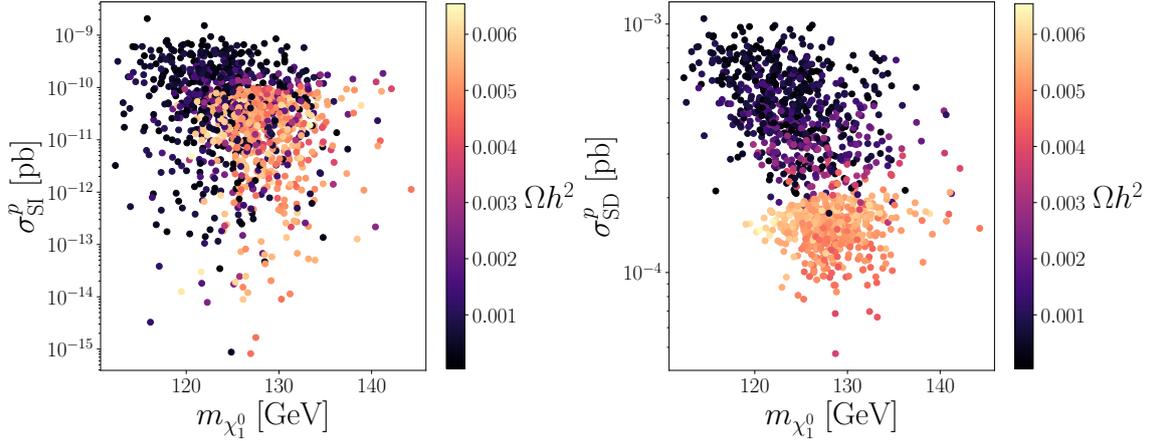
Figure 2: Left(Right): SI(SD) measurement versus the mass of the LSP. The color bar represents the relic abundance.

emphasizing that a substantial portion of the currently allowed parameter space could be probed in the near future. Regarding DM indirect detection, considering that our DM relic density tends to be of $\mathcal{O}(10^{-3})$, as shown in Fig. 2, and with most averaged annihilation cross sections times velocity typically below 10^{-25} cm^{3}s^{-1}, we do not expect indirect detection measurements [51] to be a serious concern.

3.3.3 Low Energy Observables

In the NMSSM, precision measurements in the B-meson sector impose important constraints on the allowed parameter space. Rare flavor changing neutral current processes such as $B \to X_s\gamma$, $B_s \to \mu^+\mu^-$, and $B \to \tau\nu$, as well as mass differences in neutral meson mixing like ΔM_{B_s}, are particularly sensitive to new physics contributions. These observables receive loop corrections from charged Higgs bosons, charginos, stops, and other superpartners, which can shift their predictions away from SM values. Given the high precision of current experimental results from LHCb, Belle II, and earlier B-factories, even small deviations can tightly constrain the NMSSM parameters, making B-physics an indispensable tool for probing its phenomenology in parallel with collider and dark matter searches.

The following two observables are found to play a crucial role, $B \to S\gamma$ and $B^\pm \to \tau^\pm\nu_\tau$, in constraining the values of $m_H^\pm$ and $\tan\beta$. We consider the experimental bounds as reported in [52] and allow for a 2σ deviation from their central value.

$$\mathrm{BR}(B \to X_s\gamma)_{E_\gamma \geq 1.6~\mathrm{GeV}} = (3.32 \pm 0.15) \times 10^{-4}\,, \tag{5}$$

$$\mathrm{BR}(B^\pm \to \tau^\pm\nu_\tau) = (1.11 \pm 0.165) \times 10^{-4}\,. \tag{6}$$

In the SM, the dominant contribution to $\mathrm{BR}(B \to S\gamma)$ comes from the $t - W^\pm$ loop, almost saturating the measured value. However, in SUSY theories, significant contributions may arise if the models include light SUSY particles. In the NMSSM framework, the $\tilde{t}_1 - \chi_1^\pm$ loops can provide large contributions to the above-mentioned decay mode. Similarly, decays of B-mesons to final states with τ-leptons are also sensitive to BSM effects, especially those involving a charged Higgs boson. Of course, all of these effects are mostly sensitive to the low $\tan\beta$ region. In Figure 3, we show the points allowed by these two constraints with the color mapped against values of $\tan\beta$. As expected, the allowed points have $m_H^\pm$ greater than 600 GeV with $\tan\beta$ values less than 6, thus significantly minimizing the effects of new
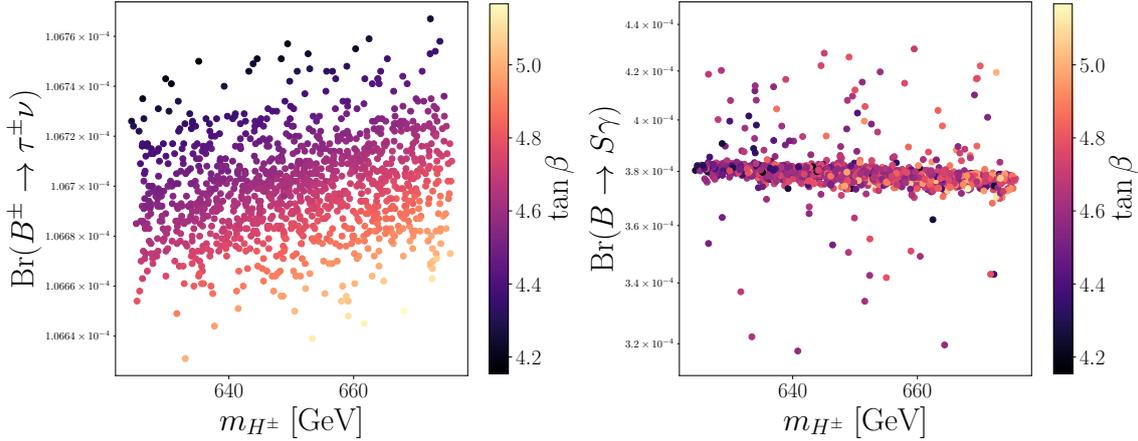
Figure 3: Left(Right): NMSSM allowed scan points following $B^{\pm} \to \tau^{\pm}\nu(B \to S\gamma)$ constraints. The color bar represents $\tan\beta$ value.

physics. The top quarks are decoupled and, therefore, the corresponding loop contribution to these flavor decays is highly suppressed. Furthermore, $\mathrm{BR}(B_s \to \mu^+\mu^-)$ has been verified and found to be consistent with the current experimental bounds of $(3.1 \pm 0.7) \times 10^{-9}$.

We also apply the constraints from EW Precision Observables (EWPOs), a powerful way to scrutinize the SM at the highest precision. We use NMSSMCALC [53], a publicly available code to calculate the oblique parameters (namely S, T, and U) for the NMSSM, and demand that the estimated values of S,T and U for the sampled points are within the 2σ of the best fit values provided by the GFITTER [54], assuming the reference mass of the Higgs boson and Top quark at 125 GeV and 173 GeV respectively,

$$S = 0.05 \pm 0.11, \quad T = 0.09 \pm 0.13, \quad \text{and} \quad U = 0.01 \pm 0.11. \tag{7}$$

3.3.4 Collider constraints

We now focus on the direct collider constraints on the NMSSM parameter space. A variety of experimental bounds on the NMSSM have been explored in the literature, although different studies have emphasized different sets of constraints. For example, Ref. [55] primarily examined LEP II limits with a particular emphasis on scenarios with small λ. The package NMSSMTools incorporates several such constraints, including Higgs searches at LEP II; however, it does not cover many direct search results from the LHC. To overcome this limitation, we employ the SModelS framework [56].

SModelS is a publicly available tool designed to interpret and constrain theories beyond the Standard Model using existing results from the LHC. Instead of performing a full simulation of complete new-physics models, SModelS decomposes their predicted collider signatures into simplified topologies defined solely by particle masses, production cross sections, and decay chains. These simplified signatures are then confronted with the extensive database of experimental limits from ATLAS and CMS. This approach allows one to efficiently determine whether a given BSM scenario is excluded or constrained by current searches without the need for dedicated recasting of each analysis.

Although in the scanned points squarks and gluinos are pushed to the TeV mass range, SModelS indicates that the strongest constraints still arise from searches for squarks and gluinos in final states with jets and missing transverse energy. In particular, the most

stringent exclusions are derived from ATLAS [57, 58] and CMS [59, 60]analyses targeting jets plus missing-energy signatures.

4 Anomalies

In this section, we present a detailed discussion of the anomalous datasets that are employed to further probe the NMSSM parameter space. The purpose of this analysis is to investigate whether the NMSSM framework can provide a consistent and viable explanation for these anomalies, thereby offering potential insights into its phenomenological relevance. For clarity and completeness, each anomaly will be examined separately and in turn, with particular emphasis on its characteristics, the methodology of its incorporation into our study and the implications it carries for the viability of the model.

4.1 The 95 GeV Excesses

Both the CMS and ATLAS collaborations at the LHC have independently observed hints of a new particle with a mass around 95 GeV. In fact, the CMS experiment has consistently seen a slight excess in this mass range, first at 97 GeV with about 2σ significance and later at 95.4 GeV with similar significance in the diphoton final state in proton-proton collisions at 13 TeV [9–11]. The ATLAS experiment has also seen a similar, albeit smaller, excess of 1.7σ at 95 GeV in the same diphoton mode [12]. Interestingly, also the ALEPH, DELPHI, L3 and OPAL (ADLO) experiments at the LEP collider have found a small excess at a mass of 98 GeV decaying to $b\bar{b}$ [61]. Given the large uncertainties in the $m_{b\bar{b}}$ resolution, the oldest observations seem consistent with the recent ones. Although none of these findings are definitive enough, the consistent appearance of this signal across different experiments is interesting on its own.

All of the above-mentioned observations are best explained in terms of the aforementioned signal strength variables. We consider two such observables, namely, $\mu_{\gamma\gamma}$ and $\mu_{b\bar{b}}$, defined as follows:

$$\mu_{\gamma\gamma} = \frac{\sigma(gg \to h_{95} \to \gamma\gamma)}{\sigma(gg \to h_{95}^{\mathrm{SM}} \to \gamma\gamma)}, \qquad \mu_{b\bar{b}} = \frac{\sigma(e^+e^- \to Zh_{95} \to Zb\bar{b})}{\sigma(e^+e^- \to Zh_{95}^{\mathrm{SM}} \to Zb\bar{b})},$$

where h_{95} denotes the new resonance of mass around 95 GeV and h_{95}^{SM} denotes the same but with SM-like couplings. The gluon-fusion process is considered as the production mode of the 95 GeV particle to analyze the $\gamma\gamma$ mode, while for the $b\bar{b}$ mode it is produced in association with a Z gauge boson. The total cross section for each point in the scan is calculated using `MadGraph` [62]. Following the combined analysis of ATLAS and CMS, we use $\mu_{\gamma\gamma} = 0.24^{+0.09}_{-0.08}$ [63] while for $\mu_{b\bar{b}}$ we use the ADLO combined limit from LEP $\mu_{b\bar{b}} = 0.117 \pm 0.057$ [64]. We demand that the signal strengths calculated for the sampled points must satisfy these limits within a 2σ range of the central value.

