

Fermilab

Machine-Learning Analysis of Radiative Decays to Dark Matter at the LHC

FERMILAB-PUB-24-0791-V

arXiv:2410.13799

This manuscript has been authored by Fermi Research Alliance, LLC
under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy,
Office of Science, Office of High Energy Physics.

Machine-Learning Analysis of Radiative Decays to Dark Matter at the LHC

ERNESTO ARGANDA^{1*}, MARCELA CARENA^{2,3,4†}, MARTÍN DE LOS RÍOS^{1,5‡}, ANDRÉS D. PEREZ^{1§}, DUNCAN ROCHA^{2,3¶}, ROSA M. SANDÁ SEOANE^{1||} AND CARLOS E. M. WAGNER^{3,4,6**}

¹*Departamento de Física Teórica and Instituto de Física Teórica UAM-CSIC,
Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain*

²*Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA*

³*Enrico Fermi Institute, Physics Department, University of Chicago, Chicago, IL 60637, USA*

⁴*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*

⁵*SISSA - International School for Advanced Studies, Via Bonomea 265, 34136 Trieste, Italy*

⁶*HEP Division, Argonne National Laboratory, 9700 Cass Ave., Argonne, IL 60439, USA*

Abstract

The search for weakly interacting matter particles (WIMPs) is one of the main objectives of the High Luminosity Large Hadron Collider (HL-LHC). In this work we use Machine-Learning (ML) techniques to explore WIMP radiative decays into a Dark Matter (DM) candidate in a supersymmetric framework. The minimal supersymmetric WIMP sector includes the lightest neutralino that can provide the observed DM relic density through its co-annihilation with the second lightest neutralino and lightest chargino. Moreover, the direct DM detection cross section rates fulfill current experimental bounds and provide discovery targets for the same region of model parameters in which the radiative decay of the second lightest neutralino into a photon and the lightest neutralino is enhanced. This strongly motivates the search for radiatively decaying neutralinos which, however, suffers from strong backgrounds. We investigate the LHC reach in the search for these radiatively decaying particles by means of cut-based and ML methods and estimate its discovery potential in this well-motivated, new physics scenario.

*ernesto.arganda@uam.es

†carena@fnal.gov

‡mdelosri@sissa.it

§andresd.perez@uam.es

¶drocha@uchicago.edu

||rosa.sanda@uam.es

**cwagner@uchicago.edu

Contents

1	Introduction	1
2	Theoretical Framework	3
3	Collider Analysis	6
3.1	Event Simulation and Characterization	6
3.2	Analysis Strategies	10
3.2.1	Cut-Based Strategy	10
3.2.2	Machine-Learning Strategy	11
4	Results	13
5	Conclusions	17
A	Kinematic Distributions	19
B	Review on MLL and BL Formulae	22
C	Results with $E_T^{\text{miss}} > 200$ GeV	23

1 Introduction

The origin of the scale of electroweak (EW) symmetry breaking and the identity of the constituent particles that make up Dark Matter (DM) are two of the big questions in particle physics. Beyond the Standard Model (BSM) theories with a weakly interacting massive particle (WIMP) [1, 2] sector provide a natural framework that connects these two important questions and has been the subject of intense exploration at colliders and direct detection experiments for the past few decades. The exploration of a WIMP sector as a solution to the DM mystery, together with a deeper understanding of the Higgs boson properties [3], are the two main objectives of the Large Hadron Collider (LHC) and its upcoming High Luminosity (HL) run. The HL-LHC will provide an unprecedented amount of data, opening new windows into DM searches, yet posing new challenges for disentangling potential new physics signals from a plethora of background signals.

Supersymmetry [4–9] has been and remains a leading BSM theory to connect the two big questions: i) it readily provides a DM candidate by rendering the lightest supersymmetric particle (LSP) stable in the presence of R -Parity [10–14]; ii) it relates the weak scale to the soft supersymmetry breaking scale by naturally explaining the smallness of the measured Higgs boson mass with respect to the Planck scale of new physics. Current searches at the LHC [15–24] provide significant constraints on the presence of weak scale supersymmetric partners of the colored Standard Model (SM) particles. WIMPs, instead, may be lighter [24–26], and due to their small production cross sections, their potential weak scale masses can be efficiently probed at the HL-LHC runs.

There is currently only indirect information about the characteristic scale of potential supersymmetric weakly interacting particles. This comes primarily from cosmological requirements to produce the observed dark matter relic density [27]. A proper contribution to the DM relic density may be achieved in the co-annihilation region [28–35], where the mass of the lightest neutralino is close to other weakly interacting particles with higher annihilation rate. A particularly interesting example that arises somewhat naturally in supersymmetric scenarios is when the LSP is the superpartner of the hypercharge gauge boson (bino), $\tilde{\chi}_1^0$, and these additional particles are the second lightest neutralino, $\tilde{\chi}_2^0$, and the lightest chargino, $\tilde{\chi}_1^\pm$, which in the Minimal Supersymmetric Standard Model (MSSM) [4–9] can be, for instance, the superpartners of the Z and W gauge bosons (winos), respectively. The mass of the LSP is constrained by direct DM detection [36–39]. It turns out that if the lightest neutralino is bino-like and the bino mass parameter, M_1 , is of the opposite sign as the μ parameter (related to the higgsino mass), the direct detection cross section is suppressed, opening the parameter space of possible LSP mass values [40–47].

Moreover, sufficiently light supersymmetric particles could play a role in providing sizable additional contributions to the muon anomalous magnetic moment $a_\mu - 2$. It has been shown in the co-annihilation scenario that demanding the wino mass parameter M_2 to be of the same sign as μ yields a positive contribution to $a_\mu - 2$, thereby alleviating the apparent discrepancy between experimental results and theoretical calculations [48–69]. Combining both conditions: $\mu \times M_1 < 0$ and $\mu \times M_2 > 0$, it happens that the second lightest neutralino tends to decay radiatively into the lightest neutralino and a photon [70–72] in a significant way. Therefore, the searches for radiative decays of neutralinos are quite intriguing and well motivated [73]; and should be pursued at the LHC (A similar enhanced radiative decay behavior has been pointed out in the next-to-minimal-MSSM, see Ref. [74]).

Electroweakinos, as they are generically called the weakly interacting supersymmetric partners of the Higgs and electroweak gauge bosons, are difficult particles to detect at the LHC. This is specially the case when they decay radiatively and to a dark matter particle, due to the large Standard Model (SM) backgrounds. A naive analysis, based on simple cuts of the main kinematic variables, generally fails to show discovery potential even at the high luminosity LHC. These analyses, however, neglect the correlation between different kinematic variables, which are most accurately captured using machine-learning (ML) methods (for reviews of ML applied to high energy physics see, for instance, [75–91]).

In this article, we perform and compare cut-based and ML analyses of the search for radiative neutralino decays at the LHC, for a center-of-mass energy of $\sqrt{s} = 14$ TeV and a total integrated luminosity of $\mathcal{L} = 100 \text{ fb}^{-1}$. The analysis is a projection of the results that can be obtained in the early stages at the Run 3 of the LHC. Additionally, this choice for the luminosity saves computational resources in terms of Monte Carlo simulations. For the ML analysis, we estimate discovery reach both within the traditional Binned-Likelihood (BL) method [92] and within the Machine-Learned Likelihoods (MLL) method [93–96]. In the latter, we perform an unbinned fit of the likelihood function obtained from the ML output using Kernel Density Estimators (KDE) [97, 98]. We assume that all colored supersymmetric particles are heavy and beyond LHC reach, and we focus on the

production of winos, that in these MSSM scenarios are the second neutralino and the lightest chargino, and their radiative decay to the lightest neutralino and a photon. Due to the small mass difference between $\tilde{\chi}_1^0$ and $\tilde{\chi}_2^0/\tilde{\chi}_1^\pm$ required by co-annihilation, most of the signatures in the proposed signal process tend to be soft, so we consider the production of a highly energetic initial state radiation (ISR) jet in association with the electroweakino pair [73]. This increases the missing energy signature, which becomes useful in the search process. Contrary to the naive cut-and-count approach, the ML analysis shows an interesting discovery potential for electroweakinos at the LHC in the coming years.

This article is organized as follows: in Section 2 we present the theoretical framework, highlighting the supersymmetric parameter space under consideration. The LSP is a bino-like neutralino that is almost degenerate in mass with the wino-like neutralino $\tilde{\chi}_2^0$ and chargino $\tilde{\chi}_1^\pm$, and the radiative decay of $\tilde{\chi}_2^0$ is significant. Section 3 is dedicated to the collider analysis we conducted. First, we explain the event simulation and characterize the signal against the background. Then, we present the two approaches considered for the analysis: one based on rectangular cuts, optimized through a sequential cut selection, and a second approach using machine learning, considering both the BL method and the unbinned MLL method for the final significance estimation. The results of our collider analysis are shown in Section 4, comparing the potential of the different strategies applied. We reserve Section 5 for our conclusions. Appendix A is devoted to complementary kinematic distributions of signal and background features used in our analysis. Appendix B reviews the main formulae for the BL and the MLL methods. Finally, in Appendix C, we show the results considering a more stringent cut in the missing transverse energy, which is compatible with current LHC triggers, although this leads to fewer statistics and larger uncertainties as a drawback.

2 Theoretical Framework

As mentioned in the introduction, several works have shown that there exist viable regions of the MSSM parameter space that simultaneously reproduce the observed DM relic density of the universe and yield a solution to the apparent tension between theory and experiment in the value of the muon anomalous magnetic moment. Addressing both of these questions restricts the MSSM parameter space substantially. For example, a stable $\mathcal{O}(200)$ GeV bino-dominated neutralino produced via freeze-out would be too abundant and severely violate DM relic density constraints. The right abundance is obtained if we consider the “compressed region” of parameter space, where there are two species that are almost degenerate in mass and the lightest of the two is a DM candidate. This allows for coannihilation to occur between these two species in the early Universe, reducing the LSP relic density to be in the ballpark of the observed value. At the same time, the sleptons play an important role in the MSSM contributions to $g_\mu - 2$, and therefore they cannot be too massive. Resolution of the muon $g_\mu - 2$ apparent tension in the scenario of a bino-like LSP requires generically that the slepton masses $M_{\tilde{l}} \lesssim 1$ TeV ¹. This requirement on the slepton sector adds to

¹Throughout this article, we shall assume the value of the muon $g_\mu - 2$ obtained by computing the hadronic vacuum polarization contributions by dispersion relations of the electron-positron hadronic cross sections [99]. If the

the previously mentioned condition that $(M_2 \times \mu) > 0$. Finally, there exist direct detection limits that constrain neutralino relics scattering off nuclei via the exchange of a CP-even Higgs. This constraint is weakened for $(M_1 \times \mu) < 0$, where the spin-independent direct detection cross sections receive contributions from MSSM diagrams which partially cancel.

Motivated by the previous considerations, we are led to a region of parameter space that contains an interesting and yet unexploited, new collider signal. In this region, the LSP is a bino-like neutralino, $\tilde{\chi}_1^0$, which has a mass of several hundred GeV. Close in mass, but slightly heavier, are the next mass eigenstates of the electroweakinos, the second neutralino ($\tilde{\chi}_2^0$) and the lightest chargino ($\tilde{\chi}_1^\pm$). In addition, the sleptons must have masses of several hundred GeV, to fulfill current LHC bounds while still allowing for a possible contribution to $g_\mu - 2$. Lastly, the higgsino masses need to be well below the TeV scale, to avoid inducing a large bino-induced negative contribution to the muon $g_\mu - 2$ [47]. For this study, we fix $M_{\tilde{t}} = 600$ GeV and $|\mu| = 800$ GeV.

Typically, for the neutralino mass splitting larger than the observed Higgs mass, $m_H \sim 125$ GeV, $m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0} \gtrsim m_H$, the second neutralino decays via on-shell Higgs or electroweak gauge bosons. Decays of this type produce multilepton final states, which have been searched for extensively [101, 102, 104–112]. However, the proper relic density can only be achieved deeper into the compressed region, for $m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0} \sim 20 - 40$ GeV. Interestingly enough, both collaborations have seen some multilepton excesses in potential electroweakino production compatible with a compressed spectra and masses in the 200 – 300 GeV range [101–103]. There have been a few works trying to give a coherent explanation of these excesses (see for example [113, 114]). In [73] it was shown that in this region, due to the small mass differences between the second lightest and the lightest neutralino, the decays via off-shell electroweak gauge bosons become suppressed and give way to a radiative decay mode ($\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 + \gamma$). Because this new mode is a 1-loop process with virtual sleptons and electroweakinos, it scales differently with the neutralino mass splitting and it is sensitive to several parameters in the MSSM parameter space, in particular the slepton masses. As mentioned before, the value of the slepton masses plays an important role in the radiative decay branching ratio of the second lightest neutralino, and larger slepton mass values will reduce it. Note that the value of $\tan(\beta)$ does not play a relevant role on the radiative decay branching ratio.²

The region of parameter space with significant, loop induced radiative decay offers a new signature for electroweakino searches at the LHC, which, as mentioned before, is dominated by wino production, $pp \rightarrow \tilde{\chi}_2^0 \tilde{\chi}_1^\pm$. Current LHC searches that consider this production channel probe the compressed region by searching for multilepton final states. As proposed in [73] it may prove beneficial to consider a soft photon ($p_T \sim \mathcal{O}(50)$ GeV) as additional evidence for MSSM activity. In this paper we explore the experimental viability of searching for these photons offering insights into the

final value tends to be the one obtained by lattice methods (see for instance Ref. [99] and Ref. [100] for a discussion of this issue), the tension would be reduced and higher slepton masses than the ones assumed in this article would be preferred.

