LOOKING FOR NEW PHYSICS: FROM DARK MATTER TO MACHINE LEARNING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Yik Chuen, San

December 2024

LOOKING FOR NEW PHYSICS: FROM DARK MATTER TO MACHINE LEARNING

Yik Chuen, San, Ph.D.

Cornell University 2024

Ever since the Standard Model (SM) has been written down, countless efforts have been made to complement/extend it with new inputs from both theoretical and experimental sides. In this work, we provide such possible extensions from two perspectives: phenomenology (model-building) and data analysis.

From the model-building perspective, dark matter has proven to be one of the most robust signs we have about new physics. As such, we propose new models involving dark matter, along with analyses of their properties and methods of probing them in experiments.

An alternative approach of discovering new physics involves the use of more sophisticated data analysis methods based on modern machine learning models and techniques. On this front, we present works regarding improvements towards two existing proposals - 'Classification Without Labels' and 'Anomaly Detection with Density Estimation'.

BIOGRAPHICAL SKETCH

Yik Chuen San obtained his Bachelor of Science degree from the Hong Kong University of Science and Technology in August 2018. Afterwards, he started pursuing his doctoral degree at Cornell University under the supervision of Prof. Maxim Perelstein, with a focus on theoretical particle physics. Outside of physics, he also enjoys playing the piano and cooking.

This work is dedicated to my parents and my bygone self, without whom I would not be here today.

ACKNOWLEDGEMENTS

My journey for the past six years has been one full of swerves - from living by myself for the first time in a foreign land, to experiencing a global pandemic, to suffering from depression, and now to me concluding this journey with a newfound understanding of what I am. Throughout this journey, I have met, both directly and indirectly, many people who, in one way or another, have helped shape me into the human I am today.

First and foremost, I thank my mother and father for giving me the treasure of life, and a chance to see a world I could not have fathomed. My journey would not have been possible without their support.

For the people I met at Cornell, I thank all my friends: Andrew and Mijo, for the countless banter we have had about both the academic and the non-academic; Ameen, for being not just an admirable colleague, but also the perfect roommate; Fernanda, Steven, and Margarita, for taking the initiatives to invite me out for fun (even though I don't always end up joining); Han and Avinash, for listening to my rants about life; and Namitha, for sharing various teaching duties and giving me help whenever I needed it.

I also thank professors in the particle phenomenology group: Csaba, for his inspiring lectures which reshaped my understanding of physics; and Yuval, for being always so cheerful and approachable; and of course, Maxim, my supervisor, for taking a chance on me and teaching me how to think about physics and do research properly, for giving me guidance when I am lost, and for letting me feel that I belong here.

Finally, I would like to express gratitude to my past self, for having a dream and deciding to pursue this path. You might be disappointed if you could see me today, but I want you to know that while this is not the destination you had

envisioned, there is no regret along this path that we have taken.

It is almost certain that I have missed many people who I should thank, and for that, I apologize - alas, my memory is not what it used to be six years ago!

Yik Chuen San

Ithaca, July 2024

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

The beauty of the Standard Model lies in its simplicity - it is specified only by its gauge groups and the corresponding fermion and scalar representations (unified by the gauge symmetry principle). Despite its simplicity and its success in explaining much of the data we have observed in experiments (both collider and astronomical), there remains a lot of unsatisfactory aspects and unanswered questions. Examples include neutrino masses [1], the hierarchy problem [2], the strong *CP* problem [3], and of course the nature of dark matter. There has been a plethora of works aiming to study these missing pieces, all of which come in different flavours - from proposing new theoretical principles to detailed analysis of experimental data. In this work, we present works on studying new physics, taking both model-building and data-driven perspectives, with emphases on dark matter-related models, and applications of machine learning.

1.1 Dark Matter

Dark matter is perhaps one of the most robust signals of physics beyond the Standard Model (BSM)[1] (in comparison to a lot of fleeting signals coming from collider experiments). Perhaps what we know most about it is the fact that we do not know much about it. The few things about dark matter we are more or less sure of are that it is stable, cold (i.e. non-relativistic), collision-less, and have zero or very small electric charge. Taking an optimistic point of view, this means that there are a lot of rooms for various models of dark matter. In what follows

[1]It is worth noting that while the majority of the community believes that nature of dark matter is not explained by the Standard Model, there are some works exploring alternative possibilities.[4, 5]

below we briefly summarize current evidence and searches for dark matter. For more detailed discussion, we recommend [6].