Interestingly, CMS has also observed an excess of local significance of 2.7σ around 95 GeV in the search for a light neutral (pseudo)scalar boson in proton-proton collisions at the LHC with $\sqrt{s} = 13$ TeV in the di-tau channel [65]. We also consider this mild but significant excess by calculating the corresponding signal strength variable as follows:

$$\mu_{\tau^+\tau^-} = \frac{\sigma(gg \to h_{95} \to \tau^+\tau^-)}{\sigma(gg \to h_{95}^{\mathrm{SM}} \to \tau^+\tau^-)} \, .$$

We thus demand that the points must further satisfy $\mu_{\tau^+\tau^-} = 1.38^{+0.69}_{-0.55}$ [65]. Figure 4 displays the distributions of the various signal strength variables together with the relevant

branching ratios of the 95 GeV resonance considered in this study. As can be seen, the majority of the scanned points lie well within the 2σ contour in the two-dimensional plane spanned by the signal strength variables. In Section 6, we present the explicit parameter space corresponding to points that remain within the 2σ level when all anomalies and constraints are simultaneously taken into account.
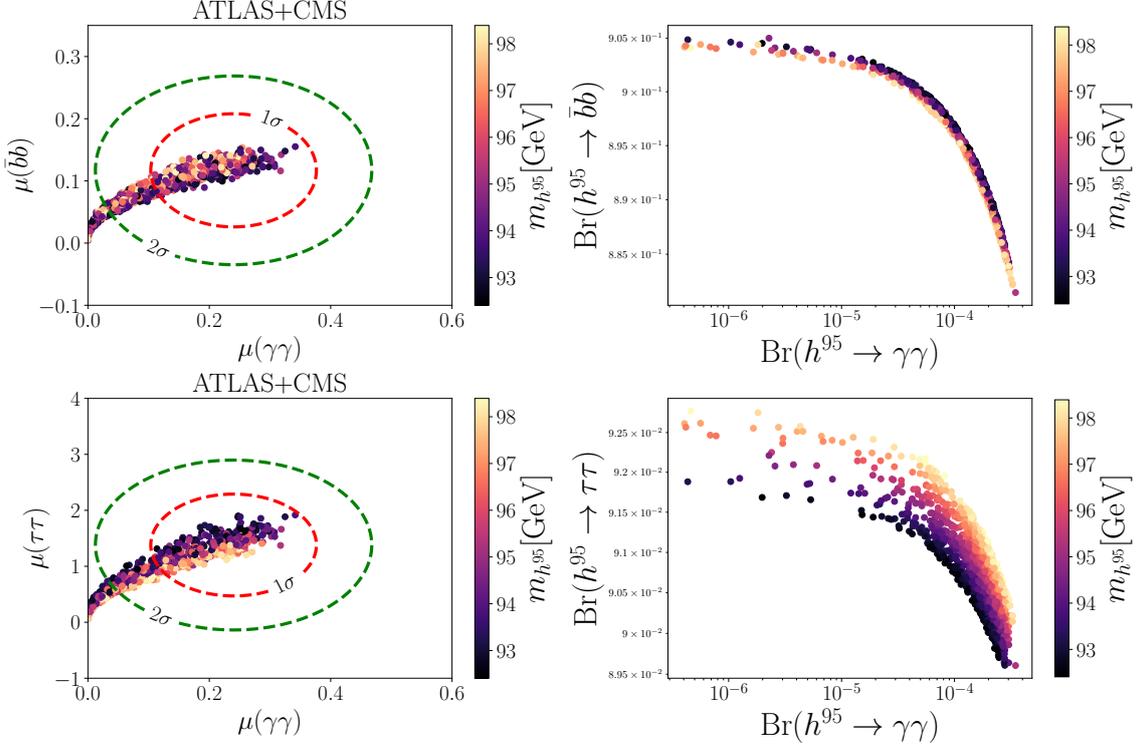

Figure 4: Distribution of points in the 2D plane of relevant signal strengths (left) and BRs (right) in order to explain the excess at 95 GeV. The color gauge represents the h_{95} mass.

4.2 The EWino Excess

We now turn our attention to another interesting observation made by the ATLAS and CMS collaborations at the LHC. Interestingly, both experiments have observed mild excesses (only at the 1-2σ level) of events over the SM expectation corresponding to the final states with multiple leptons and large $\not{E}_T$. Specifically, these events correspond to chargino-neutralino pair production processes where chargino and neutralino decays proceed via an off-shell $W^\pm, Z$ boson leading to soft leptons [20–22, 66]. The simplified SUSY models considered in these studies are of two types, wino/bino and Higgsino scenarios. In the wino/bino scenario, the $\tilde{\chi}_1^0$ (i.e., the LSP) is considered to be purely bino whereas $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^\pm$ are taken purely wino and degenerate in mass. The simplified Higgsino model assumes $\tilde{\chi}_1^0$, $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^\pm$ to be Higgsino-like states with compressed mass spectrum with $\tilde{\chi}_1^\pm$ lying approximately halfway between the two light neutralinos. In the wino/bino scenario, the excess as reported in the observed versus expected no of events appears when the mass difference $\Delta m = m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0} \sim 10\text{–}30$ GeV when $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^\pm$ decays to LSP plus the off-shell $W^\pm, Z$ boson. In the Higgsino scenario, the excess is observed in the range $\Delta m \sim 10\text{–}30$ GeV with the most pronounced effect at $\Delta m \sim 30$ GeV for the ATLAS collaboration. The CMS

collaboration has also reported a similar excess of events observed in SUSY searches with, soft, multi-lepton plus $\not{E}_T$ final state. Although CMS finds an excess of events in the range $\Delta m \sim 8$–30 GeV for the Higgsino case, an excess of events is observed in the region of $\Delta m = 20$–50 GeV for the wino/bino scenario. Considering the mass resolution, both excesses observed by the two collaborations are compatible with each other, with CMS finding excesses more on the higher mass differences. Interestingly, similar excesses have also been observed in mono-jet searches around the same mass region [23, 67].
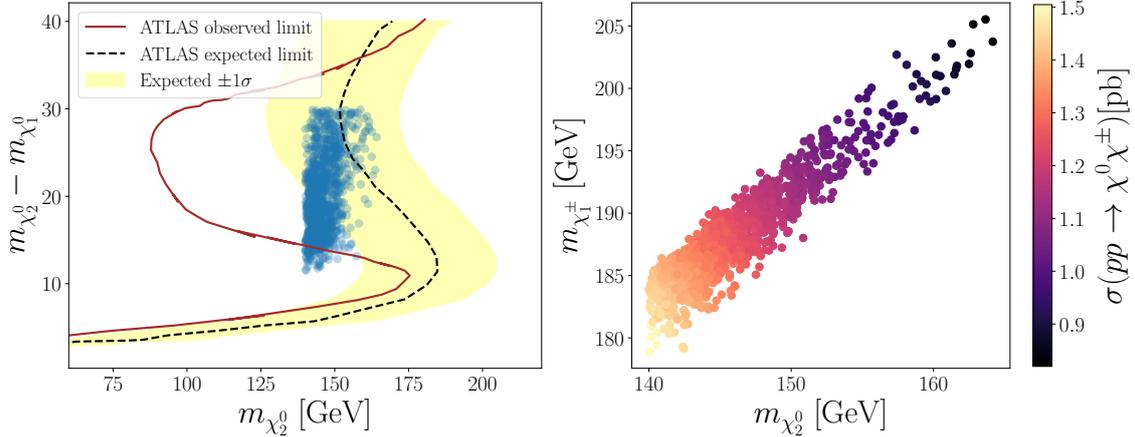

Figure 5: Left: Expected and observed ATLAS limits with scanned points overlaid in blue, which are demonstrating clear NMSSM compatibility with the EWino anomaly. ATLAS data are extracted from Figure 16 of [21]. Right: Cross section for chargino–neutralino pair production.

Motivated by Ref. [18], in our analysis, we study the Higgsino model scenario in the NMSSM and consider a conservative limit $\Delta m = 10$-30 GeV to identify the correct region of the parameter space. Figure 5 illustrates the distribution of sampled points that satisfy the previously discussed constraints. The left panel shows the exclusion limits in the $m_{\tilde{\chi}_2^0} - \Delta m$ plane, while the right panel depicts the chargino–neutralino pair-production cross section in the $m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^\pm}$ plane. The phase space regions providing the correct cross section have both the neutralinos and lightest chargino masses below 250 GeV, as pointed out in other studies [68]. The production cross section for the scanned points is done by using `MadGraph`.

For completeness, in Figure 6, we show the amount of Higgsino and singlino components in the LSP and Next-to-LSP (NLSP) in terms of the elements of the neutralino mixing matrix $\mathcal{Z}$, especially the Higgsino-singlino elements (ZNH and ZNS) with the wino-bino elements ($ZNB = ZNW$) vanishing small. Note that the NLSP has been found to be of pure Higgsino nature in order to be consistent with the observed excess.

4.3 The 650 GeV Excess

A search for the resonant production of a Standard Model Higgs boson in association with a new scalar particle, denoted as Y, via a heavy resonance X, observed a local excess of 3.8σ in the final state consisting of two photons (from the decay of the h_{SM}) and two b-quarks (from the decay of Y) [13]. The 95% confidence level (CL) upper limit on the product of the production cross section and branching ratio is approximately 0.8 fb. However, as reported in the experimental analysis, the best fit value of the cross section with the branching fraction in the largest excess is $0.35^{+0.17}_{-0.13}$ fb. Therefore, the anomaly in the $bb\gamma\gamma$
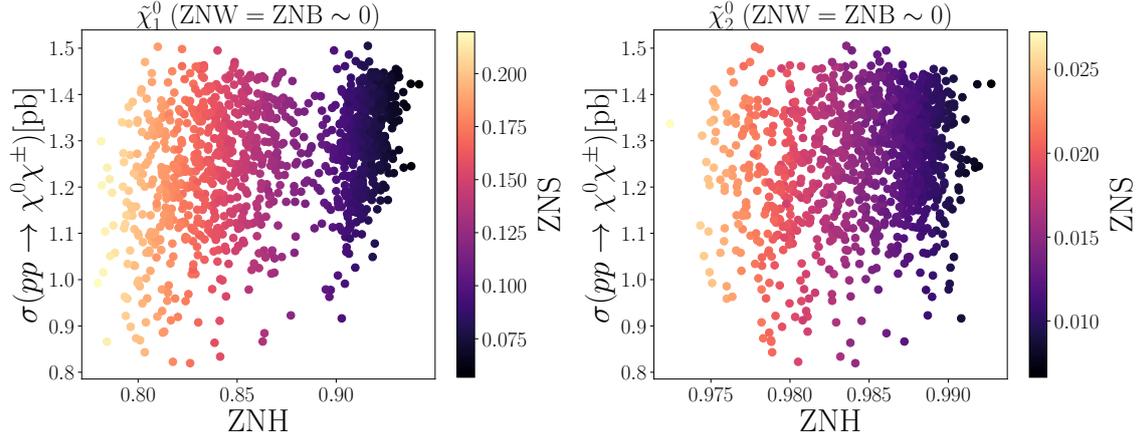
Figure 6: Plots to identify the LSP and NLSP types. The NLSP is a pure Higgsino, which matches the experimental measurement.