²The chosen value of slepton masses, $M_{\tilde{t}} = 600$ GeV, and $\tan(\beta) = 50$ reduce the reported tension between the experimental measurement of $g_\mu - 2$ and theoretical predictions based on data-driven approaches. However, if one increases the slepton masses or lowers the value of $\tan(\beta)$ one could easily minimize the supersymmetric contributions to $g_\mu - 2$, thereby disregarding the current apparent tension in this quantity.

BP	$m_{\tilde{\chi}_2^0}$ [GeV]	$(m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0})$ [GeV]	$\text{Br}(\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 \gamma)$	$\sigma(pp \rightarrow \tilde{\chi}_1^\pm \tilde{\chi}_2^0)$	$\frac{\sigma_{DD} \times \min(\Omega_{\tilde{\chi}_1^0}, \Omega_{\text{DM}})}{\sigma_{DD,95} \times \Omega_{\text{DM}}}$	$\Omega_{\tilde{\chi}_1^0} h^2$	$\Delta a_\mu^{\text{MSSM}}$ $\times 10^{10}$
1	200	34	15%	190 fb	0.34	2.0	22
2	200	19	37%	190 fb	0.47	0.12	22
3	200	10	73%	190 fb	0.09	0.02	22
4	250	37	15%	92 fb	0.74	1.2	21
5	250	22	36%	92 fb	0.96	0.12	21
6	250	13	67%	92 fb	0.19	0.03	21
7	300	39	16%	48 fb	1.3	0.8	20
8	300	24	36%	48 fb	1.6	0.12	20
9	300	15	62%	48 fb	0.36	0.04	20
10	350	41	17%	27 fb	2.2	0.60	19
11	350	26	35%	27 fb	2.5	0.12	19
12	350	17	58%	27 fb	1.0	0.04	19
13	400	43	16%	16 fb	3.2	0.47	17
14	400	27	32%	16 fb	3.6	0.12	17
15	400	18	52%	16 fb	1.7	0.05	17

Table 1: Table of benchmark points used in this analysis. For all points, $\mu = 800$ GeV, $\tan(\beta) = 50$, and $M_{\tilde{t}} = 600$ GeV, and $(M_1 \times \mu) < 0$. We set the squark soft masses to 2.5 TeV, and any additional superpartner masses to 1.5 TeV. Also, $m_{\tilde{\chi}_1^\pm} \approx m_{\tilde{\chi}_2^0}$ for all of the points. The points which produce the proper cosmological relic density are shaded in gray. Direct detection cross sections for points which produce relic $\tilde{\chi}_1^0$ abundance above Ω_{DM} are rescaled to estimate the local abundance of $\tilde{\chi}_1^0$. The preferred value of Δa_μ is 25 ± 5 . A few of the low-mass points are excluded by ATLAS or CMS [25, 101–103]. Points are chosen in a grid across the soft parameters M_2 and $M_2 - M_1$ which leads to a slightly uneven set of splittings as there is a variation of the dependence of physical parameters on soft parameters.

LHC reach potential in the radiative decay mode.

We consider 15 benchmark points (BPs) whose relevant parameters for collider phenomenology are shown in Table 1. We generate the mass spectra and decay rates for these points using `SuSpect` [115] and `SUSY-HIT` [116]. Cross sections were generated using `PROSPINO` [117], and the direct detection constraints and relic density were evaluated with `MicrOmegas` [118], including the recently published LUX-ZEPLIN results [119]. These benchmark points are chosen as representative points across the compressed region. The relic density is directly dependent on the mass splitting, and so not all the BPs necessarily produce the cosmological relic density, but a subset does (BPs 2, 5, 8, 11, and 14). Points predicting a relic density above the observed one are excluded within the MSSM but may be brought back into agreement by additional SM modifications, such as a late injection of entropy during the Big Bang or another species with which $\tilde{\chi}_1^0$ can coannihilate. For those points, we rescale the direct detection cross section by the predicted relic density to account for a modified local abundance, consistent with the observed one. A few of these BPs, with LSP below

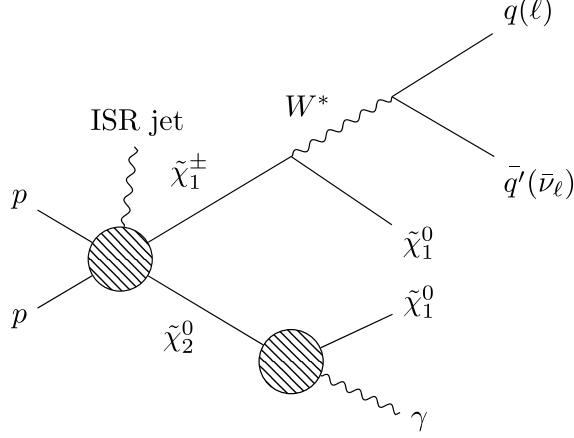


Figure 1: Feynman diagram of the production mode and final state considered.

250 GeV, are in principle excluded by previous searches. In particular, the benchmarks with lowest neutralinos masses would be excluded by ATLAS and CMS multilepton searches (e.g. [101–103]). However, the ATLAS and CMS bounds applied to our scenario are weakened by the change in the assumptions. For example, CMS assumes both 100% branching ratios of $(\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 + h^*)$ and higher squark masses, which are modified significantly in our BPs. This affects both the rates of the multilepton final states and the production cross section for $(pp \rightarrow \tilde{\chi}_2^0 \tilde{\chi}_1^\pm)$. For the LHC bounds depicted in Figure 5, we use the bounds placed in [114], which provide a more recent update to the bounds in [73] for the same region in parameter space.

Due to the large non-standard Higgs boson masses allowed by LHC searches for large values of $\tan\beta$ [120, 121], there is only a small suppression of the direct DM detection rate induced by the heavy Higgs bosons exchange and the direct DM detection amplitude is dominated by the SM-like CP-even Higgs boson exchange. A larger suppression may be obtained for smaller $\tan\beta$ values ($\tan\beta \leq \mathcal{O}(10)$), for which the LHC bounds become weaker and the blind spot conditions [44]

$$\frac{2(M_1 + \mu \sin 2\beta)}{m_h^2} \simeq -\frac{\mu \tan \beta}{m_H^2} \quad (1)$$

may be approached by lowering the heavy CP-even Higgs boson mass m_H . The supersymmetric $g_\mu - 2$ contribution becomes subdominant in this region of mass parameters. A dedicated analysis of the allowed region of parameters would be interesting, but it is beyond the scope of this article.

3 Collider Analysis

3.1 Event Simulation and Characterization

The proposed signal is the LHC production of the lightest chargino, $\tilde{\chi}_1^\pm$, and the second-lightest neutralino, $\tilde{\chi}_2^0$, (both wino-like) in addition to an initial state radiation jet, as can be seen in Figure 1. We consider that the chargino decays to the lightest neutralino plus a lepton and a neutrino through an off-shell W boson, $\tilde{\chi}_1^\pm \rightarrow \tilde{\chi}_1^0 \ell \nu_\ell$, while the second lightest neutralino decays radiatively to the

lightest neutralino plus a photon, $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 + \gamma$. Therefore the LHC signal process under study is the following:

$$pp \rightarrow \tilde{\chi}_1^\pm \tilde{\chi}_2^0 j \rightarrow \tilde{\chi}_1^0 \ell \nu_\ell + \tilde{\chi}_1^0 \gamma + j. \quad (2)$$

Signal events have been generated with **MadGraph5_aMC@NLO** [122], using the Universal Feynman Output (UFO) [123] model **MSSM_SLHA2** [124] at LO in QCD with the NNPDF2.3 LO PDF set [125], for an LHC center-of-mass energy of 14 TeV. Parton showering and hadronization processes have been performed with **Pythia8** [126, 127], while **Delphes** [128] has been used for a fast detector simulation response, using the default ATLAS card³. Spin correlations in the decays of the charginos $\tilde{\chi}_1^\pm \rightarrow \tilde{\chi}_1^0 \ell^\pm \nu_\ell$ ($\ell = e, \mu$) have been taken into account via the **MadGraph5_aMC@NLO** decay chain syntax. We have simulated events within the mass range $m_{\tilde{\chi}_2^0} \in [200, 400]$ GeV, selecting BPs with a step of 50 GeV in $m_{\tilde{\chi}_2^0}$ as shown in Table 1. For each value of $m_{\tilde{\chi}_2^0}$, we explore three different values of $m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0}$.

Background events have been simulated at LO in QCD with **MadGraph5_aMC@NLO**, and subsequently have been processed with **Pythia8** and **Delphes**. We have considered the contributions to the mono-photon search channel, characterized by the presence of large missing transverse energy, a highly boosted initial state radiation jet, and the presence of at least one charge lepton. The dominant channels are W +jets, $W\gamma$ and $t\bar{t}$ +jets. Additionally, we also considered in our sample other subdominant contributions, including Z +jets, single-top, $t\bar{t}\gamma$ and diboson (ZZ , WW , and ZW). Other contributions are expected to be negligible.

Regarding object identification criteria for light leptons, electron candidates are required to have $p_T > 10$ GeV and $|\eta| < 2.47$, and should also lie outside the transition region ($1.37 < |\eta| < 1.52$) between the barrel and endcap calorimeters of the ATLAS detector, while muons must have $p_T > 10$ GeV and $|\eta| < 2.7$. Jets are required to satisfy $p_T > 20$ GeV and $|\eta| < 4.5$. Hadronically decaying τ leptons are required to have $p_T > 20$ GeV and $|\eta| < 2.47$, excluding the η range $1.37 < |\eta| < 1.52$. Photon candidates are required to have $p_T > 10$ GeV and $|\eta| < 2.37$.

Event selection criteria require at least one charged light lepton ($\ell = e, \mu$), at least one photon, and at least one jet. The leading jet, typically the ISR jet, is required to have $p_T > 100$ GeV. Due to the presence of neutralinos and neutrinos in the final state, we require $E_T^{\text{miss}} > 100$ GeV. We have checked that requiring $E_T^{\text{miss}} > 200$ GeV do not change significantly our results (see Appendix C). In this work, we propose a looser trigger for the HL-LHC which implies a larger number of expected events and could allow a better determination of uncertainties, that could be implemented by the experimental LHC collaborations in the searches for this class of DM signals.

The expected signal and background events yields estimated for the LHC at a center-of-mass energy of 14 TeV and a total integrated luminosity of 100 fb^{-1} after applying event selection criteria can be found in Table 2. For all the BPs selected in the MSSM parameter space, the expected yield is ~ 3 orders of magnitude lower than the expected total background yield. A first approximation of the discovery significance, $S/\sqrt{B}$, in each case, is shown in the third column of Table 2, where

³For the sake of simplicity, we focus our analysis in the ATLAS detector, but results are expected to be extensive to CMS experiment.

Process	Yield	BP #	Yield	$S/\sqrt{B}$
$W + \text{jets}$	60058	1	202	0.52
$W\gamma$	58462	2	459	1.19
$t\bar{t} + \text{jets}$	18051	3	637	1.65
$Z + \text{jets}$	3360	4	111	0.28
Single-top	3214	5	235	0.61
$t\bar{t}\gamma$	2498	6	334	0.86
Diboson	2340	7	66	0.17
Total background	147983	8	129	0.33
		9	179	0.46
		10	40	0.10
		11	74	0.19
		12	102	0.26
		13	23	0.05
		14	41	0.10
		15	57	0.14

Table 2: Expected background and signal events at the LHC with a center-of-mass energy of 14 TeV and a total integrated luminosity of 100 fb^{-1} , after applying event selection criteria.

S and B are the expected numbers of signal and background events, respectively. Neither evidence nor discovery is achieved in any of the examples.