1.1.1 Evidence of Dark Matter

The evidence of existence of dark matter primarily comes from astrophysical and cosmological observations, with studies involving galaxy rotation curves being one of the most important kind. [7]

Consider a rotating galaxy. By measuring the Doppler shift of, say, the hydrogen spectral lines of different stars in the galaxy, one is able to compute speeds of the stars at various distances from the center of the galaxy. This is the so-called galactic rotation curve. Theoretically, what shape of this curve should we expect? If we assume spherical symmetry for the sake of simplicity, one can conclude from Newton's law of motion that the speed $v(r)$ of a particle at a distance r from the galactic center is given by

$$v(r) = \sqrt{\frac{GM(r)}{r}}, \tag{1.1}$$

where $M(r)$ is the total mass enclosed within a radius r.

For a spiral galaxy, much of its (visible) mass is concentrated at the center, so that $M(r)$ is basically constant at large r. This in return implies that we should expect $v(r)$ to drop as $r^{-1/2}$ at large distances. However in experiments what was observed is that $v(r)$ remains roughly constant, which is in stark contrast with our expectations above. The most natural explanation of this apparent discrepancy is that there exists invisible matter permeating throughout the galaxy that we did not take into account of - dark matter.[2]

[2]Realistic experiments are of course far more complicated than the cartoon picture we

Another, more modern, piece of evidence of the existence of dark matter comes from gravitational lensing of galaxy clusters. Very roughly, it involves directly measuring the total mass distribution via gravitational lensing, and comparing it with visible matter (most of it in the form of hot gases) distribution measured via X-ray spectroscopy. Once again, large discrepancies exist between the two measurements, implying the existence of dark matter.

1.1.2 Searches of Dark Matter

Current experiments looking for dark matter include both astrophysical/cosmological experiments and collider experiments. For astrophysical/cosmological experiments, they fall under two broad categories - direct and indirect detection.

In direct detection methods, one considers an incoming dark matter particle colliding with a target. This collision deposits energy on the target which one tries to measure. In experiments, the collision targets are typically either atomic nuclei or electrons. These experiments are most sensitive (in the sense of setting the tightest bounds on DM-matter cross-sections) when the mass of the incoming dark matter particle is comparable to the mass of the target particle: when the DM is too small, the recoil energy is too small to be detectable, while when the DM mass is too large there are fewer collisions (assuming constant DM density). Experiments exploring this front include XENON1T and SENSEI amongst others.

In indirect detection methods, instead of detecting recoils, the goal is to de-

sketched here, please see [8] for more details.

tect the pair annihilation or decay of dark matter into Standard Model particles. These particles in principle should contribute to cosmic rays one can observe near Earth on top of contributions coming from other sources. If one has a good model of the background sources, annihilation/decay of dark matter can be detected as an excess over such background in the form of a 'bump-hunting'. Therefore, when designing these kinds of experiments, it is very important to have both a good understanding of background, as well as enough signals (i.e. high enough dark matter density). Examples of indirect detection experiments include Super-Kamiokande and IceCube.

At colliders, instead of relying on pre-existing dark matter in the universe, one aims to produce them directly in collisions of particles. Due to the observational requirement that dark matter is stable and chargeless, they must be produced in pairs and subsequently escape the detectors, leaving their signatures as missing energies just like neutrinos. On the other hand, in some models dark matter is produced via massive mediators. These mediators can therefore be detected through 'bump-hunting' techniques.

1.1.3 Types of Dark Matter Models

Due to our lack of understanding of some key properties (e.g. mass, spin) of dark matter, there are so many different models of dark matter that it can be difficult to classify them all. Of course, all these models must still obey the properties we mentioned above. Here we try to list a few different popular classes of dark matter models.

- Axionic Dark Matter Models The (QCD) axion [9] was first proposed as an elegant solution to the strong CP problem. Soon it was realized that it can serve as a potential candidate of dark matter. Even though most of the parameter space of QCD axions has been ruled out by experiments, axion-like particles (similar to QCD axions, except that their coupling to the Standard Model is not controlled by QCD symmetry breaking scale) remain an attractive candidate.
- Dark Sector Models This class of models purports the existence a whole sector of 'dark' particles which only communicate with the Standard Model via a mediator (i.e. a 'dark force'). The most typical model includes augmenting the Standard Model gauge group with a dark $U(1)$ via kinetic mixing. This model is the main subject of Chapter 3.
- Weakly Interacting Massive Particles (WIMP) Models Models in this class assume that the dark matter is charged under the electroweak gauge group with zero electric charge, and is a color singlet. Assuming close-to-weak-scale mass and order 1 coupling, these particles can correctly reproduce the thermal relic abundance of dark matter observed. This is sometimes referred to as the 'WIMP miracle'. Unfortunately, experiments have severely constrained WIMP models.