channel can be fulfilled if the total cross section lies within the best fit value. This excess was found for a strong resonance X with $M_X \sim 650$ GeV and $M_Y \sim 90 - 100$ GeV. A previous study by the CMS Collaboration searched for a heavy resonance X (with 240 GeV $< M_X < 3$TeV) decaying into a SM Higgs boson (h_{SM}) and a lighter scalar h_s (with 60GeV $< m_{h_s} < 2.8$ TeV). In this search, the h_{SM} decays into a pair of τ leptons, while the lighter scalar h_s decays into a pair of b quarks [14]. No significant excess was observed, leading to model-independent 95% CL upper limits on the product of the production cross section and BRs for a wide range of X and $Y \equiv h_s$ masses ranging between 125 fb (for $m_X = 240$ GeV) to 2.7 fb (for $m_X = 1000$ GeV).
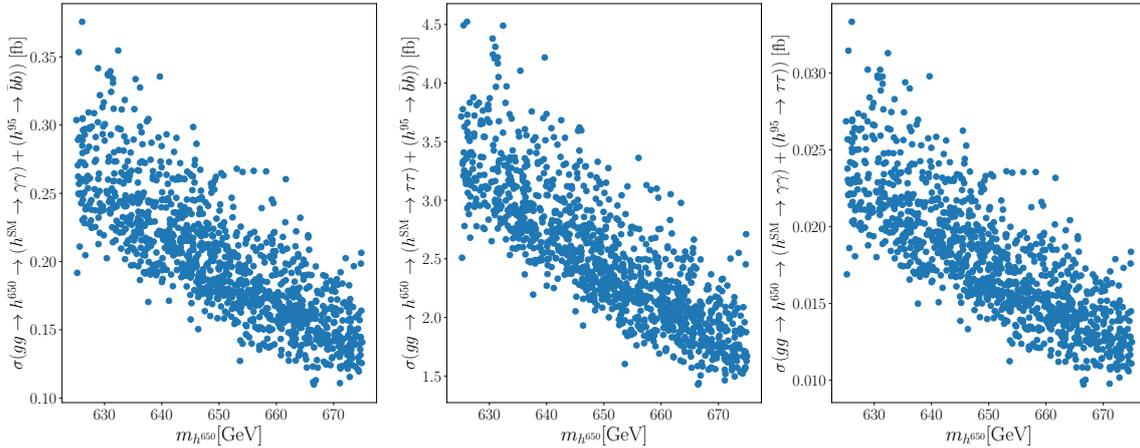

Figure 7: The distribution of $\sigma \times$ BR values as a function of the $h^{650}(\equiv H_3 \equiv X)$ mass for all the sampled points associated to the excess of events observed around 650 GeV.

Complementing these searches, CMS has recently performed a dedicated analysis targeting heavy resonances (X) decaying into the Standard Model Higgs boson (h_{SM}) and a lighter scalar (h_s), with the characteristic final state $\gamma\gamma\tau^+\tau^-$ [15]. This channel is particularly interesting as it probes scenarios with $m_{h_s} \approx 95$GeV, thereby directly testing one of the most intriguing excesses reported in recent years. The analysis sets stringent upper

limits on the production cross section times branching ratio for $X \to h_{\mathrm{SM}} h_s \to \gamma\gamma\tau^+\tau^-$, ranging from 0.69 to 15fb in the low X-mass region. When combined with previous searches in complementary final states, such as $b\bar{b}\tau^+\tau^-$ and $\gamma\gamma b\bar{b}$, these results provide a powerful handle on probing the extended Higgs sector of the NMSSM.

We have incorporated all these experimental results to constrain the parameter space relevant to our study. As shown in Figure 7, the combination of limits from different final states significantly reduces the viable regions of parameter space, particularly those predicting sizable branching fractions into $b\bar{b}\tau^+\tau^-$. This highlights the strong interplay between different Higgs decay channels in shaping the phenomenology of the NMSSM and underlines the importance of multi-channel searches in covering scenarios where one individual signature might otherwise evade detection.

4.4 The $(g-2)_\mu$ Measurement

The anomalous magnetic moment of the muon, $a_\mu \equiv (g-2)_\mu/2$, ranks among the most precisely determined quantities in physics, both experimentally and theoretically. This observable plays a pivotal role, as the persistent discrepancy between its measured value and the SM prediction could offer vital clues to the existence of BSM physics. The Fermilab Muon $g-2$ experiment (E989) has reported the results from its Run-1 dataset, collected in 2018 [69–71], which show excellent agreement with the earlier E821 measurement performed at BNL [24]. The discrepancy between the experimental measurements and the SM prediction [72] is

$$a_\mu^{\mathrm{exp}} - a_\mu^{\mathrm{SM}} = \Delta a_\mu = (24.9 \pm 4.8) \times 10^{-10}\,, \tag{8}$$

corresponding to a 5σ discrepancy. However, this result does not incorporate the lattice calculation for the leading order hadronic vacuum polarization contribution [73]. Inclusion of this theory prediction would yield a slightly higher value for Δa_μ leading to a smaller deviation from the experimental result of $< 2\sigma$. This leads to a difference of [74]

$$a_\mu^{\mathrm{exp}} - a_\mu^{\mathrm{SM}} = (38 \pm 63) \times 10^{-11}\,. \tag{9}$$

In the NMSSM, the dominant one-loop contributions to the anomalous magnetic moment of the muon, a_μ, arise from loops involving chargino–sneutrino or neutralino–smuon pairs. These contributions grow approximately linearly with $\tan\beta$ and are positive when the effective Higgsino mass parameter, μ_{eff}, is positive. Figure 8 presents the corresponding values of Δa_μ for the points selected in our scans. The left panel shows Δa_μ as a function of the sum of neutralino masses, with the color bar indicating the sum of the chargino masses. This plot demonstrates a clear trend of increasing Δa_μ for smaller values of either the chargino or neutralino masses. The right panel displays the same information in the plane of Δa_μ versus μ_{eff}, with the color bar representing $\tan\beta$, highlighting the dependence on these parameters. Overall, these results indicate that the selected points are consistent with the most recent measurement of the muon anomalous magnetic moment, as given in Equation 9, thereby providing further support for the viability of the NMSSM parameter space under consideration.

5 Our Predictions

In the spirit of not only producing NMSSM 'postdictions', i.e., of explaining existing anomalous data *a posteriori*, but also 'predictions', i.e., pointing to *new* datasets that will reveal themselves also anomalous if investigated over the viable parameter space, we chose among the latter two collider signatures of DM.
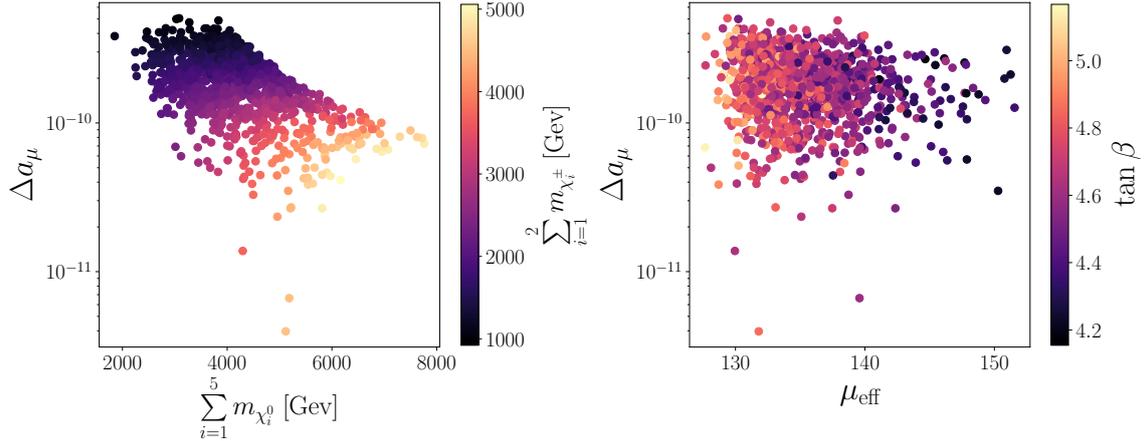
Figure 8: Left: Muon anomalous magnetic moment for the selected points plotted against the sum of neutralino masses on the x-axis, with the color scale indicating the sum of chargino masses. Right: Muon anomalous magnetic moment for the selected points plotted against μ_{eff} on the x-axis, with the color scale indicating $\tan\beta$.

5.1 Mono-H and -Z Signals

In this subsection, we investigate the sensitivity of the HL-LHC to mono-X signatures, focusing on BPs selected from the optimized NMSSM parameter space. Specifically, we study the production processes

$$pp \to \tilde{\chi}_1^0 \tilde{\chi}_1^0 H \quad \text{and} \quad pp \to \tilde{\chi}_1^0 \tilde{\chi}_1^0 Z,$$

where the lightest neutralinos ($\tilde{\chi}_1^0$) are weakly interacting massive particles that escape detection, leading to $\not{E}_T$. These channels are particularly relevant for probing EWino DM at the HL-LHC. In principle, these signatures can lead to additional anomalies in future measurements that align with the current anomalies mentioned above. For this purpose, we employ a multi-modal network, that integrates graph-based DL layers and MLP layers, to analyze these signatures and enhance discrimination between signal and background. This approach integrates both high-level physics observables and low-level detector features, offering improved performance compared to conventional cut-based or shallow Machine Learning (ML) methods.

Although mono-jet channels have been widely studied and offer complementary sensitivity to DM and EWino searches [75–77], our focus here is on the mono-H and -Z channels, which provide distinctive kinematic features and are sensitive to different regions of the NMSSM parameter space. Feynman diagrams corresponding to the leading-order contributions to these channels are shown in Figure 9, with the upper row depicting the mono-Z processes and the lower row representing mono-H production.