After applying the baseline event selection criteria, we have explored a large but simple set of variables to characterize the kinematics of the studied final state, and subsequently to design our cut-based and machine-learning strategies. We considered not only fundamental low-level detector variables such as the transverse momentum (p_T) and pseudorapidity (η) of the leading jet (j_1), lepton (ℓ_1), and photon (γ_1), the missing transverse energy (E_T^{miss}) and object multiplicities, but several high-level observables that contribute to a more nuanced understanding of the underlying physics and play a crucial role enhancing the signal-to-background discrimination. Among these observables we incorporate the hadronic activity, defined as

$$H_T^{\text{jets}} = \sum p_T^{\text{jets}}, \quad (3)$$

which provides information about the formation of hadronic particles in the final state; the total transverse energy, a measure of the hardness of the event [129], defined as

$$H_T = \sum p_T^{\text{jets}} + \sum p_T^{\tau} + \sum p_T^e + \sum p_T^{\mu} + \sum p_T^{\gamma}; \quad (4)$$

the transverse mass of each leading particle $A = \{j_1, \ell_1, \gamma_1\}$,

$$m_T^A \equiv m_T(\mathbf{p}_T(A), \mathbf{E}_T^{\text{miss}}) = \sqrt{2p_T(A)E_T^{\text{miss}}(1 - \cos \Delta\phi(\mathbf{p}_T(A), \mathbf{E}_T^{\text{miss}}))}, \quad (5)$$

where $p_T(A) = |\mathbf{p}_T(A)|$, and $E_T^{\text{miss}} = |\mathbf{E}_T^{\text{miss}}|$; the scalar sum of the transverse momentum of all the

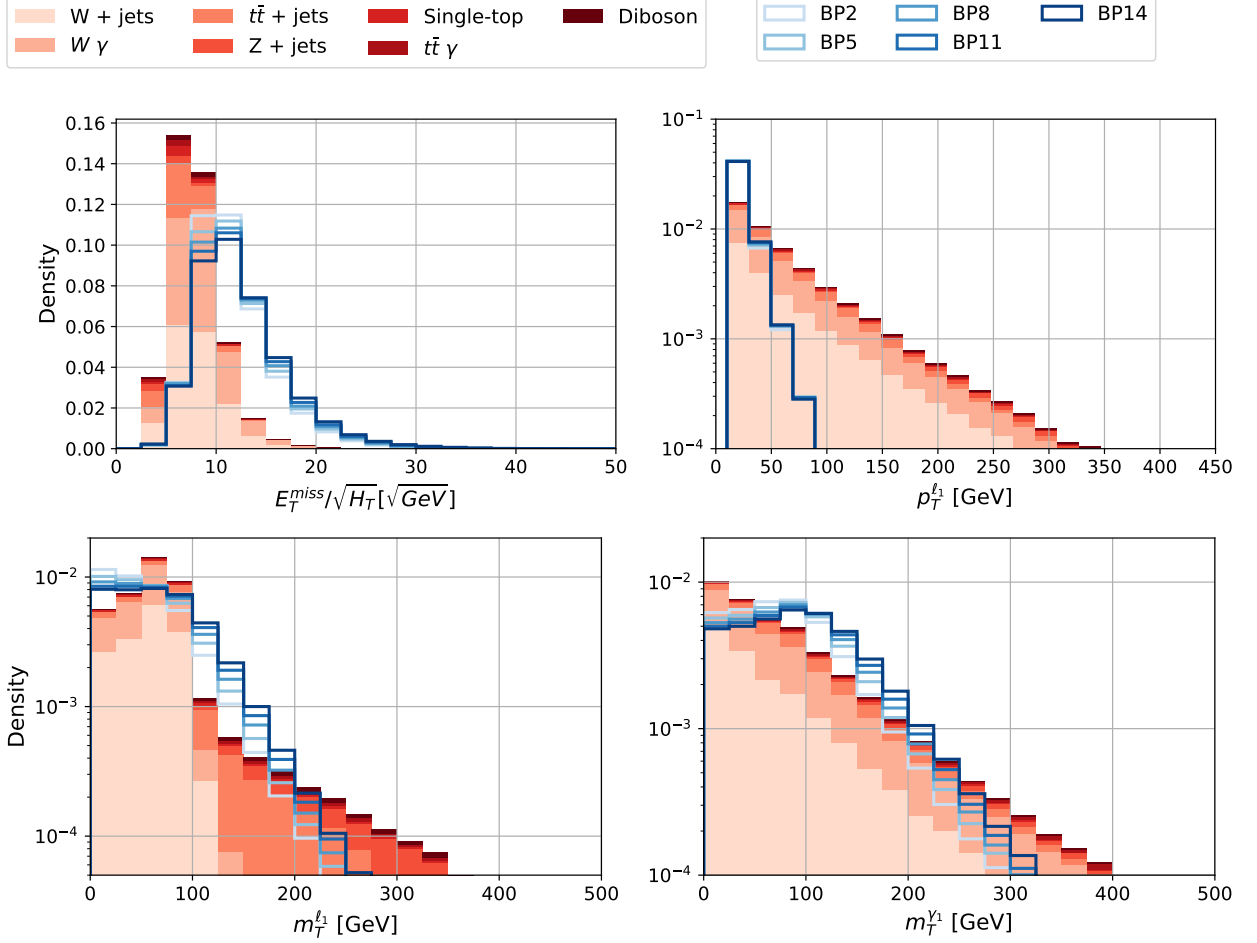


Figure 2: Distributions of the four final state kinematic variables showing the most discrimination between signal and background: missing transverse energy significance, $E_T^{\text{miss}}/\sqrt{H_T}$ (top-left panel); total transverse momentum of the leading lepton, $p_T^{\ell_1}$ (top-right panel); transverse mass of the leading lepton, $m_T^{\ell_1}$ (bottom-left panel); and transverse mass of the leading photon, $m_T^{\gamma_1}$ (bottom-right panel). We display a subset of the benchmark points, choosing those which produce the correct relic density.

leading particles,

$$s_T^1 = p_T^{\ell_1} + p_T^{j_1} + p_T^{\gamma_1}; \quad (6)$$

and the missing transverse energy significance, defined as $E_T^{\text{miss}}/\sqrt{H_T}$. A high value of the latter variable indicates that the observed E_T^{miss} cannot be explained by momentum resolution effects, then the event is more likely to contain undetected objects such as neutrinos or more exotic invisible particles [130].

In Figure 2 we present the distributions of the four most relevant variables for the discrimination task performed in the next section. The total SM background incorporates all the processes in Table 2, considering the weight of each channel by their relative cross sections. As an example, we show 5 different BPs that saturate the measured amount of DM relic density, while the rest of the BPs have similar behavior. We have found that the most discriminant variable is the missing

transverse energy significance (shown in the top-left panel), partially inherited from the harder E_T^{miss} distribution of the signal events given by the recoil of the $(\tilde{\chi}_2^0 + \tilde{\chi}_1^\pm)$ system against the ISR jet. Since $m_{\tilde{\chi}_1^0} \simeq m_{\tilde{\chi}_2^0} \simeq m_{\tilde{\chi}_1^\pm} \sim \mathcal{O}(200 - 400 \text{ GeV})$, we expect E_T^{miss} dominated by the contribution of the two lightest neutralinos in the final state. We obtain a high value of $E_T^{\text{miss}}/\sqrt{H_T}$ both for background and signal, manifesting the presence of undetected objects. In the case of the SM background, the only physical contribution comes from neutrinos produced in W decays which escape the detector, while for the signal there is a large contribution from the heavy invisible neutralinos.

The second most important feature, and the main low-level variable, is the transverse momentum of the leading lepton ($p_T^{\ell_1}$), shown in the top-right panel of Figure 2. Since the masses of the lightest chargino $\tilde{\chi}_1^\pm$ and the lightest neutralino $\tilde{\chi}_1^0$ are close, the leptons produced in the chargino decay through an off-shell W gauge boson, $\tilde{\chi}_1^\pm \rightarrow \tilde{\chi}_1^0 \ell \nu_\ell$, have a soft distribution. On the other hand, the leptons arise mostly from on-shell W -boson decays in the background case, with larger phase space.

The transverse mass of the leading lepton (bottom-left panel of Figure 2) is also important. In the considered SM processes, this variable peaks around the W -boson mass since the only contribution to E_T^{miss} comes from the neutrinos. The back-to-back configuration between the leading lepton and the reconstructed E_T^{miss} is mostly favored for background events, where the neutrinos can recoil with the corresponding charged leptons from the SM W -boson decays. As mentioned above, the E_T^{miss} in the signal process is dominated by the neutralinos, producing a broader $m_T^{\ell_1}$ distribution.

Finally, on the bottom-right panel of Figure 2 we show the fourth most relevant feature, the transverse mass of the leading photon. In the signal case, this feature peaks around 100 GeV, the minimum p_T value set in the event selection criteria for the leading jet. Since the $(\tilde{\chi}_2^0 + \tilde{\chi}_1^\pm)$ system decays to missing energy plus a photon and a soft lepton (the latter can be seen in the top-right panel of Figure 2), the transverse mass of the photon seems to reconstruct the system transverse energy that recoils against the ISR jet. On the other hand, in the background processes the photon transverse mass is not directly associated with the leading jet. The distributions for the rest of the low-level and high-level variables can be found in Figures 6 and 7 of Appendix A.

3.2 Analysis Strategies

In this section, we present two different approaches for our LHC exclusion and discovery reach analyses: a sequential cut-based analysis, that optimizes cuts in all the variables considered in the signal and background characterization; and a machine-learning one, which, after using a binary supervised classifier for signal and background discrimination, computes the final discovery significance using the entire ML output by means of the BL method and the unbinned MLL method. It is important to note that each analysis is trained with events from the complete set of BPs, and hence, can be applied to all the compressed parameter space.

3.2.1 Cut-Based Strategy

To refine the cut-and-count method, we propose a sequential cut-based (SCB) strategy, that automatically searches the best cuts in all variables by assessing the impact of each cut to the overall

significance. This implies the exploration of a high dimensional space that can be computationally expensive. In our case we have 22 variables, and we need to define lower and upper cuts, leading to 44 degrees of freedom. To simplify the problem, in this work we explore one variable at a time in an iterative way. Although this solution could be improved, this is beyond the scope of the paper as, in any case, the machine-learning approach described in the following section is more flexible than the rectangular cut-based method.

The SCB procedure is based on the following steps:

1. Perform a grid with 10 points evenly distributed in the available range, including its borders, of a randomly selected variable.
2. Use each point as an upper cut for the signal region and estimate the significance, $Z = S/\sqrt{B}$, where S and B are the numbers of signal and background expected events below the threshold associated with that point.
3. Use each point as a lower cut for the signal region and estimate the significance, $Z = S/\sqrt{B}$, where S and B are the numbers of signal and background expected events above the threshold associated with that point.
4. Apply and save the cut (superior or inferior) that leaves the best significance and has more than 5 signal and 5 background events.
5. Repeat items 1 to 4 taking into account the new updated range of each variable.

We have performed 100 iterations to allow each variable to be analyzed more than one time and hence the cuts can be updated. To make our analysis applicable to the entire parameter space, in steps 2 and 3 we have applied the upper and lower cut to all BP datasets and computed the resulting significance in each one. Then, the optimal cut decided in step 4 takes into account the significance average over all the BPs. Although the final signal region would not be the optimal one for each BP, the approach is more suited to explore the parameter space of interest.

3.2.2 Machine-Learning Strategy

For the ML analysis, we used a supervised per-event binary classifier based on the **XGBoost** [131,132] toolkit, trained to distinguish between signal and background events using as input features all the low-level and high-level variables described in Section 3.1. **XGBoost** or extreme gradient boosting is a ML method used in supervised problems either for classification or regression. It is also very popular in the high-energy physics community and widely used by ATLAS and CMS Collaborations. The algorithm is based on parallel tree boosting, i.e. it gives a prediction model in the form of an ensemble of weak prediction models (decision trees).

For the background dataset, we employed 200k events considering the relative weight of each background channel, while for the signal we used 200k events (equal number of events per BP). In this way, we ensure to have balanced datasets for the training of the machine-learning models. Notice that we train a single ML model using samples from all BPs. While this approach may not be

optimal for each particular BP, it is intentional. As mentioned before, the goal is to apply the same ML classifier across the entire parameter space, as the true underlying model would be unknown in an experimental search. In order to avoid overfitting, we have split each dataset into two independent datasets: a training set consisting of 80% of the original dataset; and a validation set consisting of the remaining 20% data. In addition, we have also prepared a third independent dataset, dubbed test set with ~ 400 background and signal events. It is used to assess the performance of the algorithm once it is already trained and to compute the significances as explained below.

The one-dimensional output of the binary classifier, $o(x) \in [0, 1]$, quantifies the probability of being signal-like of a given event, i.e., for signal-like events, the output $o(x)$ should be close to 1 while for background-like events $o(x)$ should be close to 0. The output of the ML classifier can be seen in the left panel of Figure 3 when tested with only pure background or pure signal samples, orange and green histograms, respectively. As expected, it can be seen that signal and background events peak at probabilities of 1 and 0 respectively, highlighting the classifier performance on the discrimination task (AUC= 0.94). For completeness, using the the mentioned ML classifier, we analyzed its performance distinguishing the signal against each one of the backgrounds. We found that the discrimination powers are very similar with a slightly better performance considering the more abundant ones. This comes from the fact that the training set was build weighting the number of events by the cross sections of each background process, hence the algorithm focuses more on those channels. In the right panel of Figure 3 we show the feature importance score for the XGBoost classifier. This magnitude is estimated from the gain metric that measures the relative contribution of the corresponding feature for generating a prediction. A higher value of this metric implies a higher impact on discrimination. The most relevant features, as already anticipated, are the missing transverse energy significance, $E_T^{\text{miss}}/\sqrt{H_T}$; the momentum of the leading lepton, $p_T^{\ell_1}$; the transverse mass of the leading lepton, $m_T^{\ell_1}$; and the transverse mass of the leading photon, $m_T^{\gamma_1}$.