1.2 Machine Learning in Particle Physics

While historically model-building has been very successful, in recent years the lack of significant discoveries has ushered another school of thought in which one seeks to learn/discover new physics in a more data-driven approach. Ma-

chine learning soon stand to become a very popular research topics for experimentalists and theorists alike.

Particle physicists have been early adopters of machine learning methods - the 'Artificial Intelligence in High-Energy and Nuclear Physics' workshop started in 1990! Initially, these methods were employed mostly on the experimental side in terms of event selections and data analysis. Over the years, their applications have spread to more theoretical works as well. For example, they have played a significant role in Monte Carlo simulations, which are very important for theorists studying different models.

Traditionally machine techniques can be roughly classified into being supervised or unsupervised. In this section, we provide a brief overview of how each category can be applied to different aspects of particle physics.

1.2.1 Supervised Learning

In supervised learning, the goal is to perform regression or classification tasks based on labelled data. To put it simply, imagine we have a bunch of data points of the form $(\vec{x}_i, y_i)$, the goal is to use to data to model a function $f : \vec{x} \mapsto y$. In the context of particle physics, one of the most typical use cases of supervised learning is 'b-tagging', where each data point corresponds to a collision event, with $\vec{x}$ corresponding to the various observables built from kinematic information, and y denotes whether the event contains a b-quark.

Traditionally, the inputs $\vec{x}$ are hand-crafted based on physically sound considerations (in a process commonly referred to as 'feature engineering'), and are

subsequently fed to different supervised models such as boosted decision trees and neural networks. However, with the advent of deep learning, the feature engineering stage has been automated so the the input data consists of much lower-level information. For example, it has been demonstrated that the resulting b-tagger from such an approach has about twice as much rejection rate at the same signal efficiency compared to traditional methods! [10] The implications of this advancement is significant: it means that the machine learning algorithms have learnt something physically useful that us physicists were unable to figure out. If one views this from another angle, perhaps it is possible to directly learn physics from experimental data!

Another application more directly related to looking for new physics involves anomaly detection - the idea of discerning events containing a new particle (the 'anomaly') from the Standard Model background. Typically, given a BSM model, one can generate Monte Carlo samples of collider events for both the signal and background, one can then train a specific model to detect new particles predicted by this model. Another direction of anomaly detection makes less assumption about the model, and is the subject of Chapter 4.

1.2.2 Unsupervised Learning

In contrast to supervised learning, in unsupervised learning we are not given the label y of the data. Instead, the goal is to learn something useful about the data distribution itself. The two most common objectives are clustering and density estimation.

In clustering tasks, one seeks to group the data into different categories shar-

ing similar properties. In the context of collider experiments, suppose one wants to discern two underlying processes which are not modelled very well theoretically. In such a case it is inadvisable to rely on Monte Carlo samples for use in supervised learning [11]. Instead, one can first perform clustering of events which can then be used for more downstream analysis. This way, inaccuracies due synthetic data can be mitigated.

For density estimation tasks, the goal is to model the distribution from which the data is drawn, and perhaps also generate new data from the learned distributions. This is very useful both experimentally and theoretically. Experimentally, it can help us better model the Standard Model background, which is essential in various searches of new particles. Theoretically, it can help reduce resources needed to perform Monte Carlo simulations. In particle physics, the Monte Carlo simulation processes involve a lot of computationally-intensive steps - from matrix element generations all the way down to parton showering and detector response simulations. One way unsupervised learning can speed up this process is to first generate Monte Carlo samples as one would traditionally, and then build a surrogate model of the underlying data distribution. One can then use the learned distribution to sample new events. [12]

Another way unsupervised learning can aid Monte Carlo simulations lies in directly generating parton level events from matrix elements, where by 'generating' we mean sampling from the (normalized) matrix element function. In order to improve efficiency of the generator, importance sampling is employed, in which an algorithm seeks to approximate the given distribution function. Traditionally the VEGAS algorithm [13] (and its variants) have been the backbone of many Monte Carlo simulation software in particle physics. However, one sig-

nificant drawback of this algorithm is that it is not very efficient whenever the distribution function has a complicated shape, which happens more and more frequently as one increases the number of final state particles. In view of this issue, there has been proposals of replacing/augmenting the existing generating algorithm with more modern machine learning models such as normalizing flow or generative adversarial network [14, 15, 16]. These methods have been shown to work wonderfully where the traditional algorithm fails, improving both efficiency and accuracy of Monte Carlo simulations.