To quantify the expected signal rates, we compute the production cross sections at the HL-LHC for each of the mono-X channels. Figure 10, left panel, shows the total cross sections as a function of the sum of the masses of the lightest and next-to-lightest neutralinos. The mono-jet, mono-Z and mono-H channels are shown in blue, orange and green, respectively. The red and black markers indicate the specific BPs selected for the DL analysis in the mono-Z and mono-H channels, respectively. The right panel displays the cross sections as a function of the NMSSM parameter μ_{eff}, further illustrating the correlation between model parameters and signal strengths.
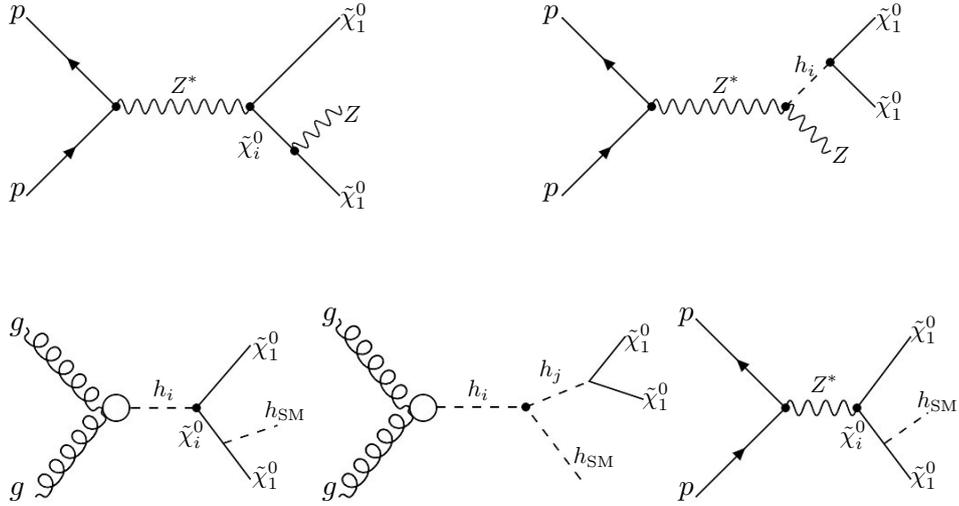
17

Figure 9: Feynman diagrams illustrating the leading contributing channels. The top row shows mono-Z processes, while the bottom row displays mono-H processes.
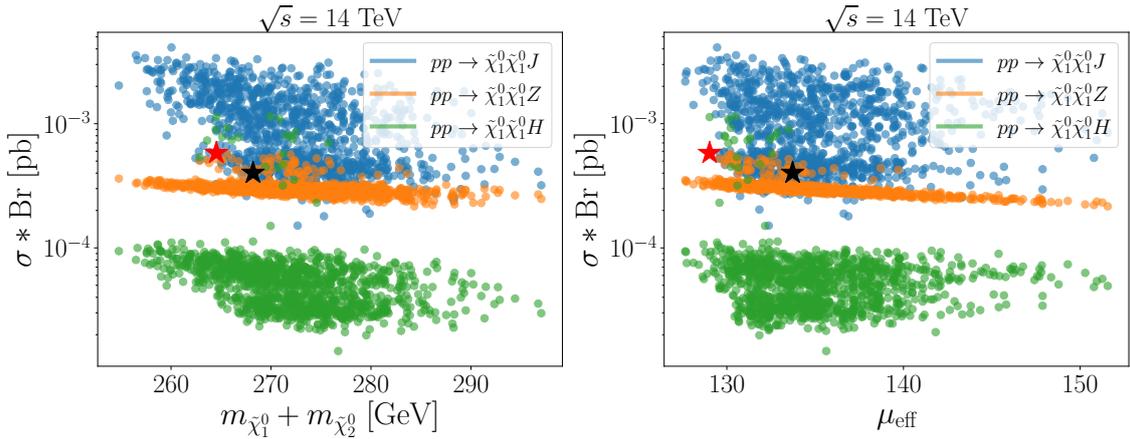

Figure 10: Left: Total production cross sections for mono-jet (blue), mono-Z (orange) and mono-H (green) channels as a function of the combined mass of the LSP and NLSP neutralinos. Right: Cross sections versus the NMSSM parameter μ_{eff}. In both plots, the red and black points correspond to the chosen BPs for the mono-Z and -H analyses, respectively.

The full set of NMSSM input parameters for the selected BPs is provided in Table 2. These points are chosen to optimize the signal significance while satisfying existing experimental constraints from DM relic density, direct detection limits and collider bounds.

5.2 Background

Having selected the BPs for the signal, we now turn to the analysis of the corresponding background processes. The most significant background arises from top quark production, including both $t\bar{t}$ and single top processes. In particular, $t\bar{t}$ production contributes to

	$\tan\beta$	λ	κ	A_λ	A_κ	μ_{eff}	M_1	M_2	M_3	A_t	$m_{\tilde\chi_1^0}$
mono-H	4.62	2.6×10^{-1}	1.8×10^{-1}	6.1×10^2	-2.97×10^2	1.33×10^2	2.7×10^3	1.08×10^3	3×10^3	-4.9×10^3	1.26×10^2
mono-Z	4.6	2.9×10^{-1}	2.2×10^{-1}	5.6×10^2	2.9×10^2	1.29×10^2	2.49×10^3	9.1×10^2	1.4×10^3	-4.6×10^3	1.2×10^2

Table 2: NMSSM input parameters for the BPs used in the mono-H and -Z analyses.

the final state with $b\bar{b}$ pairs through its hadronic, semileptonic and leptonic decay modes. Among these, the semileptonic decay dominates due to its relatively high production rate and the presence of substantial $\not{E}_T$. The leptonic decay mode, while less prominent, also contributes non-negligibly, with $\not{E}_T$ arising from the two neutrinos produced in leptonic W decays. However, the application of a lepton veto with a transverse momentum threshold of $p_T > 10$ GeV effectively reduces this background at the selection level. In contrast, the semileptonic background is less suppressed by such a veto. The hadronic decay of $t\bar{t}$, despite having the largest cross-section, contributes primarily through fake $\not{E}_T$ originating from jet energy mismeasurement. A detailed simulation shows that the hadronic $t\bar{t}$ background plays only a sub-dominant role. The single top background is also considered, though its impact remains minor due to a much lower production cross-section compared to the semileptonic and leptonic $t\bar{t}$ channels.

Another significant source of background in the signal region is vector boson production in association with jets, denoted as $V+$jets (where $V = W^\pm, Z$). These processes feature large production cross-sections of order $\sim 10^3$ pb and can generate considerable $\not{E}_T$ via the semileptonic decays of the weak bosons. However, these backgrounds rely on the misidentification of two light-flavor jets as b-jets, a scenario with a low double-mistag probability of approximately 0.04% as estimated from our Monte Carlo simulations. Notably, the contribution from $W+$jets is found to be sub-dominant relative to $Z+$jets. This is primarily due to the higher $\not{E}_T$ typically observed in $Z+$jets events and the suppression of $W^\pm+$jets through the lepton veto.

The QCD production of $b\bar{b}$ pairs presents a major challenge for this analysis, given its extremely large cross-section of around 10^5 pb. However, the impact of this background is highly sensitive to the $\not{E}_T$, which in this case arises from jet energy mismeasurement. By optimizing the cut on $\not{E}_T$ this background becomes less relevant relative to those from $t\bar{t}$ and $Z+$jets. In this analysis we do not consider this background.

Another irreducible background also comes from the process $Z(\to \nu\bar{\nu})h_{\text{SM}}(\to b\bar{b})$. This process yields a final state with both significant $\not{E}_T$ and a $b\bar{b}$ invariant mass distribution similar to that of the signal. However, due to its comparatively small cross-section, approximately 100 fb, this background remains less significant than the others discussed above.

Based on the kinematics of signal and backgrounds, in order to generate the events, more effectively, we apply the following cuts at the simulation/generation (i.e., parton) level: $p_T(\text{jet}) \geq 20$ GeV, $p_T(l) \geq 20$ GeV, $p_T(b) \geq 20$, MET ≥ 30 GeV and $|\eta(l,\text{jet},b)| \leq 2.5$ (hereafter, $j = e, \mu$). Moreover, we require the events to have at least two b-jets with cone radius $R = 0.4$ using a flat tagging efficiency of 75% where as for the mis-tagging efficiency of gluon- and light-quark-jets as b-ones, we adopt a flat rate of 10^{-3}. After generating the events we reconstruct nine kinematic distributions for both signal and backgrounds as detailed below. To collect the two b-jets from the Higgs boson decay we do not require any mass window cut rather we collect the b-jet pair that has the closest invariant mass to the SM-like Higgs boson mass.

Selection cuts and the corresponding signal and backgrounds efficiency for mono-H and -Z are reported in Tables 3 and 4, respectively. The selection cuts are optimized to

Selection cuts	$\bar{t}t$ leptonic	$V+jj$	VV	Zh_{SM}	mono-H
Cross section [pb]	26.9	516	0.587	0.106	4×10^{-4}
$\not{E}_T \geq 200$ GeV	1.5×10^{-2}	1.5×10^{-3}	2×10^{-2}	4.7×10^{-2}	0.436
$N(b)\geq 2$	1.2×10^{-2}	1.1×10^{-3}	1.4×10^{-2}	3.5×10^{-2}	0.325
$p_T(b_1)\geq 50$, $p_T(b_2)\geq 30$ GeV	1.8×10^{-3}	1.7×10^{-4}	3×10^{-3}	1.4×10^{-2}	0.13
80 GeV $\leq$ $m(b\bar{b})$ $\leq$ 140 GeV	5×10^{-4}	6×10^{-5}	2.6×10^{-3}	1.2×10^{-2}	0.121
$\Delta\phi(b_{1,2},\not{E})>0.35$	4×10^{-4}	4×10^{-5}	2.4×10^{-3}	1.1×10^{-2}	0.120

Table 3: Signal and background efficiencies in the $H(\to b\bar{b})+\not{E}_T$ final state after applying selection cuts at 14 TeV. Selection efficiencies are computed as the fraction of events remaining after applying the selection cut.

Selection cuts	$\bar{t}t$ leptonic	$V+jj$	VV	Zh_{SM}	mono-Z
Cross section [pb]	26.9	516	0.587	0.106	5×10^{-4}
$\not{E}_T \geq 200$ GeV	1.5×10^{-2}	1.5×10^{-3}	2×10^{-2}	4.7×10^{-2}	0.38
$N(b)\geq 2$	1.2×10^{-2}	1.1×10^{-3}	1.4×10^{-2}	3.5×10^{-2}	0.29
$p_T(b_1)\geq 30$, $p_T(b_2)\geq 20$ GeV	1.1×10^{-2}	5.5×10^{-4}	4.8×10^{-2}	2.8×10^{-2}	0.182
50 GeV $\leq$ $m(b\bar{b})$ $\leq$ 120 GeV	6×10^{-3}	3×10^{-4}	4.3×10^{-2}	2.6×10^{-2}	0.12
$\Delta\phi(b_{1,2},\not{E})>0.35$	5.5×10^{-3}	3×10^{-4}	2.5×10^{-2}	1.1×10^{-2}	0.11

Table 4: Signal and background efficiencies in the $Z(\to b\bar{b})+\not{E}_T$ final state after applying selection cuts at 14 TeV. Selection efficiencies are computed as the fraction of events remaining after applying the selection cut.

maximize the signal-to-background yield resulting in different ranges for both analyses.

5.3 Analysis Setup and Data Structure

To exploit the differences in QCD color flow between signal and background processes, we incorporate both reconstructed high-level kinematic variables and color flow information as input features to a multi-modal DL network. This approach allows the model to learn complementary aspects of the event topology: one related to momentum-space features and the other rooted in the QCD dynamics governing the QCD color flow.

The color flow of a QCD process is dictated by the underlying $SU(3)$ color symmetry, which enforces color conservation at each vertex of the interaction. The way color indices are contracted determines the radiation pattern observed in the final state. For example, in processes where jets originate from a color singlet, such as a Higgs or Z boson, the color indices of the resulting quarks are contracted with each other, forming a color dipole. This typically results in soft hadronic radiation between the two jets. In contrast, for jets originating from the decay of a colored particle, such as a top quark, the color indices of the outgoing quarks are connected to those of the initial-state protons. This leads to isolated color poles and a soft radiation pattern that aligns more closely with the beam directions [78–80].