After the binary classification, the computation of the exclusion and discovery reach ($Z = 2$ or $Z = 5$, respectively; with Z defined in Appendix B) is done using the entire one-dimensional classifier output in both binned and unbinned approaches, based on the traditional BL method and the MLL method, respectively. The binning process inherently sacrifices information on probability densities within each bin, impacting the likelihood estimation and potentially reducing the significance. While reducing bin width can mitigate this loss, finite statistics render this approximation unreliable. In contrast, unbinned methods maintain individual data point granularity, offering a potentially more accurate data representation without averaging information across bins, although it may suffer from numerical fluctuations that in binned methods are smoothed out.

On one hand, for the BL approach, the likelihood function is built as the product of the Poisson probability functions modeling the population of each bin, as shown in Appendix B. Therefore the full ML output is binned and turned into a histogram to find the expected number of signal and background events, to finally compute the significance with a likelihood ratio test.

On the other hand, the unbinned approach applied in this work is based on the MLL framework [93–96]. It uses KDE, a non-parametric method for density estimation (thus it does not assume a specific functional form for the underlying distribution), to fit the classifier output when

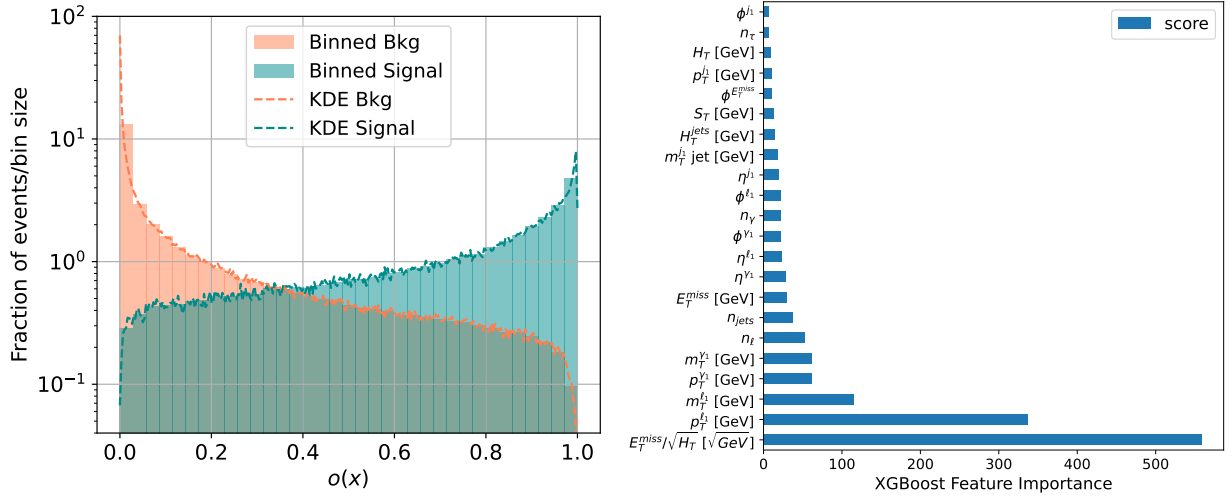


Figure 3: Left panel: output of the **XGBoost** classifier trained using all the BPs as signal (AUC= 0.94), when tested with only pure background (orange) or pure signal (green) samples. The dashed curves correspond to the PDFs obtained with the KDE fit. Right panel: feature importance score (gain metric) for the same **XGBoost** classifier.

tested with only pure background or pure signal samples, respectively. These fitted distributions are used as a one-dimensional approximation of the background and signal probability density functions (PDF), $\tilde{p}_s(o(x))$ and $\tilde{p}_b(o(x))$, and then introduced into an unbinned likelihood function to compute the significance with a likelihood ratio test, as can be seen in Appendix B. The KDE fit only involves the bandwidth parameter, which controls the degree of smoothness of the estimated density function. To obtain this parameter, we have trained independent signal and background KDEs with 50k input points each⁴. Then, the value of the bandwidth in each case has been found through a grid search employing the `GridSearchCV` function in the `sklearn.model_selection` Python package [133]. This method selects the bandwidth that maximizes data likelihood in a 5-fold cross-validation strategy. For further details, we refer the reader to the MLL references [93–96].

In the left panel of Figure 3 we also show, as dashed lines, the corresponding KDE fit for the background (orange lines) and signal (green lines) probability distributions. As expected, it can be seen that both KDE fits describes the probability distributions of the signal and background ML outputs. These probabilities are used in turn to assess the significance of the MLL method as explained in Appendix B.

4 Results

To compare the two analysis strategies, in Figure 4 we present the SCB signal enriched regions and the ML classifier behavior, applied to the BP1 dataset as an example. To ease visualization we show only four features: $E_T^{\text{miss}}/\sqrt{H_T}$, $p_T^{\ell_1}$, $m_T^{\ell_1}$, and $m_T^{\gamma_1}$, the most relevant variables according

⁴The density estimation tends to converge to the true underlying distribution for large datasets. We have checked that increasing the number of events used in this work for the KDE fits yields no significant changes.

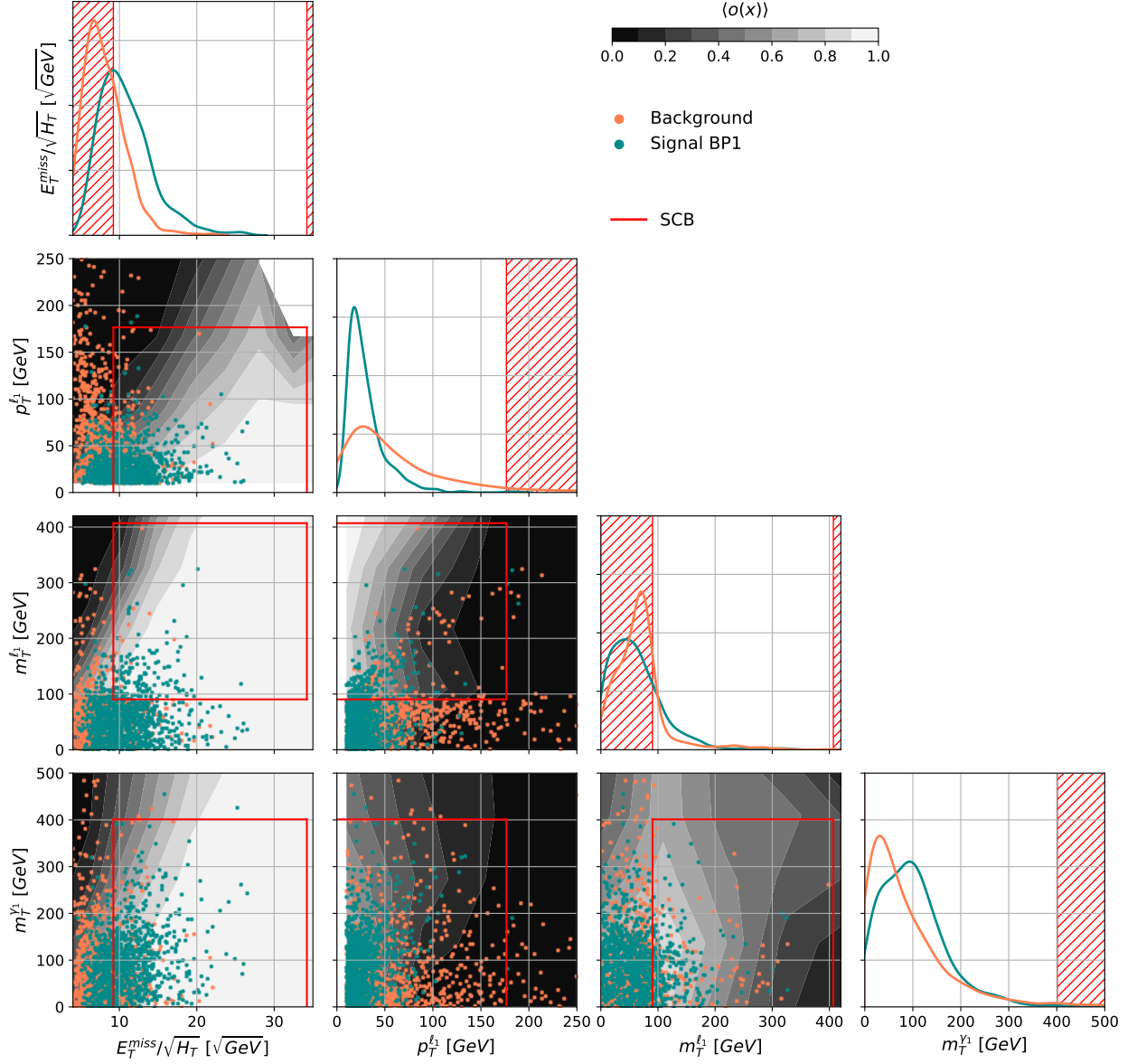


Figure 4: Correlation plot of the four most important features according to the ML algorithm (see right panel of Figure 3) for signal (green dots) and background (orange dots) events. The gray scale corresponds to the average machine-learning output $\langle o(x) \rangle$ in 2D bins. The signal-enriched regions found with the SCB method are demarcated with solid red lines. In the diagonal we show the PDF of each feature distinguishing signal and background events. Finally, the hatched regions correspond to the excluded areas of the SCB method.

to the ML algorithm. The points in orange correspond to background events, and the green ones to signal events (notice that background and signal events may overlap as can be seen from the distributions in the diagonal). We depict the cuts corresponding to the SCB algorithm as red solid boxes. The points outside the boxes are rejected and not included in the SCB significance computation. Regarding the correlations between the variables, we can see that the SCB has not established any lower cut for $m_T^{\gamma_1}$, even when one could assume that setting $m_T^{\gamma_1} \gtrsim 75$ GeV would be preferred by analyzing only the $m_T^{\gamma_1}$ distribution (bottom-right panel of Figure 2) dismissing its impact on the other features.

Concerning the ML-based approach, in Figure 4 the contour plots represent the average of the ML output of all the events in 2D bins, which we denote $\langle o(x) \rangle$. To compute this magnitude, we average over events with the same values of the two considered features while the other features are not fixed and cover the entire analyzed ranges. If the color is the darkest (black), an event in that region has $\langle o(x) \rangle \sim 0$, meaning that the algorithm determines that the event is background-like, whereas the lightest color (white) is associated with $\langle o(x) \rangle \sim 1$, implying that the event is more likely to be classified as signal-like. We highlight again that we are considering the average output, as the prediction for an event depends on all the variables. Nonetheless, this representation is useful to provide an idea of the behavior of the ML taking into account some correlations that otherwise are difficult to plot in a large parameter space. From Figure 4 we can see that the ML algorithm performs well identifying the events not only inside the signal-enriched regions obtained with the SCB process but also in the tails of the distributions. In this case, the correlations between variables are described better, for example in the $[p_T^{\ell_1}, E_T^{\text{miss}}/\sqrt{H_T}]$ plane the rectangular cuts do not seem to be optimal. Importantly, since the ML-based methods do not need to define a region to maximize the signal-to-background ratio, they can take advantage of the entire parameter space including the events hard to classify, with $\langle o(x) \rangle \sim 0.5$, as well as the tails of the distributions.

Our main results are summarized in Figure 5 where we show the projected exclusion and discovery reach, $Z = 2$ (dashed lines) and $Z = 5$ (solid lines), respectively, in the $[m_{\tilde{\chi}_2^0}, m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0}]$ plane, for the three strategies considered in this work. To obtain the contour levels, we have interpolated the significance values obtained with the 15 BPs. To assess the impact of including statistical uncertainties, we have adopted a conservative approach in which we have computed how the $Z = 2$ and $Z = 5$ regions are weakened considering $1\sigma_{\text{stat}}$, shown as hatched regions.⁵ To start, we find an improvement factor of $\sim 3 - 6$ in the discovery reach compared with the baseline significances reported in Table 2. For a fix $\tilde{\chi}_2^0$ mass, the sensitivity increases for smaller mass splittings. This behavior can be explained by the fact that smaller mass differences between the neutralinos result in larger branching ratios, leading to a higher number of expected signal events, as shown in Tables 1 and 2. In addition, the machine-learning strategies are significantly more sensitive than the cut-based one, with the unbinned MLL method providing a slightly larger discovery reach than the BL procedure. This improvement can be illustrated by noticing that for the SCB strategy, only the

⁵We have employed 2k pseudo-experiments for each BP to evaluate the test statistic and compute the significance, while its statistical uncertainty can be obtained from the variance of the resulting distribution. We have checked that increasing the amount of pseudo-experiments does not significantly modify the results.

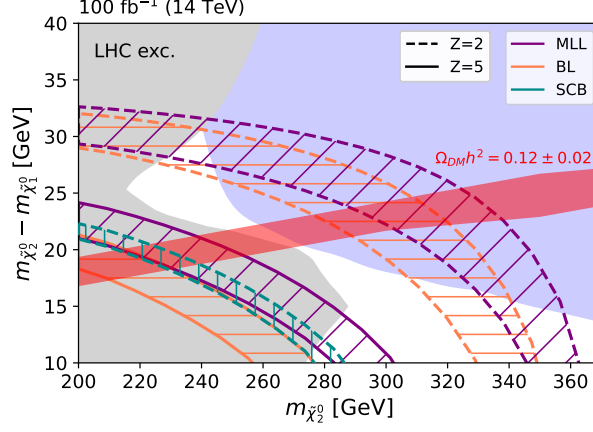


Figure 5: Projected exclusion ($Z = 2$, dashed contour lines) and discovery reach ($Z = 5$, solid contour lines) in the $[m_{\tilde{\chi}_2^0}, m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0}]$ plane applying the optimized sequential cut-based strategy (SCB, green), the Binned-Likelihood approach (BL, orange), and the Machine-Learned Likelihood method (MLL, violet). All points in this parameter space lead to values of Δa_μ close to 2×10^{-9} (see Table 1). Direct detection constraints are shown as a light-blue region, including the recent LUX-ZEPLIN results [119]. LHC bounds [114] are shown as a gray-shaded region. The red band indicates where the model parameters explain the observed DM relic density.