The global color flow structure manifests itself in the detector as a color string, a flux tube of soft hadrons connecting color correlated final-state particles. In particular, the distribution of soft hadrons surrounding hard jets encodes valuable information about the color connections of the event [81, 82]. This soft radiation pattern is sensitive to whether the two jets are color-connected to each other or to the beam remnants and can thus serve as a powerful discriminator between signal and background topologies. In our case, the signal consists of two b-quarks arising from the hadronic decay of a color singlet boson. These quarks form a color dipole and their associated radiation tends to fill the region between them. Conversely, in the dominant $t\bar{t}$ background, the two b-quarks are produced from the decay of color triplet top quarks. These b-quarks are individually color-connected to the incoming beams, resulting in two largely isolated radiation cones and a depletion of soft hadronic activity between them. To capture this contrast, we design a multi-modal DL network with two distinct input streams: one capturing high-level kinematic variables and the other designed to characterize color flow features. For the latter, we utilize observables that are sensitive to soft radiation patterns surrounding the b-jets, such as energy flow and the transverse momentum density of all charged and neutral hadrons. Specifically, we consider six features for each final-state hadron: transverse momentum, energy, mass, charge, pseudorapidity and azimuthal angle. These features are effective in capturing the differences in QCD color flow between signal and background events.

To complement the color flow observables, we use a set of high-level reconstructed kinematic variables as inputs to the DL network. These variables are chosen based on their ability to distinguish signal events from the dominant backgrounds. The list of kinematic observables is as follows:

- MET or $\not{E}_T$: This is defined as

$$\text{MET} = -\left| \sum_{v_i} \vec{p}_T(v_i) \right|,$$

  where the sum runs over all visible particles. Signal events involving neutralino DM, with mass around 100 GeV, typically exhibit large MET due to the escaping LSPs, which helps to suppress $t\bar{t}$ and di-boson backgrounds.

- $p_T(b_1)$: Transverse momentum of the leading b-jet. In the signal process, the b-jets originate from the decay of a boosted SM boson resulting from a heavy mediator. This results in a p_T spectrum that closely resembles that from $t\bar{t}$ events (where b-jets arise from top decays), but differs significantly from other background processes.

- $p_T(b_2)$: Transverse momentum of the subleading b-jet.

- $\eta(b_1)$: Pseudorapidity of the leading b-jet. While this variable shows similar distributions for signal and most backgrounds, the $pp \to Zb\bar{b}$ process exhibits a broader η distribution.

- $\eta(b_2)$: Pseudorapidity of the subleading b-jet, generally following the same trend as $\eta(b_1)$ across all processes.

- $p_T(b\bar{b})$: Transverse momentum of the b-jet pair that best reconstructs the boson mass, either Higgs or Z. This variable highlights the boost of the Higgs or Z boson in the signal events.

- $\eta(b\bar{b})$: Pseudorapidity of the b-jet pair.

- $\Delta R(b, \bar{b})$: Angular distance between the two b-jets reconstructing the Higgs or Z boson, defined as
$$\Delta R(b, \bar{b}) = \sqrt{(\Delta\eta(b, \bar{b}))^2 + (\Delta\phi(b, \bar{b}))^2}.$$

  In ZZ and hZ background processes, where the b-jets come from the decay of a gauge or Higgs boson produced near resonance, the jets are mostly back-to-back, peaking at $\Delta R(b, \bar{b}) \sim 3$. For QCD-induced $Zb\bar{b}$ events, the angular separation has a broader distribution. In contrast, signal events with boosted bosons produce b-jets with smaller separation.

- $m(b\bar{b})$: Invariant mass of the b-jet pair closest to the Higgs or Z boson mass.

To prepare the data for DL training, we construct two distinct datasets. The first dataset captures QCD color flow information and has the shape $(N, 80, 6)$, where N is the number of training events, 80 is the fixed number of hadrons considered per event and 6 is the number of features for each hadron. For events with more than 80 hadrons, only the 80 with the highest momenta are selected; for events with fewer hadrons, the remaining slots are zero-padded. The second dataset encodes high-level kinematic variables of the final-state particles and has the shape $(N, 9)$, where the 9 input features are as described earlier. To ensure balanced training, we use equal-sized datasets for signal and background events, with $N = 600{,}000$ for each class. Since this is a supervised classification task, we assign binary labels: $Y = 1$ for signal events and $Y = 0$ for background events. During training, the network iteratively adjusts its weights to minimize classification error, learning to correctly associate the input features with their corresponding labels until satisfactory accuracy is achieved.

For event simulation, we employ `MadGraph5` [62, 83] for cross section estimation and generating parton-level events. Then `Pythia8.3` [84] is utilized to include parton showering and hadronization effects. The factorization and renormalization scales have been kept at the default `MadGraph` event by event dynamic choice. Jets are formed using `FastJet` package [85] utilizing anti-KT algorithm [86] with $R = 0.4$. Detector effects are taken into account with the `Delphes` package [87] using the default ATLAS card. For the DL analysis, we use `PyTorch Geometric` [88] for building the GNN layers, while standard `PyTorch` [89] is used for the MLP. Finally, the `Scikit-Learn` package [90] is used to facilitate network training and evaluation.

5.4 DL Network Structure

To effectively analyze two distinct input datasets, we employ a multi-modal NN that combines a Graph Convolutional Network (GCN) with a Multi-Layer Perceptron (MLP). The GCN component is particularly well-suited for learning from graph-structured data, as it incorporates topological relationships among nodes and edges, enabling the model to capture spatial and relational patterns. In our framework, we represent the final-state hadrons as a fully connected graph, where each node corresponds to a hadron and is described by a feature vector $x = (M, p_T, E, \eta, \phi, \text{charge})$. This vector encapsulates the properties of the corresponding particle, where p_T denotes transverse momentum, E is the energy, η represents pseudorapidity and ϕ is the azimuthal angle. The graph is fully connected, with edges weighted by the angular distance $\Delta R_{(x_i, x_j)}$ between the particles in nodes i and j, thereby enabling the model to capture intricate spatial relationships between such particles. The primary objective of the GCN is to learn a function that maps the input features to new representations, effectively capturing the complex relationships between the graph nodes [91–93]. The core idea behind GCNs is the extension of the convolution operation

from regular grid structures, as in CNNs, to irregular graph domains. A graph convolution can be interpreted as a local aggregation of information from neighboring nodes, combining both the structural properties of the graph and the features associated with each node. Given an input graph $G = (V, E)$, the graph convolution operation is defined as

$$H^{(l+1)} = \sigma \left(\hat{D}^{-\frac{1}{2}} \hat{A} \hat{D}^{-\frac{1}{2}} H^{(l)} W^{(l)} \right),$$

where $H^{(l)} \in \mathbb{R}^{N \times F_l}$ denotes the feature matrix at layer l, with N being the number of vertices in the graph and F_l the dimensionality of the feature space at that layer. The matrix $W^{(l)} \in \mathbb{R}^{F_l \times F_{l+1}}$ is the trainable weight matrix at layer l and σ represents the non-linear activation function. The adjacency matrix of the graph, is given by $\hat{A} \in \mathbb{R}^{N \times N}$, defined as $\hat{A} = A + I_N$, where A is the original adjacency matrix and I_N is the identity matrix of size N. The diagonal degree matrix corresponding to $\hat{A}$ is denoted by $\hat{D} \in \mathbb{R}^{N \times N}$, with elements $\hat{D}_{ii} = \sum_j \hat{A}_{ij}$, which represent the number of connected edges to vertex i. The graph convolution operation can be interpreted as a message-passing scheme, where each node aggregates information from its neighbors and updates its feature representation using the learned weights. By stacking multiple GCN layers, the model is able to learn higher-order interactions among nodes, thereby capturing increasingly complex relational structures within the graph representing the color flow for different processes. The second stream of the network, which processes the kinematic information, begins with an input layer of nine neurons. This is followed by a series of fully connected layers. The outputs from both streams are then concatenated and passed to the final output layer for classification. Illustration of the schematic architecture of the network is shown in Figure 11.
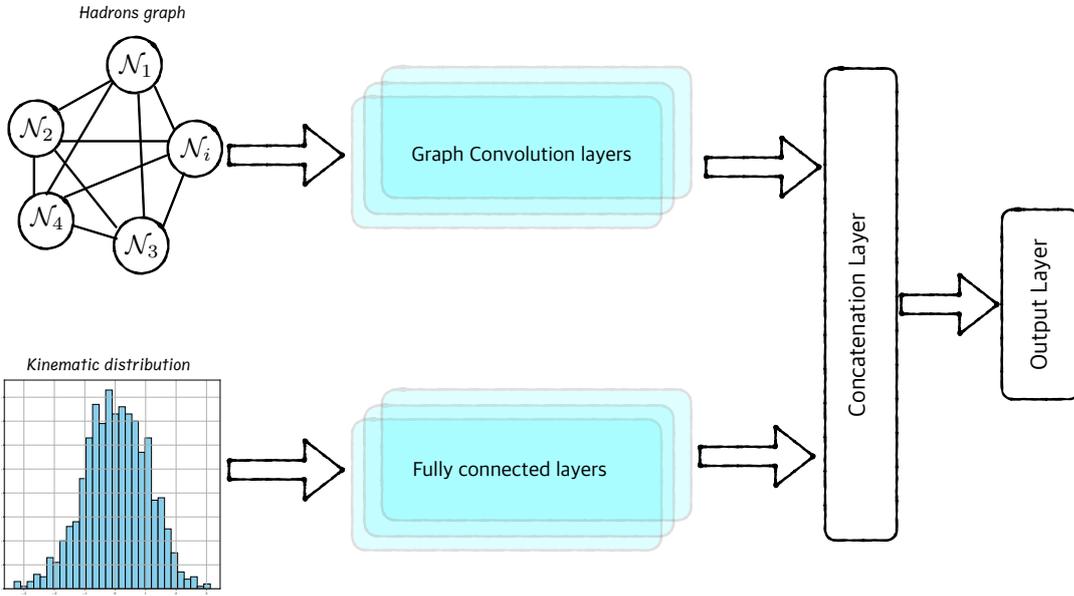

Figure 11: Schematic diagram of the multi-modal GNN architecture, illustrating the data flow from the input graph of hadrons through the GCN to produce intermediate representations. These representations are then concatenated with high-level kinematic variables and passed through an MLP, resulting in the final classification output.

Designing a DL network involves specifying its architecture, which includes a number of parameters whose values must be determined before training. These are known as hyperparameters and, unlike model weights, they cannot be learned during training and must

be set manually. The performance of the network is highly sensitive to the choice of these hyperparameters, making their careful selection essential for achieving optimal results. A common method for hyperparameter tuning is grid search, one of the most straightforward techniques. In this approach, the model is trained and evaluated for every possible combination of predefined hyperparameter values. The combination that yields the best performance is then selected. While grid search can be effective in identifying optimal configurations, it is computationally expensive and often impractical for large search spaces. To address this limitation, we adopt a more efficient alternative known as random search. Instead of exhaustively testing all combinations, random search samples hyperparameter values from specified probability distributions. This significantly reduces the computational cost while still allowing for the discovery of well-performing configurations, especially when only a subset of hyperparameters strongly influences performance. After hyperparameters optimization the best network structure consists of 3 GCN layers with ReLU activation function. This is followed by max pooling to aggregate the node embedding. The other part of the model is a basic set of 3 fully connected layers with number of neurons of 128, 512 and 256. Output from both streams are concatenated in on layer and passed to an output layer with two neurons and softmax function.

5.5   Sensitivity at the (HL-)LHC

After preparing the datasets, we train the networks to learn the complex, non-linear mapping between the input features and their corresponding labels. Signal events are assigned the label $Y = 1$, while background events are labeled $Y = 0$. To eliminate any bias related to the ordering of signal and background events, we combine both classes into a single dataset and shuffle the events along with their assigned labels. During training, the network proceeds through multiple epochs, where each epoch is defined as one complete pass over the entire dataset. For every iteration, the network updates its weights using backpropagation, which propagates the prediction error backward to adjust the network parameters. The training objective is to minimize a loss function that quantifies the discrepancy between predicted outputs and the true labels, thereby guiding the network toward an optimal set of weights. In this network we employ the categorical cross-entropy loss function, defined as $-\sum_i Y_i \log(\hat{Y}_i)$, where $i = 0, 1$ corresponds to the background and signal classes, respectively, plus Y_i and $\hat{Y}_i$ denote the true and predicted label probabilities, respectively. Training is conducted over 22 epochs with a batch size of 500 samples. The final network output is a probability vector $\hat{Y}$ of dimension 1×2, given by $(\mathcal{P}_{\text{sig}}, \mathcal{P}_{\text{bkg}})$, where $\mathcal{P} \in [0, 1]$. An event is classified as signal if $\mathcal{P}_{\text{sig}} > 0.5$ (equivalently, $\mathcal{P}_{\text{bkg}} < 0.5$) and as background if $\mathcal{P}_{\text{sig}} < 0.5$ (i.e., $\mathcal{P}_{\text{bkg}} > 0.5$). Once training is complete, the model performance is evaluated on a completely unseen test dataset to assess its generalization capability.

The effectiveness of each network in distinguishing signal from background events depends on how well it captures and disentangles the underlying physical features encoded in complex kinematic distributions and color flow patterns. This can be quantified using different metrics such as the Receiver Operating Characteristic (ROC) curves and the Area Under the Curve (AUC), which measures the model ability to correctly differentiate between signal and background events across varying decision thresholds. Figure 12, left plot, shows the ROC curves for the Mono-H and -Z analyses. The mono-H analysis has a better performance, with AUC = 0.985, over the mono-Z analysis, which has AUC = 0.973. The signal significances as a function of the network output are computed using
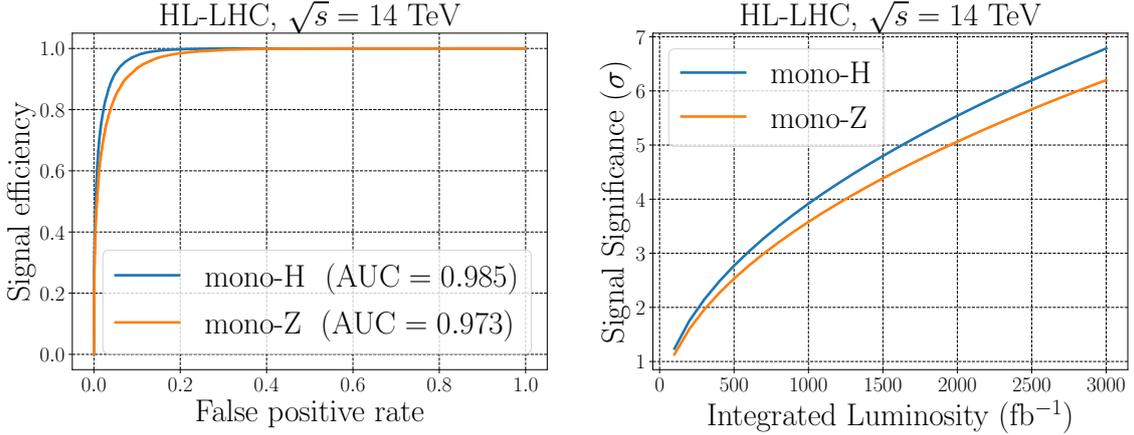
Figure 12: Left: ROC curves for the mono-H analysis in blue and the mono-Z analysis in orange. Right: signal significance at the (HL-)LHC for an integrated luminosity ranging from 100 to 3000fb^{-1}.

the formula [94, 95]:

$$\sigma = \left[2 \left((N_s + N_b) \ln \frac{(N_s + N_b)(N_b + \sigma_b^2)}{N_b^2 + (N_s + N_b)\sigma_b^2} - \frac{N_b^2}{\sigma_b^2} \ln(1 + \frac{\sigma_b^2 N_s}{N_b(N_b + \sigma_b^2)}) \right) \right]^{1/2} , \qquad (10)$$

with N_s, N_b being the number of signal and background events, respectively, and σ_b parameterizing the systematic uncertainty on the latter which is considered as 10%. The right plot of Figure 12 displays the expected signal significance as a function of the integrated luminosity. At a projected luminosity of $\mathcal{L} = 3000$ fb^{-1}, the mono-Z channel reaches a significance of 6.1σ, while the mono-H channel attains 6.8σ. These results highlight the potential of probing joint anomalies linking the existing LHC excesses, as discussed in Section 3, with the mono-H and -Z channels, by the end of the HL-LHC run.

6   Results and Discussion

Through an extensive scan of the NMSSM parameter space, aimed at accommodating all current theoretical and experimental constraints as well as various observed anomalies, we have identified a "golden region" of it that satisfies all such requirements. Specifically, leveraging this result, we have demonstrated that this region can naturally account for the observed 95 GeV and 650 GeV excesses while also providing a plausible interpretation of the reported EWino excess, including being consistent with the most recent measurement of $(g-2)_\mu$. Over such a portion of the scanned parameter space, we have found agreement between theoretical predictions and experimental data at the 2σ level.

This result has been aided by the use of a dedicated DL-informer parameter scanner. Initially, the DL scan explored broad parameter ranges, which were subsequently narrowed to accommodate the combined constraints and anomalies. The final accepted points constitute approximately 10% of the collected points from the narrow-range scan. Specifically, 1179 points were collected from this narrower scan, all of which satisfy all bounds and anomalies except for the 95 GeV and 650 GeV excesses. Imposing further the requirements for the 95 GeV and 650 GeV anomalies eliminates 90.33% of the collected points. The ranges of the twelve scanned parameters for the remaining allowed points are summarized in Table 5.

25

$\tan\beta$	λ	κ	A_λ
[4.15, 4.68]	[0.23, 0.36]	[0.16, 0.23]	[531.91, 620.72]
A_κ	μ_{eff}	M_1	M_2
[−297.87, −202.34]	[128.11, 151.25]	[559.75, 2988.48]	[768.88, 3971.61]
M_3	A_t	M_{Q_3}	M_{U_3}
[831.99, 4730.45]	[−4999.07, −3882.34]	[957.64, 4385.60]	[944.90, 4667.27]

Table 5: Allowed ranges of the scanned parameters that satisfy all the stated anomalies and constraints within the 2σ level.

To further quantify our results, we have evaluated the χ^2 values for the final set of accepted points, incorporating all the constraints and anomalies discussed above. The combined χ^2 is calculated using the standard expression: $\chi^2 = \sum_i \frac{(\text{observed}_i - \text{expected}_i)^2}{\sigma_i^2}$, where σ_i denotes the uncertainty associated with the observed measurement.

$\tan\beta$	λ	κ	A_λ	A_κ	μ_{eff}
4.61	0.32	0.21	620.72	−212.41	133.75
M_1	M_2	M_3	A_t	M_{Q_3}	M_{U_3}
2169.49	1574.89	1932.62	−4967.71	4280.64	3529.45

Table 6: Values of the scanned parameters corresponding to the best-fit point, which yields a minimum χ^2 value of 1.85. This point is illustrated by a blue star in Figure 13.

Figure 13 illustrates the distribution of the final accepted points across the scanned parameter space, with the color bar indicating the corresponding χ^2 values. The blue star marks the point with the minimum χ^2, for which detailed information is provided in Table 6. According to our findings, this point is crucial and serves as a golden benchmark for experimental groups to pursue.

7 Conclusions

In this paper, we have tested the possibility that an attractive realization of SUSY, the so-called NMSSM, can be the theoretical scenario underpinning some (current) experimental data that have revealed themselves anomalous to a reasonable degree of confidence. This choice has clear theoretical motivations. On the one hand, being a SUSY construct, the NMSSM can resolve the hierarchy problem of the SM while simultaneously generating the Higgs potential dynamically, predicting a low mass Higgs state, enabling coupling unification as well as providing a viable DM candidate. On the other hand, the NMSSM does not suffer from the 'little hierarchy' and μ problems of the minimal SUSY realization, the MSSM, while simultaneously embedding enlarged Higgs and sparticle sectors, including a different DM candidate.

Armed with these features, we have exploited the NMSSM parameter space to interpret several excesses reported at the LHC, namely the 95 GeV and 650 GeV anomalies, as well as the so-called EWino excess observed by ATLAS and CMS. To this end, we have carried out an extensive scan of the NMSSM parameter space under the requirement that all theoretical and experimental constraints be satisfied, including the most recent measurement
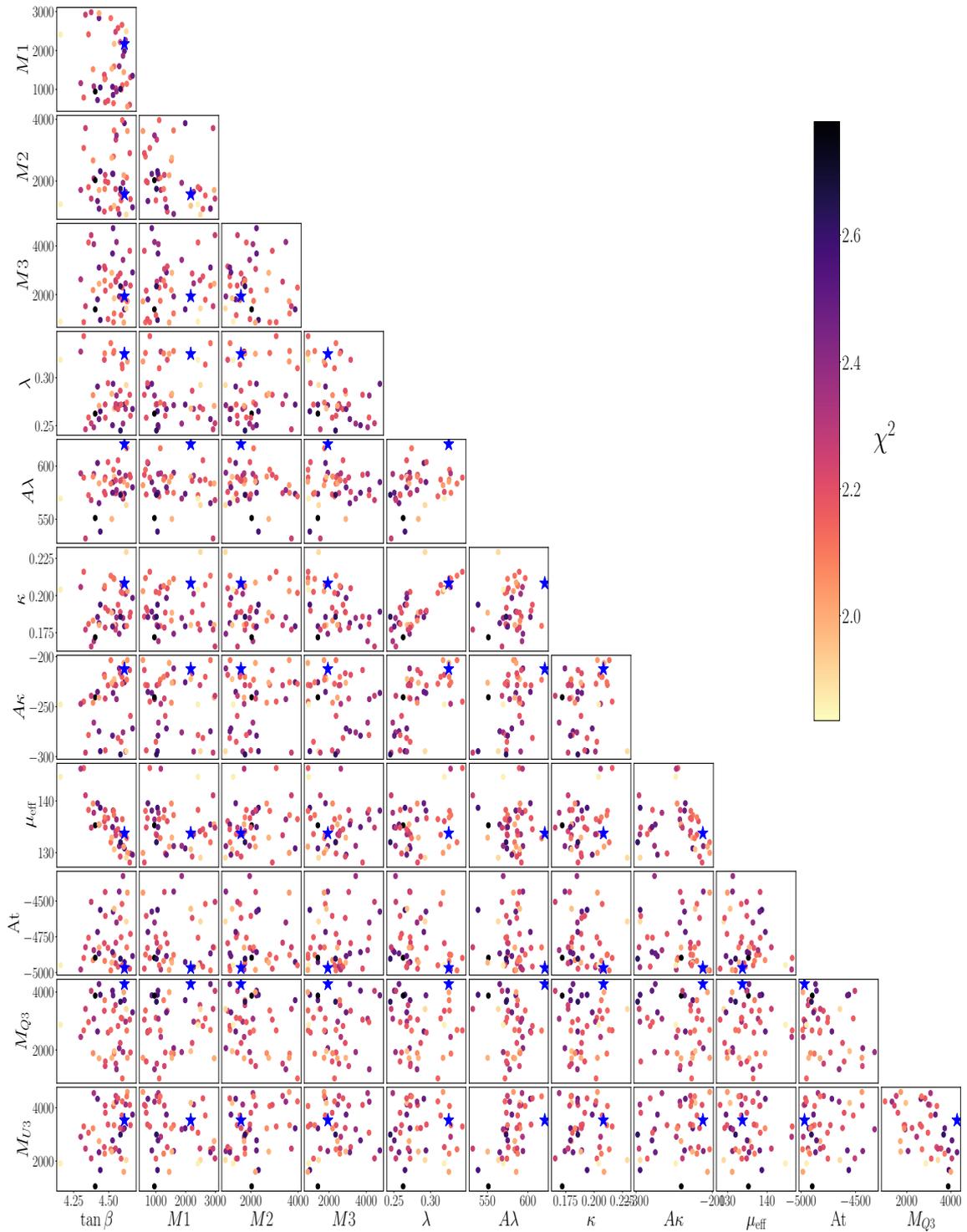
Figure 13: Ranges of the scanned variables in this analysis, with the color bar indicating the combined χ^2 values. The minimum χ^2 value is 1.85, marked by a blue star.

of the anomalous magnetic moment of the muon, $(g-2)_\mu$. This analysis has enabled us to identify a "golden region" of parameter space that simultaneously accommodates all such requirements.