$Z = 2$ level is within the parameter space of interest and in the same region as the $Z = 5$ reach provided by ML methods.

As can be seen in Table 1, the entire parameter space reproduces the value of the muon $g_\mu - 2$ within 2σ . The models where the neutralinos reproduce the observed DM abundance is shown as a red band, while points below (above) produce DM under (over) abundance, but may be viable by additional SM modifications without affecting the collider signatures. LHC exclusion bounds taken from Ref. [114] are shown in gray and covers almost all the region with $Z \geq 5$, but a small space with $275 \leq m_{\tilde{\chi}_2^0} \leq 305$ GeV and $10 \leq m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0} \leq 15$ GeV. Direct detection limits including the recently announced LUX-ZEPLIN results [119] are shown in light-blue, and are complementary to the LHC bounds excluding important regions with high neutralino masses and mass differences. We would like to highlight that the new direct detection results are almost an impressive one order of magnitude stronger than previous constraints that only excluded the region above $m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0} \simeq 40$ GeV. In Appendix C we show the results considering a more stringent selection criteria consistent with current LHC triggers, specifically, a cut on the missing transverse energy of $E_T^{\text{miss}} > 200$ GeV. It can be seen that both results are compatible, but the latter suffer from larger statistical uncertainties due to a lower number of events.

Remarkably, almost all the allowed parameter space in Figure 5 can be probed with $Z \geq 2$ by the machine-learning methods described in this article, and can be divide in three regions:

- A region with $275 \leq m_{\tilde{\chi}_2^0} \leq 365$ GeV and $m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0} \leq 20$ GeV, where an additional DM component besides the lightest neutralino would be needed.
- A small region where the lightest neutralino can reproduce the right amount of dark matter

abundance (red band) with $250 \leq m_{\tilde{\chi}_2^0} \leq 280$ GeV and $20 \leq m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0} \leq 22$ GeV. This region is particularly interesting, not only because these points could satisfy all experimental constraints considering only MSSM processes, provide an explanation for the entire DM budget within the standard freeze-out paradigm and a potential solution to the small multilepton excesses [101–103], but also because it could be probed by complementary strategies, both LHC searches and direct dark matter experiments, in the near future.

- A region where one needs to invoke a dilution process to sufficiently reduce the DM overdensity in the mass range $m_{\tilde{\chi}_2^0} \sim 220–250$ GeV and $22 \leq m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0} \leq 32$ GeV, but could also explain the ATLAS and CMS collaboration multilepton excesses.

It is worth noting that the direct detection constraint are applied in the vanilla freeze-out scenario with standard cosmology if $\Omega_{\tilde{\chi}_1^0} h^2 \leq 0.12$. As mentioned above, for models with DM overdensity we assume a dilution factor by additional SM modifications, such as a late injection of entropy during the Big Bang or the presence of another species with which the $\tilde{\chi}_1^0$ can coannihilate, to obtain $\Omega_{\tilde{\chi}_1^0} h^2 = 0.12$. However, these limits may be relaxed even further in scenarios where the dilution process leads to less than the observed relic density, thereby weakening direct detection constraints and opening the entire light-blue region of Figure 5. In that scenario, the proposed machine-learning methods can explore up to $m_{\tilde{\chi}_2^0} \leq 365$ GeV and $m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0} \leq 33$ GeV. On the other hand, collider constraints are more robust since they are independent of the particular processes assumed during the early universe.

Finally, a few comments are relevant to the discussion on the potential sensitivity presented in this work. In our analysis, we have not included systematic uncertainties that, for example, could be incorporated as variations in the dataset after their modeling using control regions. Therefore, we expect that the sensitivity would be reduced taking into account these effects. Nonetheless, the potential coverage of the parameter space presented in this proof-of-concept study is quite impressive and highlights the power of ML techniques. We expect our study to encourage the experimental collaborations to perform a complete and more detailed analysis. We find this novel channel to be very promising and capable of complementing or even improving on the results of the usual di-lepton and tri-lepton plus missing energy searches [25, 101–103] to shed light on the challenging, yet well motivated, compressed mass scenario.

5 Conclusions

The search for weakly interacting massive particles at the LHC is well motivated and, due to their small production cross section, will highly benefit from the large luminosities expected at the LHC in the coming years. Additionally, new tools made available by machine learning have allowed for more refined and precise approaches to separating signal events from large Standard Model backgrounds. A particular case of interest is when new weakly interacting massive particles decay to a potential dark matter candidate, whose experimental signature involves significant amounts of missing energy. Traditional current searches at the LHC in the framework of Supersymmetric

theories include those associated with di-lepton and trilepton signatures plus missing energy and are very efficient. Recent theoretical studies, however, have proposed a new search channel, where a supersymmetric WIMP has a significant radiative decay to the dark matter candidate. In this work we performed a first study of this new search channel and show that machine-learning techniques are promising and may complement or even improve on the results of traditional search strategies in the search for dark matter.

Our study explores electroweakino searches at the LHC, particularly for models that produce interesting but soft final states. There exists such a region in the MSSM, the “compressed region,” which contains both a dark matter candidate, that reproduces the observed DM relic density, and yields a contribution to $g_\mu - 2$ which can relieve the apparent tension between SM predictions and experimental observations. This region favors negative values of $\mu \times M_1$, leading to the suppression of direct DM detection cross sections, making the DM candidate a possible target for direct DM detection experiments in the coming years. This region also favors positive values of $\mu \times M_2$ and not too heavy slepton masses (below a TeV). Combining these conditions it happens that the second lightest neutralino tends to have a significant radiative decay into the lightest neutralino and a photon. Traditional searches, focus on multilepton final states and have placed limits on electroweakino searches in the compressed region. Here, however, we concentrate in the interesting, but yet unused signature of the radiative decay of the second neutralino ($\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 + \gamma$). Due to the low energy of the resulting photon, cut-and-count methods typically fail to precisely carve out the kinematic region where the signal is strongest.

In this article, we showed that the significance of electroweakino searches in the compressed region may be greatly improved employing machine-learning methods that explore the full correlation between the kinematic variables of these processes. Our analysis shows significant potential for the search of weak scale particles in this radiative decay channel. If neutralinos reproduce the observed DM abundance, ML techniques could probe ($Z \geq 2$) the allowed region between current LHC and direct detection constraints, $250 \leq m_{\tilde{\chi}_2^0} \leq 280$ GeV and $20 \leq m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0} \leq 22$ GeV, covering the MSSM parameter space consistent with radiative neutralino decays. If one allows additional processes which would not affect collider signatures, such as late time entropy injection or the presence of more DM candidate to dilute or complement the lightest neutralino relic production, machine-learning methods can explore up to $m_{\tilde{\chi}_2^0} \leq 365$ GeV and $m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0} \leq 33$ GeV, including the entire region where ATLAS and CMS have found small multilepton excesses. The analysis performed in this work ignores systematic errors and must be therefore corroborated by a full experimental analysis. The LHC experimental collaborations should be encouraged to proceed with such a study, which is complementary to the traditional di-lepton and tri-lepton plus missing energy ones.

Acknowledgments. The authors thank Ana Cueto Gómez, Fernando Monticelli, and Hernán Wahlberg for useful discussions on background and triggers. MC, DR, and CW would like to thank Sebastian Baum, Tong Ou, and Nausheen Shah for their collaboration in the early stages of this analysis, and Sebastian Wagner-Carena for insightful comments on our work. This work is partially supported by the Spanish Research Agency (Agencia Estatal de Investigación) through the grants

IFT Centro de Excelencia Severo Ochoa No CEX2020-001007-S (EA, MdIR, AP, RMSS), PID2021-124704NB-I00 (EA, RMSS), RYC-2017-22986 (RMSS), CNS2023-144536 (RMSS), and PID2021-125331NB-I00 (MdIR, AP), funded by MCIN/AEI/10.13039/501100011033, and by the Comunidad Autónoma de Madrid through the grant SI2/PBG/2020-00005 (MdIR, AP). MdIR is supported by the Next Generation EU program, in the context of the National Recovery and Resilience Plan, Investment PE1 – Project FAIR “Future Artificial Intelligence Research”. MC and CW would like to thank the Aspen Center for Physics, which is supported by National Science Foundation grant No. PHY-1607611, where part of this work has been done. Fermilab is operated by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy. The work of CW at the University of Chicago is also supported by the DOE grant DE-SC0013642. Work at Argonne is partially financed by the U.S. Department of Energy (DOE), Div. of HEP, Contract DE-AC02-06CH11357. This work is partially supported by the DOE under Task TeV of contract DE-FGO2-96-ER40956. This work is supported by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists, and Office of Science Graduate Student Research (SCGSR) program. The SCGSR program is administered by the Oak Ridge Institute for Science and Education for the DOE under contract number DE-SC0014664.

Appendix

A Kinematic Distributions

In this appendix, we present the distributions of the features employed to analyze the proposed signal (the four most relevant variables according to the ML classifier are shown in Figure 2).

In Figure 6 we show the p_T distributions for the leading photon and the leading jet, the missing transverse energy E_T^{miss} (upper row), the pseudorapidity η for all the leading visible objects (middle row), as well as the object multiplicities (bottom row), both for the background and five different benchmark points leading to the correct DM relic density.

Since the characteristic feature of the proposed signal is the presence of a soft photon, the naive approach would be to focus and establish dedicated cuts on the transverse momentum of the leading photon ($p_T^{\gamma_1}$). However, by comparison with the kinematic variables shown in Figure 3, it turns out to be the fifth most important feature for ML discrimination. The ISR boost leads to an increase of the kinematics of the events, and consequently of the radiative decay, increasing $p_T^{\gamma_1}$. On the other hand, the background photons tend to be much more energetic than the signal ones and, similarly to the lepton case, this is caused by the fact that the photon comes from the radiative decay of the wino-like second lightest neutralino $\tilde{\chi}_2^0$, with small mass splitting with the bino-like lightest neutralino $\tilde{\chi}_1^0$. Moreover, the E_T^{miss} is much larger in the signal than in the background, as expected by the production of a pair of $\tilde{\chi}_1^0$ in the former, while in the background it is only generated by the presence of neutrinos from the leptonic W -boson decays, or purely instrumental by the presence of soft jets and the defects in the reconstruction of the events. In the case of $p_T^{j_1}$, the distribution is slightly softer for the dominant backgrounds since the jet is recoiling with objects that are lighter

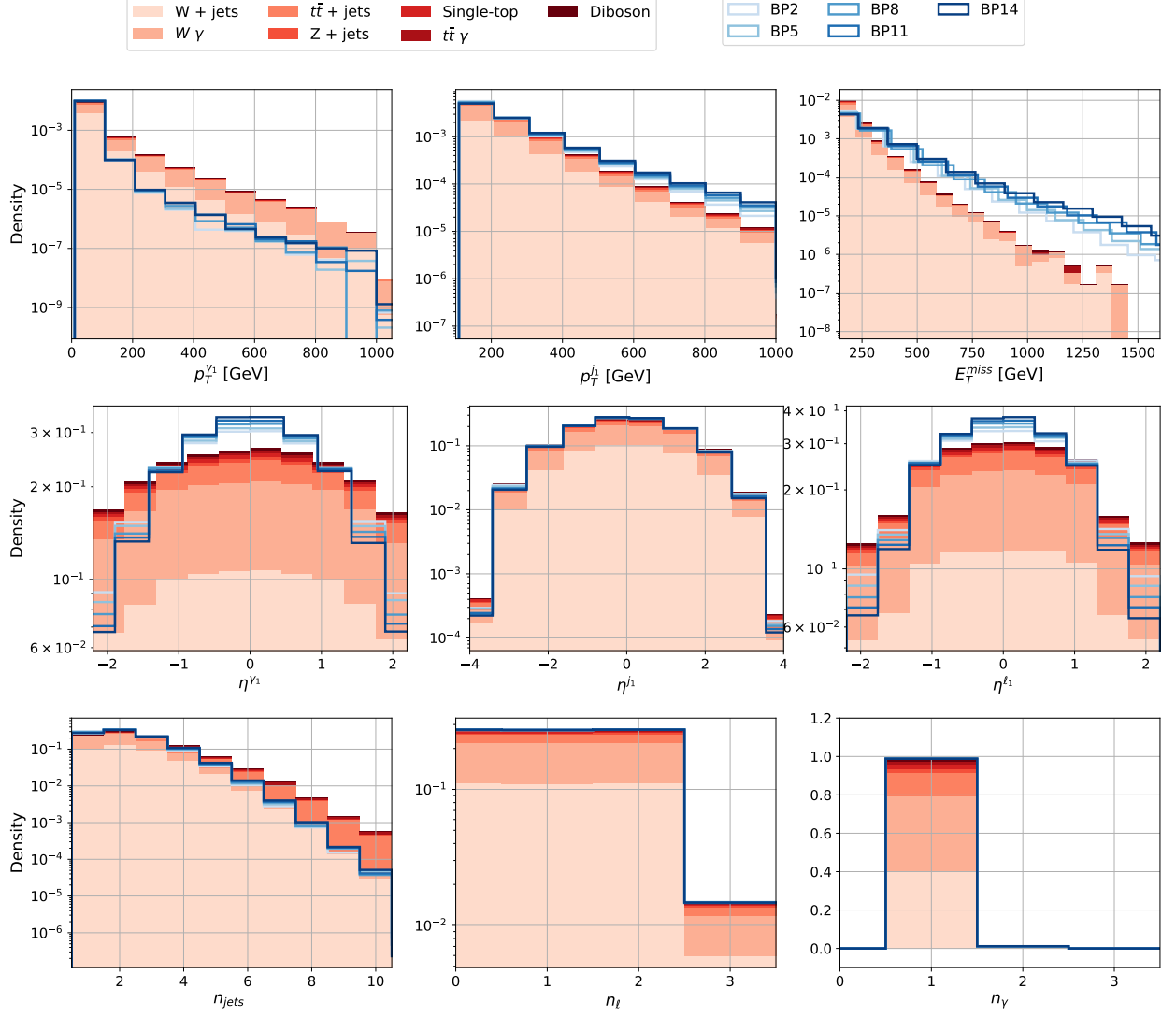


Figure 6: Total transverse momentum of the leading photon ($p_T^{\gamma_1}$) and the leading jet ($p_T^{j_1}$), missing transverse energy (E_T^{miss}), pseudorapidity (η) of the three leading visible objects in the final state ($\{\gamma_1, j_1, \ell_1\}$) and object multiplicities, after applying event selection criteria.

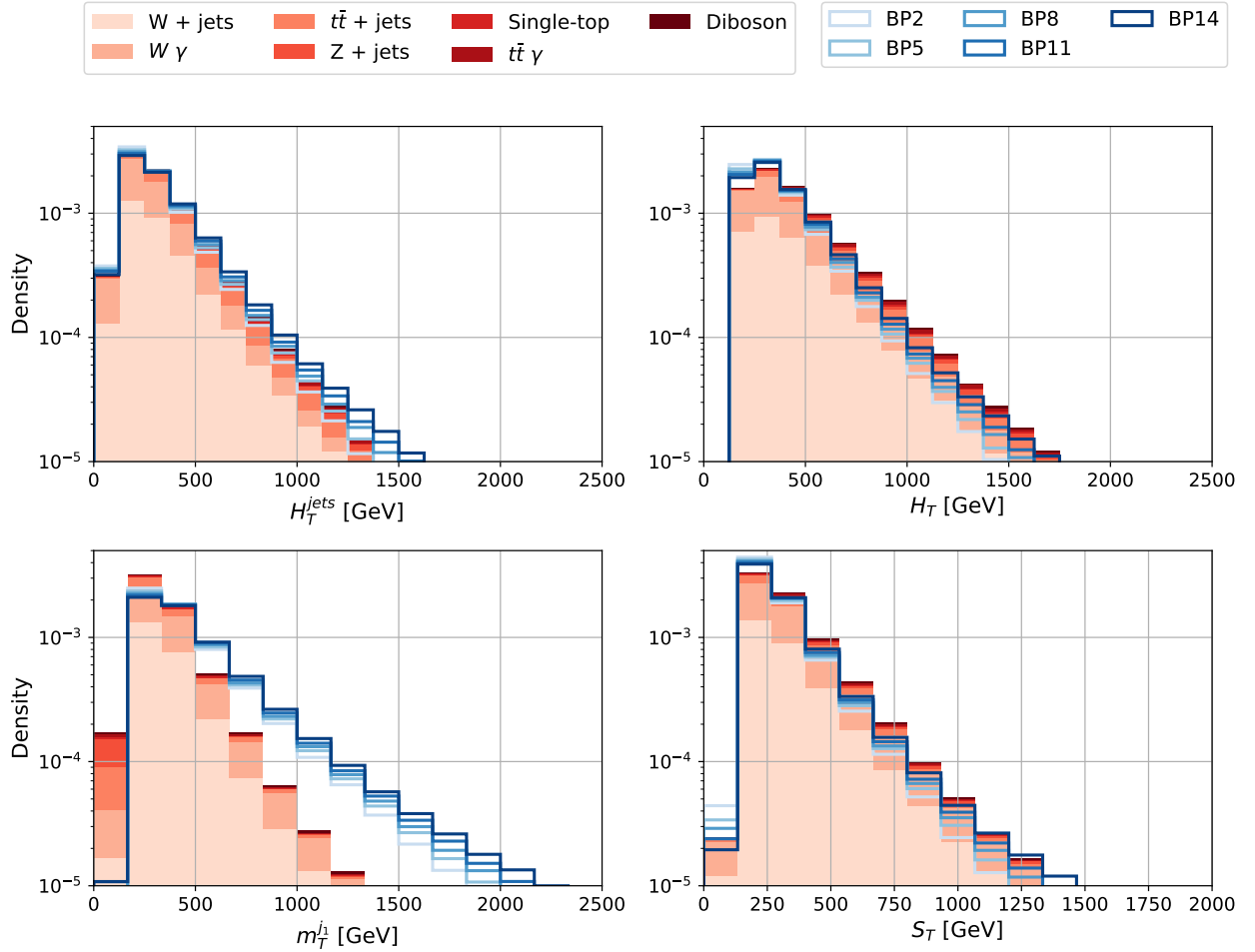


Figure 7: Hadronic activity (H_T^{jets}), total transverse energy (H_T), transverse mass of the leading jet ($m_T^{j_1}$), and scalar sum of the transverse momentum of all the leading particles (S_T) in the final state, after applying event selection criteria.

than in the signal case.

Regarding angular distributions, the leading photon and leading lepton are more centered in the η distributions because they are decay products from the electroweakino-pair system, which is generated back-to-back with respect to the ISR jet, while for the background the pseudorapidity distributions are more homogeneous. In the case of the leading jet, no angular difference is observed between background and signal distributions as expected.

Regarding object multiplicities, we observe a higher n_{jets} spectrum in the background. The dominant contribution to this effect is the $t\bar{t}$ +jets channel, coming from the hadronic decays of the $t\bar{t}$ system (which provides an extra W boson compared to the signal and the other backgrounds).

An overall insight into the behavior of the rest of high-level objects can be obtained in Figure 7. Among them, it is worth mentioning the invariant mass of the leading jet ($m_T^{j_1}$), which is higher for signal events. This behavior is expected given the size of the initial boost (higher than 100 GeV), which favors the back-to-back configuration between the leading jet and the reconstructed E_T^{miss} .

B Review on MLL and BL Formulae

In this appendix, we summarize the formulae for the MLL and the BL procedures to estimate the exclusion and discovery potential of the LHC. For the MLL method, the likelihood function $\mathcal{L}$ of N independent measurements, described by an arbitrarily high-dimensional set of observations, x , is given by

$$\mathcal{L}(\mu, s, b) = \text{Poiss}(N|\mu S + B) \prod_{i=1}^N p(x_i|\mu, s, b), \quad (7)$$

where S (B) is the expected total signal (background) yield and μ defines the hypothesis we are testing for. Additionally, Poiss stands for Poisson probability distribution, and $p(x|\mu, s, b)$ is the probability density for a single measurement x ,

$$p(x|\mu, s, b) = \frac{B}{\mu S + B} p_b(x) + \frac{\mu S}{\mu S + B} p_s(x), \quad (8)$$

where $p_s(x) = p(x|s)$ ($p_b(x) = p(x|b)$) is the signal (background) probability density function (PDF) for x , encoding the event-by-event information. The discovery reach analysis corresponds to studying the background-only hypothesis ($\mu = 0$), then the log-likelihood ratio test statistic becomes

$$\tilde{q}_0 = \begin{cases} 0 & \text{if } \hat{\mu} < 0 \\ -2 \text{Ln} \frac{\mathcal{L}(0, s, b)}{\mathcal{L}(\hat{\mu}, s, b)} = -2\hat{\mu}S + 2 \sum_{i=1}^N \text{Ln} \left(1 + \frac{\hat{\mu} S p_s(x_i)}{B p_b(x_i)} \right) & \text{if } \hat{\mu} \geq 0, \end{cases} \quad (9)$$

where $\hat{\mu}$ is the parameter that maximizes the likelihood. The value of $\hat{\mu}$ can be estimated numerically by calculating the zero of the partial derivative of Eq. (7) with respect to μ .

Since $p_{s,b}(x)$ are unknown, the base idea used by the MLL method is to replace these densities for the one-dimensional classification score $o(x)$ that can be obtained from a machine-learning binary classifier,

$$p_s(x) \rightarrow \tilde{p}_s(o(x)), \quad p_b(x) \rightarrow \tilde{p}_b(o(x)), \quad (10)$$

where $\tilde{p}_s(o(x))$ ($\tilde{p}_b(o(x))$) is the distribution of $o(x)$ for signal (background), obtained by evaluating the output of the classifier on a set of pure signal (background) events. Since the distributions are one-dimensional, they can be easily handled. The KDE method is then used to extract these distributions without binning and incorporate them into Eq.(9).

The test statistic in Eq. (9) is estimated through a finite data set of N events, with the probability distribution conditioned on the true unknown signal strength μ' . When the true hypothesis is assumed to be the signal-plus-background ($\mu' = 1$), the median expected discovery significance $\text{med} [Z_0|1]$ is defined as

$$Z_{\text{MLL}} = \text{med} [Z_0|1] = \sqrt{\text{med} [\tilde{q}_0|1]}, \quad (11)$$

where we estimate the $\tilde{q}_0$ distribution numerically by generating a set of pseudo-experiments with signal-plus-background events. The value of Z_{MLL} encodes how likely is to the background-only hypothesis to explain data that follows the signal-plus-background hypothesis, and a higher value of Z_{MLL} excludes the background-only hypothesis with higher confidence.

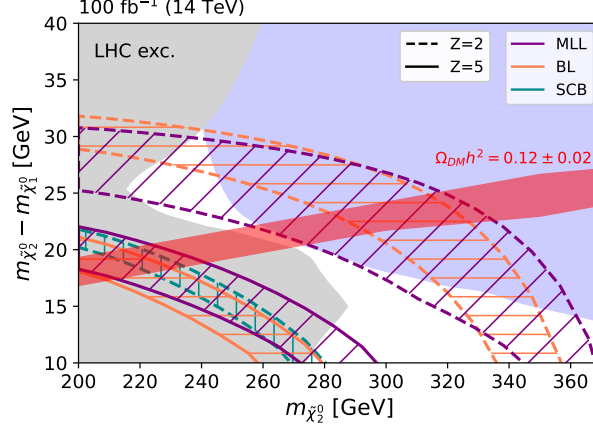


Figure 8: Projected discovery significance as shown in Fig. 5 but considering the event selection criteria with $E_T^{\text{miss}} > 200$ GeV.

For the BL method, we consider the output of the ML, $o(x) \in [0, 1]$, and divide its range in a histogram-based approach. We can describe the population of each bin as an independent Poissonian distribution, then likelihood function $\mathcal{L}$ is given by

$$\mathcal{L}(\mu, s, b) = \prod_{i=1}^N \text{Poiss}(N_d | \mu S_d + B_d), \quad (12)$$

with S_d and B_d defined as the expected number of signal and background events in each bin d . We can repeat the previous procedure to obtain the log-likelihood ratio test statistic, however through the use of Asimov data sets [92] the final formula for the discovery potential can be estimated as

$$Z_{\text{BL}} = \text{med } [Z_0 | 1] = \sqrt{\text{med } [\tilde{q}_0 | 1]} = \left[2 \sum_{d=1}^D \left((S_d + B_d) \text{Ln} \left(1 + \frac{S_d}{B_d} \right) - S_d \right) \right]^{1/2}. \quad (13)$$

C Results with $E_T^{\text{miss}} > 200$ GeV

In this appendix we present the results employing the current LHC selection triggers. The only difference with the cuts described in the main body is that we tighten the minimum value of E_T^{miss} from 100 GeV to 200 GeV. This value satisfies current LHC selection triggers for E_T^{miss} [134]. In Table 3 we show the expected number of events with the new selection cuts. Notice that we expect approximately half of the signal events compared with the ones in Table 2 but 4 times less background events, which results in a similar $S/\sqrt{B}$.

Finally, in Fig. 8 we present the discovery limits considering the selection criteria with $E_T^{\text{miss}} > 200$ GeV. Comparing the results with the ones presented in Fig. 5, we can see that the overall reach is only slightly less sensitive than those derived with $E_T^{\text{miss}} > 100$ GeV.