A crucial ingredient in achieving these results has been the use of DL assisted scans.

The NMSSM parameter space is intrinsically high-dimensional and tightly constrained by a set of correlated theoretical and experimental requirements, which makes exhaustive scans computationally prohibitive. Traditional random or grid-based sampling methods are typically inefficient in capturing finely tuned regions where several constraints overlap. In contrast, DL assisted approaches can efficiently learn the structure of the viable parameter space, guiding the exploration toward regions of phenomenological relevance while avoiding redundant sampling of disallowed configurations. This strategy has proven vital in uncovering the aforementioned golden region where multiple anomalies and constraints converge. In this sense, DL not only accelerates the scanning procedure but also improves its precision, enabling us to probe regions that would otherwise remain inaccessible to conventional techniques.

Through this approach, we have established agreement between theoretical predictions and experimental observations at the 2σ level across sizeable portions of the NMSSM parameter space. Building on this result, we predict that, within these regions, mono-H (with $H = h_{\mathrm{SM}}$) and mono-Z processes constitute promising channels for testing our hypothesis at the HL-LHC. Their distinctive kinematic features would provide a unique hallmark of such a non-minimal SUSY framework, offering a concrete path toward experimental validation in future collider searches.

Taken together, our results highlight the NMSSM as a theoretically robust and phenomenologically viable framework capable of addressing both longstanding theoretical challenges of the SM and a range of current experimental anomalies, while also demonstrating the power of modern DL assisted methodologies in tackling the complexity of high-dimensional SUSY parameter spaces.

Acknowledgments

Dedication

We would like to dedicate this paper to the late Prof. Ulrich Ellwanger, one of the most passionate advocates of the NMSSM and a truly inspiring figure for us all.

References

[1] S. Moretti and S. Khalil, *Supersymmetry Beyond Minimality: From Theory to Experiment*. CRC Press, 2019.

[2] U. Ellwanger, C. Hugonie and A. M. Teixeira, *The Next-to-Minimal Supersymmetric Standard Model*, *Phys. Rept.* **496** (2010) 1 [0910.1785].

[3] M. Maniatis, *The Next-to-Minimal Supersymmetric extension of the Standard Model reviewed*, *Int. J. Mod. Phys. A* **25** (2010) 3505 [0906.0777].

[4] U. Ellwanger, J. F. Gunion and C. Hugonie, *NMHDECAY: A Fortran code for the Higgs masses, couplings and decay widths in the NMSSM*, *JHEP* **02** (2005) 066 [hep-ph/0406215].

[5] U. Ellwanger and C. Hugonie, *NMHDECAY 2.0: An Updated program for sparticle masses, Higgs masses, couplings and decay widths in the NMSSM*, *Comput. Phys. Commun.* **175** (2006) 290 [hep-ph/0508022].

[6] U. Ellwanger and C. Hugonie, *NMSPEC: A Fortran code for the sparticle and Higgs masses in the NMSSM with GUT scale boundary conditions*, *Comput. Phys. Commun.* **177** (2007) 399 [hep-ph/0612134].

[7] B. C. Allanach et al., *SUSY Les Houches Accord 2*, *Comput. Phys. Commun.* **180** (2009) 8 [0801.0045].

[8] A. Djouadi et al., *Benchmark scenarios for the NMSSM*, *JHEP* **07** (2008) 002 [0801.4321].

[9] CMS collaboration, *Search for new resonances in the diphoton final state in the mass range between 80 and 115 GeV in pp collisions at $\sqrt{s} = 8$ TeV*, .

[10] CMS collaboration, A. M. Sirunyan et al., *Search for a standard model-like Higgs boson in the mass range between 70 and 110 GeV in the diphoton final state in proton-proton collisions at $\sqrt{s} = 8$ and 13 TeV*, *Phys. Lett. B* **793** (2019) 320 [1811.08459].

[11] CMS collaboration, *Search for a standard model-like Higgs boson in the mass range between 70 and 110 GeV in the diphoton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*, .

[12] ATLAS collaboration, *Search for diphoton resonances in the 66 to 110 GeV mass range using 140 fb^{-1} of 13 TeV pp collisions collected with the ATLAS detector*, .

[13] CMS collaboration, A. Tumasyan et al., *Search for a new resonance decaying into two spin-0 bosons in a final state with two photons and two bottom quarks in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **05** (2024) 316 [2310.01643].

[14] CMS collaboration, A. Tumasyan et al., *Search for a heavy Higgs boson decaying into two lighter Higgs bosons in the $\tau\tau bb$ final state at 13 TeV*, *JHEP* **11** (2021) 057 [2106.10361].

[15] CMS collaboration, A. Hayrapetyan et al., *Search for the nonresonant and resonant production of a Higgs boson in association with an additional scalar boson in the $\gamma\gamma\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*, 2506.23012.

[16] S. Banik, A. Crivellin, S. Iguro and T. Kitahara, *Asymmetric di-Higgs signals of the next-to-minimal 2HDM with a U(1) symmetry*, *Phys. Rev. D* **108** (2023) 075011 [2303.11351].

[17] U. Ellwanger and C. Hugonie, *Additional Higgs Bosons near 95 and 650 GeV in the NMSSM*, *Eur. Phys. J. C* **83** (2023) 1138 [2309.07838].

[18] U. Ellwanger, C. Hugonie, S. F. King and S. Moretti, *NMSSM explanation for excesses in the search for neutralinos and charginos and a 95 GeV Higgs boson*, *Eur. Phys. J. C* **84** (2024) 788 [2404.19338].

[19] A. Le Yaouanc and F. Richard, *New resonances at LHC*, in *3rd ECFA workshop on e+e- Higgs, Electroweak and Top Factories*, 6, 2025, 2506.09490.

[20] ATLAS collaboration, G. Aad et al., *Searches for electroweak production of supersymmetric particles with compressed mass spectra in $\sqrt{s}$ = 13 TeV pp collisions with the ATLAS detector*, *Phys. Rev. D* **101** (2020) 052005 [1911.12606].

[21] ATLAS collaboration, G. Aad et al., *Search for chargino–neutralino pair production in final states with three leptons and missing transverse momentum in $\sqrt{s}$ = 13 TeV pp collisions with the ATLAS detector*, *Eur. Phys. J. C* **81** (2021) 1118 [2106.01676].

[22] CMS collaboration, A. Tumasyan et al., *Search for supersymmetry in final states with two or three soft leptons and missing transverse momentum in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, *JHEP* **04** (2022) 091 [2111.06296].

[23] CMS collaboration, A. Tumasyan et al., *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 [2107.13021].

[24] Muon g-2 collaboration, G. W. Bennett et al., *Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL*, *Phys. Rev. D* **73** (2006) 072003 [hep-ex/0602035].

[25] A. Hammad and R. Ramos, *DLScanner: A parameter space scanner package assisted by deep learning methods*, *Comput. Phys. Commun.* **314** (2025) 109659 [2412.19675].

[26] G. P. Lepage, *Adaptive multidimensional integration: VEGAS enhanced*, *J. Comput. Phys.* **439** (2021) 110386 [2009.05112].

[27] P. Fayet, *Supergauge Invariant Extension of the Higgs Mechanism and a Model for the electron and Its Neutrino*, *Nucl. Phys. B* **90** (1975) 104.

[28] M. Dine, W. Fischler and M. Srednicki, *A Simple Solution to the Strong CP Problem with a Harmless Axion*, *Phys. Lett. B* **104** (1981) 199.

[29] H. P. Nilles, M. Srednicki and D. Wyler, *Weak Interaction Breakdown Induced by Supergravity*, *Phys. Lett. B* **120** (1983) 346.

[30] J. M. Frere, D. R. T. Jones and S. Raby, *Fermion Masses and Induction of the Weak Scale by Supergravity*, *Nucl. Phys. B* **222** (1983) 11.

[31] J. P. Derendinger and C. A. Savoy, *Quantum Effects and SU(2) x U(1) Breaking in Supergravity Gauge Theories*, *Nucl. Phys. B* **237** (1984) 307.

[32] A. Dedes, C. Hugonie, S. Moretti and K. Tamvakis, *Phenomenology of a new minimal supersymmetric extension of the standard model*, *Phys. Rev. D* **63** (2001) 055009 [hep-ph/0009125].

[33] C. Panagiotakopoulos and A. Pilaftsis, *Higgs scalars in the minimal nonminimal supersymmetric standard model*, *Phys. Rev. D* **63** (2001) 055003 [hep-ph/0008268].

[34] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440.

[35] R. D. Peccei and H. R. Quinn, *Constraints Imposed by CP Conservation in the Presence of Instantons*, *Phys. Rev. D* **16** (1977) 1791.

[36] Y. B. Zeldovich, I. Y. Kobzarev and L. B. Okun, *Cosmological Consequences of the Spontaneous Breakdown of Discrete Symmetry*, *Zh. Eksp. Teor. Fiz.* **67** (1974) 3.

[37] C. Panagiotakopoulos and K. Tamvakis, *Stabilized NMSSM without domain walls*, *Phys. Lett. B* **446** (1999) 224 [`hep-ph/9809475`].

[38] U. Ellwanger, C. C. Jean-Louis and A. M. Teixeira, *Phenomenology of the General NMSSM with Gauge Mediated Supersymmetry Breaking*, *JHEP* **05** (2008) 044 [`0803.2962`].

[39] F. Domingo, *A New Tool for the study of the CP-violating NMSSM*, *JHEP* **06** (2015) 052 [`1503.07087`].

[40] D. Das, U. Ellwanger and A. M. Teixeira, *NMSDECAY: A Fortran Code for Supersymmetric Particle Decays in the Next-to-Minimal Supersymmetric Standard Model*, *Comput. Phys. Commun.* **183** (2012) 774 [`1106.5633`].