Process	Yield	BP #	Yield	$S/\sqrt{B}$
$W + \text{jets}$	10834	1	81	0.53
$W\gamma$	11002	2	206	1.37
$t\bar{t}\gamma$	700	3	344	2.29
Total background	22536	4	49	0.32
		5	112	0.74
		6	179	1.19
		7	32	0.21
		8	65	0.43
		9	97	0.64
		10	20	0.13
		11	38	0.25
		12	56	0.37
		13	12	0.07
		14	22	0.14
		15	32	0.21

Table 3: Expected background and signal events at the LHC with a center-of-mass energy of 14 TeV and a total integrated luminosity of 100 fb^{-1} , after applying same event selection criteria described in the main text but with $E_T^{\text{miss}} > 200 \text{ GeV}$.

References

- [1] G. Steigman and M.S. Turner, *Cosmological Constraints on the Properties of Weakly Interacting Massive Particles*, *Nucl. Phys. B* **253** (1985) 375.
- [2] G. Arcadi, M. Dutra, P. Ghosh, M. Lindner, Y. Mambrini, M. Pierre et al., *The waning of the WIMP? A review of models, searches, and constraints*, *Eur. Phys. J. C* **78** (2018) 203 [[1703.07364](#)].
- [3] PARTICLE DATA GROUP collaboration, *Review of particle physics*, *Phys. Rev. D* **110** (2024) 030001.
- [4] P. Fayet, *Supersymmetry and Weak, Electromagnetic and Strong Interactions*, *Phys. Lett. B* **64** (1976) 159.
- [5] P. Fayet, *Spontaneously Broken Supersymmetric Theories of Weak, Electromagnetic and Strong Interactions*, *Phys. Lett. B* **69** (1977) 489.
- [6] H.P. Nilles, *Supersymmetry, Supergravity and Particle Physics*, *Phys. Rept.* **110** (1984) 1.
- [7] H.E. Haber and G.L. Kane, *The Search for Supersymmetry: Probing Physics Beyond the Standard Model*, *Phys. Rept.* **117** (1985) 75.
- [8] J. Gunion and H.E. Haber, *Higgs Bosons in Supersymmetric Models. 1.*, *Nucl. Phys. B* **272** (1986) 1.
- [9] S.P. Martin, *A Supersymmetry primer*, *Adv. Ser. Direct. High Energy Phys.* **18** (1998) 1 [[hep-ph/9709356](#)].

- [10] G.R. Farrar and P. Fayet, *Phenomenology of the Production, Decay, and Detection of New Hadronic States Associated with Supersymmetry*, *Phys. Lett. B* **76** (1978) 575.
- [11] S. Dimopoulos and H. Georgi, *Softly Broken Supersymmetry and SU(5)*, *Nucl. Phys. B* **193** (1981) 150.
- [12] S. Weinberg, *Supersymmetry at Ordinary Energies. 1. Masses and Conservation Laws*, *Phys. Rev. D* **26** (1982) 287.
- [13] N. Sakai and T. Yanagida, *Proton Decay in a Class of Supersymmetric Grand Unified Models*, *Nucl. Phys. B* **197** (1982) 533.
- [14] S. Dimopoulos, S. Raby and F. Wilczek, *Proton Decay in Supersymmetric Models*, *Phys. Lett. B* **112** (1982) 133.
- [15] ATLAS collaboration, *Search for supersymmetry in final states with two same-sign or three leptons and jets using 36 fb^{-1} of $\sqrt{s} = 13\text{ TeV}$ pp collision data with the ATLAS detector*, *JHEP* **09** (2017) 084 [[1706.03731](#)].
- [16] ATLAS collaboration, *Search for photonic signatures of gauge-mediated supersymmetry in 13 TeV pp collisions with the ATLAS detector*, *Phys. Rev. D* **97** (2018) 092006 [[1802.03158](#)].
- [17] CMS collaboration, *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](#)].
- [18] ATLAS collaboration, *Search for squarks and gluinos in final states with jets and missing transverse momentum using 139 fb^{-1} of $\sqrt{s} = 13\text{ TeV}$ pp collision data with the ATLAS detector*, *JHEP* **02** (2021) 143 [[2010.14293](#)].
- [19] ATLAS collaboration, *Search for squarks and gluinos in final states with one isolated lepton, jets, and missing transverse momentum at $\sqrt{s} = 13$ with the ATLAS detector*, *Eur. Phys. J. C* **81** (2021) 600 [[2101.01629](#)].
- [20] ATLAS collaboration, *Search for new phenomena in events with an energetic jet and missing transverse momentum in pp collisions at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector*, *Phys. Rev. D* **103** (2021) 112006 [[2102.10874](#)].
- [21] CMS collaboration, *Combined searches for the production of supersymmetric top quark partners in proton-proton collisions at $\sqrt{s} = 13\text{ TeV}$* , *Eur. Phys. J. C* **81** (2021) 970 [[2107.10892](#)].
- [22] CMS collaboration, *Search for top squark pair production in a final state with at least one hadronically decaying tau lepton in proton-proton collisions at $\sqrt{s} = 13\text{ TeV}$* , *JHEP* **07** (2023) 110 [[2304.07174](#)].
- [23] ATLAS collaboration, *SUSY August 2023 Summary Plot Update*, *ATL-PHYS-PUB-2023-025* (2023) .
- [24] ATLAS collaboration, *The quest to discover supersymmetry at the ATLAS experiment*, [2403.02455](#).
- [25] CMS collaboration, *Combined search for electroweak production of winos, binos, higgsinos, and sleptons in proton-proton collisions at $\sqrt{s} = 13\text{ TeV}$* , [2402.01888](#).
- [26] ATLAS collaboration, *A statistical combination of ATLAS Run 2 searches for charginos and neutralinos at the LHC*, [2402.08347](#).

- [27] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [[1807.06209](#)].
- [28] J.R. Ellis, T. Falk and K.A. Olive, *Neutralino - Stau coannihilation and the cosmological upper limit on the mass of the lightest supersymmetric particle*, *Phys. Lett. B* **444** (1998) 367 [[hep-ph/9810360](#)].
- [29] J.R. Ellis, T. Falk, K.A. Olive and M. Srednicki, *Calculations of neutralino-stau coannihilation channels and the cosmologically relevant region of MSSM parameter space*, *Astropart. Phys.* **13** (2000) 181 [[hep-ph/9905481](#)].
- [30] M.R. Buckley, D. Hooper and J. Kumar, *Phenomenology of Dirac Neutralino Dark Matter*, *Phys. Rev. D* **88** (2013) 063532 [[1307.3561](#)].
- [31] T. Han, Z. Liu and A. Natarajan, *Dark matter and Higgs bosons in the MSSM*, *JHEP* **11** (2013) 008 [[1303.3040](#)].
- [32] M.E. Cabrera, J.A. Casas, A. Delgado, S. Robles and R. Ruiz de Austri, *Naturalness of MSSM dark matter*, *JHEP* **08** (2016) 058 [[1604.02102](#)].
- [33] M.J. Baker and A. Thamm, *Leptonic WIMP Coannihilation and the Current Dark Matter Search Strategy*, *JHEP* **10** (2018) 187 [[1806.07896](#)].
- [34] T.T. Yanagida, W. Yin and N. Yokozaki, *Bino-wino coannihilation as a prediction in the E_7 unification of families*, *JHEP* **12** (2019) 169 [[1907.07168](#)].
- [35] H. Baer, T. Krupovnickas, A. Mustafayev, E.-K. Park, S. Profumo and X. Tata, *Exploring the BWCA (bino-wino co-annihilation) scenario for neutralino dark matter*, *JHEP* **12** (2005) 011 [[hep-ph/0511034](#)].
- [36] PANDAX-4T collaboration, *Dark Matter Search Results from the PandaX-4T Commissioning Run*, *Phys. Rev. Lett.* **127** (2021) 261802 [[2107.13438](#)].
- [37] LZ collaboration, *First Dark Matter Search Results from the LUX-ZEPLIN (LZ) Experiment*, *Phys. Rev. Lett.* **131** (2023) 041002 [[2207.03764](#)].
- [38] XENON collaboration, *First Dark Matter Search with Nuclear Recoils from the XENONnT Experiment*, *Phys. Rev. Lett.* **131** (2023) 041003 [[2303.14729](#)].
- [39] PICO collaboration, *Dark Matter Search Results from the PICO-60 C_3F_8 Bubble Chamber*, *Phys. Rev. Lett.* **118** (2017) 251301 [[1702.07666](#)].
- [40] J.R. Ellis, A. Ferstl and K.A. Olive, *Reevaluation of the elastic scattering of supersymmetric dark matter*, *Phys. Lett. B* **481** (2000) 304 [[hep-ph/0001005](#)].
- [41] J.R. Ellis, A. Ferstl and K.A. Olive, *Exploration of elastic scattering rates for supersymmetric dark matter*, *Phys. Rev. D* **63** (2001) 065016 [[hep-ph/0007113](#)].
- [42] J.R. Ellis, K.A. Olive, Y. Santoso and V.C. Spanos, *Update on the direct detection of supersymmetric dark matter*, *Phys. Rev. D* **71** (2005) 095007 [[hep-ph/0502001](#)].
- [43] H. Baer, A. Mustafayev, E.-K. Park and X. Tata, *Target dark matter detection rates in models with a well-tempered neutralino*, *JCAP* **01** (2007) 017 [[hep-ph/0611387](#)].
- [44] P. Huang and C.E.M. Wagner, *Blind Spots for neutralino Dark Matter in the MSSM with an intermediate m_A* , *Phys. Rev. D* **90** (2014) 015018 [[1404.0392](#)].

- [45] P. Huang, R.A. Roglans, D.D. Spiegel, Y. Sun and C.E.M. Wagner, *Constraints on Supersymmetric Dark Matter for Heavy Scalar Superpartners*, *Phys. Rev. D* **95** (2017) 095021 [[1701.02737](#)].
- [46] T. Han, H. Liu, S. Mukhopadhyay and X. Wang, *Dark Matter Blind Spots at One-Loop*, *JHEP* **03** (2019) 080 [[1810.04679](#)].
- [47] S. Baum, M. Carena, N.R. Shah and C.E.M. Wagner, *The tiny ($g-2$) muon wobble from small- μ supersymmetry*, *JHEP* **01** (2022) 025 [[2104.03302](#)].
- [48] T. Aoyama et al., *The anomalous magnetic moment of the muon in the Standard Model*, *Phys. Rept.* **887** (2020) 1 [[2006.04822](#)].
- [49] T. Aoyama, M. Hayakawa, T. Kinoshita and M. Nio, *Complete Tenth-Order QED Contribution to the Muon $g-2$* , *Phys. Rev. Lett.* **109** (2012) 111808 [[1205.5370](#)].
- [50] T. Aoyama, T. Kinoshita and M. Nio, *Theory of the Anomalous Magnetic Moment of the Electron*, *Atoms* **7** (2019) 28.
- [51] A. Czarnecki, W.J. Marciano and A. Vainshtein, *Refinements in electroweak contributions to the muon anomalous magnetic moment*, *Phys. Rev. D* **67** (2003) 073006 [[hep-ph/0212229](#)].
- [52] C. Gnendiger, D. Stöckinger and H. Stöckinger-Kim, *The electroweak contributions to $(g-2)_\mu$ after the Higgs boson mass measurement*, *Phys. Rev. D* **88** (2013) 053005 [[1306.5546](#)].
- [53] M. Davier, A. Hoecker, B. Malaescu and Z. Zhang, *Reevaluation of the hadronic vacuum polarisation contributions to the Standard Model predictions of the muon $g-2$ and $\alpha(m_Z^2)$ using newest hadronic cross-section data*, *Eur. Phys. J. C* **77** (2017) 827 [[1706.09436](#)].
- [54] A. Keshavarzi, D. Nomura and T. Teubner, *Muon $g-2$ and $\alpha(M_Z^2)$: a new data-based analysis*, *Phys. Rev. D* **97** (2018) 114025 [[1802.02995](#)].
- [55] G. Colangelo, M. Hoferichter and P. Stoffer, *Two-pion contribution to hadronic vacuum polarization*, *JHEP* **02** (2019) 006 [[1810.00007](#)].
- [56] M. Hoferichter, B.-L. Hoid and B. Kubis, *Three-pion contribution to hadronic vacuum polarization*, *JHEP* **08** (2019) 137 [[1907.01556](#)].
- [57] M. Davier, A. Hoecker, B. Malaescu and Z. Zhang, *A new evaluation of the hadronic vacuum polarisation contributions to the muon anomalous magnetic moment and to $\alpha(m_Z^2)$* , *Eur. Phys. J. C* **80** (2020) 241 [[1908.00921](#)].
- [58] A. Keshavarzi, D. Nomura and T. Teubner, *$g-2$ of charged leptons, $\alpha(M_Z^2)$, and the hyperfine splitting of muonium*, *Phys. Rev. D* **101** (2020) 014029 [[1911.00367](#)].
- [59] A. Kurz, T. Liu, P. Marquard and M. Steinhauser, *Hadronic contribution to the muon anomalous magnetic moment to next-to-next-to-leading order*, *Phys. Lett. B* **734** (2014) 144 [[1403.6400](#)].
- [60] K. Melnikov and A. Vainshtein, *Hadronic light-by-light scattering contribution to the muon anomalous magnetic moment revisited*, *Phys. Rev. D* **70** (2004) 113006 [[hep-ph/0312226](#)].
- [61] P. Masjuan and P. Sanchez-Puertas, *Pseudoscalar-pole contribution to the $(g_\mu-2)$: a rational approach*, *Phys. Rev. D* **95** (2017) 054026 [[1701.05829](#)].

- [62] G. Colangelo, M. Hoferichter, M. Procura and P. Stoffer, *Dispersion relation for hadronic light-by-light scattering: two-pion contributions*, *JHEP* **04** (2017) 161 [[1702.07347](#)].
- [63] M. Hoferichter, B.-L. Hoid, B. Kubis, S. Leupold and S.P. Schneider, *Dispersion relation for hadronic light-by-light scattering: pion pole*, *JHEP* **10** (2018) 141 [[1808.04823](#)].
- [64] A. Gérardin, H.B. Meyer and A. Nyffeler, *Lattice calculation of the pion transition form factor with $N_f = 2 + 1$ Wilson quarks*, *Phys. Rev. D* **100** (2019) 034520 [[1903.09471](#)].
- [65] J. Bijnens, N. Hermansson-Truedsson and A. Rodríguez-Sánchez, *Short-distance constraints for the HLbL contribution to the muon anomalous magnetic moment*, *Phys. Lett. B* **798** (2019) 134994 [[1908.03331](#)].
- [66] G. Colangelo, F. Hagelstein, M. Hoferichter, L. Laub and P. Stoffer, *Longitudinal short-distance constraints for the hadronic light-by-light contribution to $(g - 2)_\mu$ with large- N_c Regge models*, *JHEP* **03** (2020) 101 [[1910.13432](#)].
- [67] T. Blum, N. Christ, M. Hayakawa, T. Izubuchi, L. Jin, C. Jung et al., *Hadronic Light-by-Light Scattering Contribution to the Muon Anomalous Magnetic Moment from Lattice QCD*, *Phys. Rev. Lett.* **124** (2020) 132002 [[1911.08123](#)].
- [68] G. Colangelo, M. Hoferichter, A. Nyffeler, M. Passera and P. Stoffer, *Remarks on higher-order hadronic corrections to the muon $g-2$* , *Phys. Lett. B* **735** (2014) 90 [[1403.7512](#)].
- [69] MUON G-2 collaboration, *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*, *Phys. Rev. Lett.* **126** (2021) 141801 [[2104.03281](#)].
- [70] H.E. Haber and D. Wyler, *Radiative neutralino decay*, *Nucl. Phys. B* **323** (1989) 267.
- [71] S. Ambrosanio and B. Mele, *Supersymmetric scenarios with dominant radiative neutralino decay*, *Phys. Rev. D* **55** (1997) 1399 [[hep-ph/9609212](#)].
- [72] H. Baer and T. Krupovnickas, *Radiative neutralino decay in supersymmetric models*, *JHEP* **09** (2002) 038 [[hep-ph/0208277](#)].
- [73] S. Baum, M. Carena, T. Ou, D. Rocha, N.R. Shah and C.E.M. Wagner, *Lighting up the LHC with Dark Matter*, *JHEP* **11** (2023) 037 [[2303.01523](#)].
- [74] S. Roy and C.E.M. Wagner, *Dark Matter searches with photons at the LHC*, *JHEP* **04** (2024) 106 [[2401.08917](#)].
- [75] A.J. Larkoski, I. Moutl and B. Nachman, *Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Machine Learning*, *Phys. Rept.* **841** (2020) 1 [[1709.04464](#)].
- [76] P. Mehta, M. Bukov, C.-H. Wang, A.G.R. Day, C. Richardson, C.K. Fisher et al., *A high-bias, low-variance introduction to Machine Learning for physicists*, *Phys. Rept.* **810** (2019) 1 [[1803.08823](#)].
- [77] D. Guest, K. Cranmer and D. Whiteson, *Deep Learning and its Application to LHC Physics*, *Ann. Rev. Nucl. Part. Sci.* **68** (2018) 161 [[1806.11484](#)].
- [78] K. Albertsson et al., *Machine Learning in High Energy Physics Community White Paper*, *J. Phys. Conf. Ser.* **1085** (2018) 022008 [[1807.02876](#)].

- [79] A. Radovic, M. Williams, D. Rousseau, M. Kagan, D. Bonacorsi, A. Himmel et al., *Machine learning at the energy and intensity frontiers of particle physics*, *Nature* **560** (2018) 41.
- [80] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby et al., *Machine learning and the physical sciences*, *Rev. Mod. Phys.* **91** (2019) 045002 [[1903.10563](#)].
- [81] D. Bourilkov, *Machine and Deep Learning Applications in Particle Physics*, *Int. J. Mod. Phys. A* **34** (2020) 1930019 [[1912.08245](#)].
- [82] M. Feickert and B. Nachman, *A Living Review of Machine Learning for Particle Physics*, [2102.02770](#).
- [83] M.D. Schwartz, *Modern Machine Learning and Particle Physics*, [2103.12226](#).
- [84] G. Karagiorgi, G. Kasieczka, S. Kravitz, B. Nachman and D. Shih, *Machine Learning in the Search for New Fundamental Physics*, [2112.03769](#).
- [85] Y. Coadou, *Boosted decision trees*, [2206.09645](#).
- [86] P. Shanahan et al., *Snowmass 2021 Computational Frontier CompF03 Topical Group Report: Machine Learning*, [2209.07559](#).
- [87] T. Plehn, A. Butter, B. Dillon and C. Krause, *Modern Machine Learning for LHC Physicists*, [2211.01421](#).
- [88] V. Belis, P. Odagiu and T.K. Aarrestad, *Machine learning for anomaly detection in particle physics*, *Rev. Phys.* **12** (2024) 100091 [[2312.14190](#)].
- [89] J. Bardhan, T. Mandal, S. Mitra, C. Neeraj and M. Patra, *Unsupervised and lightly supervised learning in particle physics*, [2403.13676](#).
- [90] S. Mondal and L. Mastrolorenzo, *Machine Learning in High Energy Physics: A review of heavy-flavor jet tagging at the LHC*, [2404.01071](#).
- [91] J. Halverson, *TASI Lectures on Physics for Machine Learning*, [2408.00082](#).
- [92] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) 1554 [[1007.1727](#)].
- [93] E. Arganda, X. Marcano, V.M. Lozano, A.D. Medina, A.D. Perez, M. Szewc et al., *A method for approximating optimal statistical significances with machine-learned likelihoods*, *Eur. Phys. J. C* **82** (2022) 993 [[2205.05952](#)].
- [94] E. Arganda, M. de los Rios, A.D. Perez and R.M. Sandá Seoane, *Imposing exclusion limits on new physics with machine-learned likelihoods*, *PoS ICHEP2022* (2022) 1226.
- [95] E. Arganda, A.D. Perez, M. de los Rios and R.M. Sandá Seoane, *Machine-learned exclusion limits without binning*, *Eur. Phys. J. C* **83** (2023) 1158 [[2211.04806](#)].
- [96] E. Arganda, D.A. Díaz, A.D. Perez, R.M. Sandá Seoane and A. Szykman, *LHC study of third-generation scalar leptoquarks with machine-learned likelihoods*, *Phys. Rev. D* **109** (2024) 055032 [[2309.05407](#)].
- [97] M. Rosenblatt, *Remarks on Some Nonparametric Estimates of a Density Function*, *Ann. Math. Statist.* **27** (1956) 832.
- [98] E. Parzen, *On Estimation of a Probability Density Function and Mode*, *Ann. Math. Statist.* **33** (1962) 1065.

- [99] S. Borsanyi et al., *Leading hadronic contribution to the muon magnetic moment from lattice QCD*, *Nature* **593** (2021) 51 [[2002.12347](#)].
- [100] N.M. Coyle and C.E.M. Wagner, *Resolving the muon $g - 2$ tension through Z' -induced modifications to σ_{had}* , *JHEP* **12** (2023) 071 [[2305.02354](#)].
- [101] ATLAS collaboration, *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](#)].
- [102] ATLAS collaboration, *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](#)].
- [103] CMS collaboration, *Search for supersymmetry in final states with two or three soft leptons and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **04** (2022) 091 [[2111.06296](#)].
- [104] ATLAS collaboration, *Search for electroweak production of supersymmetric particles in final states with two or three leptons at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **78** (2018) 995 [[1803.02762](#)].
- [105] ATLAS collaboration, *Search for electroweak production of charginos and sleptons decaying into final states with two leptons and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector*, *Eur. Phys. J. C* **80** (2020) 123 [[1908.08215](#)].
- [106] CMS collaboration, *Search for supersymmetry in final states with two oppositely charged same-flavor leptons and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **04** (2021) 123 [[2012.08600](#)].
- [107] CMS collaboration, *Search for electroweak production of charginos and neutralinos in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **04** (2022) 147 [[2106.14246](#)].
- [108] CMS collaboration, *Search for chargino-neutralino production in events with Higgs and W bosons using 137 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **10** (2021) 045 [[2107.12553](#)].
- [109] ATLAS collaboration, *Search for charginos and neutralinos in final states with two boosted hadronically decaying bosons and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **104** (2021) 112010 [[2108.07586](#)].
- [110] ATLAS collaboration, *Searches for new phenomena in events with two leptons, jets, and missing transverse momentum in 139 fb^{-1} of $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, [2204.13072](#).
- [111] ATLAS collaboration, *Search for electroweak production of supersymmetric states in scenarios with compressed mass spectra at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **97** (2018) 052010 [[1712.08119](#)].
- [112] ATLAS collaboration, *Search for chargino-neutralino production with mass splittings near the electroweak scale in three-lepton final states in $\sqrt{s}=13$ TeV pp collisions with the ATLAS detector*, *Phys. Rev. D* **101** (2020) 072001 [[1912.08479](#)].

- [113] D. Agin, B. Fuks, M.D. Goodsell and T. Murphy, *Seeking a coherent explanation of LHC excesses for compressed spectra*, [2404.12423](#).
- [114] M. Chakraborti, S. Heinemeyer and I. Saha, *Consistent excesses in the search for $\tilde{\chi}_2^0 \tilde{\chi}_1^\pm$: wino/bino vs. Higgsino dark matter*, *Eur. Phys. J. C* **84** (2024) 812 [[2403.14759](#)].
- [115] J.-L. Kneur, G. Moultaka, M. Ughetto, D. Zerwas and A. Djouadi, *SuSpect3: A C++ code for the supersymmetric and Higgs particle spectrum of the MSSM*, *Comput. Phys. Commun.* **291** (2023) 108805 [[2211.16956](#)].
- [116] A. Djouadi, M.M. Mühlleitner and M. Spira, *Decays of supersymmetric particles: The Program SUSY-HIT (SUspect-SdecaY-Hdecay-InTerface)*, *Acta Phys. Polon. B* **38** (2007) 635 [[hep-ph/0609292](#)].
- [117] W. Beenakker, R. Hopker and M. Spira, *PROSPINO: A Program for the production of supersymmetric particles in next-to-leading order QCD*, [hep-ph/9611232](#).
- [118] G. Bélanger, F. Boudjema, A. Goudelis, A. Pukhov and B. Zaldivar, *micrOMEGAs5.0 : Freeze-in*, *Comput. Phys. Commun.* **231** (2018) 173 [[1801.03509](#)].
- [119] LZ COLLABORATION collaboration, *Dark Matter Search Results from 4.2 Tonne-Years of Exposure of the LUX-ZEPLIN (LZ) Experiment*, [2410.17036](#).
- [120] CMS collaboration, *Search for additional neutral MSSM Higgs bosons in the $\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **09** (2018) 007 [[1803.06553](#)].
- [121] ATLAS collaboration, *Search for heavy Higgs bosons decaying into two tau leptons with the ATLAS detector using pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. Lett.* **125** (2020) 051801 [[2002.12223](#)].
- [122] 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](#)].
- [123] L. Darmé et al., *UFO 2.0: the ‘Universal Feynman Output’ format*, *Eur. Phys. J. C* **83** (2023) 631 [[2304.09883](#)].
- [124] C. Duhr and B. Fuks, *A superspace module for the FeynRules package*, *Comput. Phys. Commun.* **182** (2011) 2404 [[1102.4191](#)].
- [125] R.D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244 [[1207.1303](#)].
- [126] T. Sjöstrand, S. Ask, J.R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [[1410.3012](#)].
- [127] T. Sjostrand, S. Mrenna and P.Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852 [[0710.3820](#)].
- [128] 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](#)].
- [129] ATLAS collaboration, *Object-based missing transverse momentum significance in the ATLAS detector*, .

- [130] T.C. collaboration, *Missing transverse energy performance of the cms detector*, *Journal of Instrumentation* **6** (2011) P09001–P09001.
- [131] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, [1603.02754](#).
- [132] T. Chen and C. Guestrin, *XGBoost: A scalable tree boosting system*, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, (New York, NY, USA), pp. 785–794, ACM, 2016, [DOI](#).
- [133] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel et al., *Scikit-learn: Machine learning in Python*, *Journal of Machine Learning Research* **12** (2011) 2825.
- [134] ATLAS collaboration, *Trigger menu in 2018*, [ATL-DAQ-PUB-2019-001](#) (2019) .