[41] G. Alguero, G. Belanger, F. Boudjema, S. Chakraborti, A. Goudelis, S. Kraml et al., *micrOMEGAs 6.0: N-component dark matter*, *Comput. Phys. Commun.* **299** (2024) 109133 [`2312.14894`].

[42] G. Alguero, J. Heisig, C. K. Khosa, S. Kraml, S. Kulkarni, A. Lessa et al., *Constraining new physics with SModelS version 2*, *JHEP* **08** (2022) 068 [`2112.00769`].

[43] H. Bahl, T. Biekötter, S. Heinemeyer, C. Li, S. Paasch, G. Weiglein et al., *HiggsTools: BSM scalar phenomenology with new versions of HiggsBounds and HiggsSignals*, *Comput. Phys. Commun.* **291** (2023) 108803 [`2210.09332`].

[44] P. Bechtle, S. Heinemeyer, T. Klingl, T. Stefaniak, G. Weiglein and J. Wittbrodt, *HiggsSignals-2: Probing new physics with precision Higgs measurements in the LHC 13 TeV era*, *Eur. Phys. J. C* **81** (2021) 145 [`2012.09197`].

[45] P. Bechtle, D. Dercks, S. Heinemeyer, T. Klingl, T. Stefaniak, G. Weiglein et al., *HiggsBounds-5: Testing Higgs Sectors in the LHC 13 TeV Era*, *Eur. Phys. J. C* **80** (2020) 1211 [`2006.06007`].

[46] PLANCK collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [`1807.06209`].

[47] LZ collaboration, J. Aalbers et al., *First Dark Matter Search Results from the LUX-ZEPLIN (LZ) Experiment*, *Phys. Rev. Lett.* **131** (2023) 041002 [`2207.03764`].

[48] XENON collaboration, E. Aprile et al., *Dark Matter Search Results from a One Ton-Year Exposure of XENON1T*, *Phys. Rev. Lett.* **121** (2018) 111302 [`1805.12562`].

[49] XENON collaboration, E. Aprile et al., *Projected WIMP sensitivity of the XENONnT dark matter experiment*, *JCAP* **11** (2020) 031 [`2007.08796`].

[50] LZ collaboration, D. S. Akerib et al., *Projected WIMP sensitivity of the LUX-ZEPLIN dark matter experiment*, *Phys. Rev. D* **101** (2020) 052002 [1802.06039].

[51] FERMI-LAT collaboration, M. Ackermann et al., *Dark Matter Constraints from Observations of 25 Milky Way Satellite Galaxies with the Fermi Large Area Telescope*, *Phys. Rev. D* **89** (2014) 042001 [1310.0828].

[52] HFLAV collaboration, Y. S. Amhis et al., *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018*, *Eur. Phys. J. C* **81** (2021) 226 [1909.12524].

[53] J. Baglio, R. Gröber, M. Mühlleitner, D. T. Nhung, H. Rzehak, M. Spira et al., *NMSSMCALC: A Program Package for the Calculation of Loop-Corrected Higgs Boson Masses and Decay Widths in the (Complex) NMSSM*, *Comput. Phys. Commun.* **185** (2014) 3372 [1312.4788].

[54] GFITTER GROUP collaboration, M. Baak, J. Cúth, J. Haller, A. Hoecker, R. Kogler, K. Mönig et al., *The global electroweak fit at NNLO and prospects for the LHC and ILC*, *Eur. Phys. J. C* **74** (2014) 3046 [1407.3792].

[55] V. Barger, P. Langacker, H.-S. Lee and G. Shaughnessy, *Higgs Sector in Extensions of the MSSM*, *Phys. Rev. D* **73** (2006) 115010 [hep-ph/0603247].

[56] F. Ambrogi, S. Kraml, S. Kulkarni, U. Laa, A. Lessa, V. Magerl et al., *SModelS v1.1 user manual: Improving simplified model constraints with efficiency maps*, *Comput. Phys. Commun.* **227** (2018) 72 [1701.06586].

[57] ATLAS collaboration, G. Aad et al., *Search for squarks and gluinos in final states with jets and missing transverse momentum using 139 fb^{-1} of $\sqrt{s}$ =13 TeV pp collision data with the ATLAS detector*, *JHEP* **02** (2021) 143 [2010.14293].

[58] ATLAS collaboration, M. Aaboud et al., *Search for squarks and gluinos in final states with jets and missing transverse momentum using 36 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data with the ATLAS detector*, *Phys. Rev. D* **97** (2018) 112001 [1712.02332].

[59] CMS collaboration, T. C. Collaboration et al., *Search for supersymmetry in proton-proton collisions at 13 TeV in final states with jets and missing transverse momentum*, *JHEP* **10** (2019) 244 [1908.04722].

[60] CMS collaboration, A. M. Sirunyan et al., *Search for top squark production in fully-hadronic final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* **104** (2021) 052001 [2103.01290].

[61] LEP WORKING GROUP FOR HIGGS BOSON SEARCHES, ALEPH, DELPHI, L3, OPAL collaboration, R. Barate et al., *Search for the standard model Higgs boson at LEP*, *Phys. Lett. B* **565** (2003) 61 [hep-ex/0306033].

[62] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [1405.0301].

[63] T. Biekötter, S. Heinemeyer and G. Weiglein, *95.4 GeV diphoton excess at ATLAS and CMS*, *Phys. Rev. D* **109** (2024) 035005 [2306.03889].

[64] J. Cao, X. Guo, Y. He, P. Wu and Y. Zhang, *Diphoton signal of the light Higgs boson in natural NMSSM*, *Phys. Rev. D* **95** (2017) 116001 [`1612.08522`].

[65] CMS collaboration, A. Tumasyan et al., *Searches for additional Higgs bosons and for vector leptoquarks in $\tau\tau$ final states in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, *JHEP* **07** (2023) 073 [`2208.02717`].

[66] CMS collaboration, *Combined search for electroweak production of winos, binos, higgsinos, and sleptons in proton-proton collisions at sqrts = 13 TeV*, .

[67] ATLAS collaboration, G. Aad et al., *Search for new phenomena in events with an energetic jet and missing transverse momentum in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector*, *Phys. Rev. D* **103** (2021) 112006 [`2102.10874`].

[68] M. Chakraborti, S. Heinemeyer and I. Saha, *Consistent excesses in the search for wino/bino vs. Higgsino dark matter*, *Eur. Phys. J. C* **84** (2024) 812 [`2403.14759`].

[69] MUON G-2 collaboration, B. Abi et al., *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*, *Phys. Rev. Lett.* **126** (2021) 141801 [`2104.03281`].

[70] MUON G-2 collaboration, T. Albahri et al., *Magnetic-field measurement and analysis for the Muon $g - 2$ Experiment at Fermilab*, *Phys. Rev. A* **103** (2021) 042208 [`2104.03201`].

[71] MUON G-2 collaboration, T. Albahri et al., *Beam dynamics corrections to the Run-1 measurement of the muon anomalous magnetic moment at Fermilab*, *Phys. Rev. Accel. Beams* **24** (2021) 044002 [`2104.03240`].

[72] T. Aoyama et al., *The anomalous magnetic moment of the muon in the Standard Model*, *Phys. Rept.* **887** (2020) 1 [`2006.04822`].

[73] S. Borsanyi et al., *Leading hadronic contribution to the muon magnetic moment from lattice QCD*, *Nature* **593** (2021) 51 [`2002.12347`].

[74] R. Aliberti et al., *The anomalous magnetic moment of the muon in the Standard Model: an update*, `2505.21476`.

[75] D. Agin, B. Fuks, M. D. Goodsell and T. Murphy, *A joint explanation for the soft lepton and monojet LHC excesses in the wino-bino model*, `2506.21676`.

[76] B. Fuks, M. D. Goodsell and T. Murphy, *Monojets from compressed weak frustrated dark matter*, *Phys. Rev. D* **111** (2025) 055010 [`2409.03014`].

[77] D. Agin, B. Fuks, M. D. Goodsell and T. Murphy, *Monojets reveal overlapping excesses for light compressed higgsinos*, *Phys. Lett. B* **853** (2024) 138597 [`2311.17149`].

[78] F. Maltoni, K. Paul, T. Stelzer and S. Willenbrock, *Color Flow Decomposition of QCD Amplitudes*, *Phys. Rev. D* **67** (2003) 014026 [`hep-ph/0209271`].

[79] K. Hagiwara and Y. Takaesu, *Generating QCD amplitudes in the color-flow basis with MadGraph*, *Eur. Phys. J. C* **71** (2011) 1668 [`1010.0748`].

[80] W. Kilian, T. Ohl, J. Reuter and C. Speckner, *QCD in the Color-Flow Representation*, *JHEP* **10** (2012) 022 [`1206.3700`].

[81] J. H. Kim, M. Kim, K. Kong, K. T. Matchev and M. Park, *Portraying Double Higgs at the Large Hadron Collider*, *JHEP* **09** (2019) 047 [1904.08549].

[82] A. Hammad, S. Khalil and S. Moretti, *Search for mono-Higgs signals in* $b\bar{b}$ *final states using deep neural networks*, *Phys. Rev. D* **107** (2023) 075027 [2208.10133].

[83] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H. S. Shao and M. Zaro, *The automation of next-to-leading order electroweak calculations*, *JHEP* **07** (2018) 185 [1804.10017].

[84] C. Bierlich et al., *A comprehensive guide to the physics and usage of PYTHIA 8.3*, *SciPost Phys. Codeb.* **2022** (2022) 8 [2203.11601].

[85] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J. C* **72** (2012) 1896 [1111.6097].

[86] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [0802.1189].

[87] DELPHES 3 collaboration, J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens et al., *DELPHES 3, A modular framework for fast simulation of a generic collider experiment*, *JHEP* **02** (2014) 057 [1307.6346].

[88] M. Fey and J. E. Lenssen, *Fast graph representation learning with pytorch geometric*, *arXiv preprint arXiv:1903.02428* (2019) .

[89] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan et al., *Pytorch: An imperative style, high-performance deep learning library*, *Advances in neural information processing systems* **32** (2019) .

[90] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel et al., *Scikit-learn: Machine learning in python*, *the Journal of machine Learning research* **12** (2011) 2825.

[91] M. Bachlechner, T. Birkenfeld, P. Soldin, A. Stahl and C. Wiebusch, *Partition pooling for convolutional graph network applications in particle physics*, *JINST* **17** (2022) P10004 [2208.05952].

[92] W. Esmail, A. Hammad and S. Moretti, *Sharpening the* $A \to Z^{(*)}h$ *signature of the Type-II 2HDM at the LHC through advanced Machine Learning*, *JHEP* **11** (2023) 020 [2305.13781].

[93] R. Sahu, *CapsLorentzNet: Integrating Physics Inspired Features with Graph Convolution*, 2403.11826.

[94] LHC DARK MATTER WORKING GROUP collaboration, T. Abe et al., *LHC Dark Matter Working Group: Next-generation spin-0 dark matter models*, *Phys. Dark Univ.* **27** (2020) 100351 [1810.09420].

[95] S. Antusch, E. Cazzato, O. Fischer, A. Hammad and K. Wang, *Lepton Flavor Violating Dilepton Dijet Signatures from Sterile Neutrinos at Proton Colliders*, *JHEP* **10** (2018) 067 [1805.11400].