CHAPTER 2

DARK MATTER AS A SOLUTION TO MUONIC PUZZLES

2.1 Introduction

Observational evidence for the existence of dark matter (DM) is overwhelming. While DM comprises most of the matter in today's universe, and contributes about 20% of the total energy density, there is no known elementary particle that can account for it. Many candidate theories have been proposed, extending the Standard Model (SM) of particle physics to include one or more dark matter particles. In many theories, DM particles have potential experimental or observational signatures going beyond the purely gravitational effects that have been observed. However, no non-gravitational signature of DM has been conclusively established so far.

In this paper, we propose that dark matter particles are directly responsible for explaining a long-standing puzzle in particle physics, the proton charge radius anomaly. The value of the proton charge radius measured using Lamb shift in muonic hydrogen [17, 18] does not agree, at about 5σ level, with the value obtained from electron-proton scattering and electron hydrogen spectroscopy [19, 20]. A similar discrepancy was observed in muonic deuterium [21]. We note that some recent data on ep scattering [22] and H spectroscopy [23, 24] are in better agreement with muonic hydrogen results, while others [25] confirm the discrepancy. At present, the sources of disagreement among electronic experiments are not understood, and in this paper, we take the point of view that the anomaly is real and demands an explanation in terms of Beyond-the-SM (BSM) phenomena. (For previous works that proposed BSM explanations of the proton

charge radius puzzle, see *e.g.* [26, 27, 28, 29, 30].)

The key ingredient of our proposal is the idea that loop diagrams involving light dark matter states can induce a new "quantum" force between SM particles [31, 32]. Since the flavor structure of DM couplings to the SM is unconstrained, it is plausible that this new force may be felt by muons but not electrons. As will be shown below, such flavor-dependent quantum force can account for the proton charge radius puzzle, without conflict with any existing experimental constraints. At the same time, the dark matter particle responsible for this force can be a thermal relic consistent with the measured cosmological DM abundance, as well as with all known bounds on DM properties. The DM particle is predicted to have a mass of about 1 MeV, an interesting range from the point of view of direct detection experiments.

Another prominent experimental anomaly involving the muon is the anomalous magnetic moment a_μ, whose measured value differs from the SM prediction by about 3σ [33, 34, 35]. While non-perturbative SM contributions may account for some or all of this discrepancy, in this paper we take the point of view that it is real and requires a BSM explanation. It turns out that in our model, the effect of DM loops on a_μ is subdominant to the shift induced by the mediator particles which connect DM and SM sectors. We show that this shift can indeed account for the observed discrepancy, and demonstrate a set of parameters for which a_μ and proton charge puzzles are simultaneously solved, DM particle is a thermal relic with correct relic density, and all experimental and observational constraints are satisfied. It is remarkable that the simple model presented here can account for such a diverse set of data pointing to physics beyond the SM.

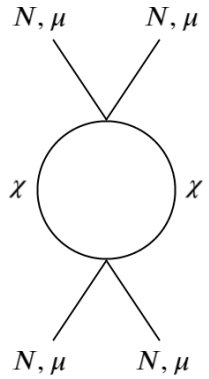


Figure 2.1: One loop diagram involving exchange of dark matter particle χ that induces a new force between muons and protons.

2.2 Model

We introduce a DM field χ, a Dirac fermion with no SM gauge charges. DM is coupled to the SM via two real-scalar mediator fields, a leptophilic mediator X and a leptophobic mediator X'. Both mediators are also SM gauge singlets. At energies below the QCD confinement scale, where all physics relevant for this study takes place, the interaction Lagrangian is given by

$$\mathcal{L}_{\rm int} = -g\bar{\mu}\mu X - (g'_p\bar{p}p + g'_n\bar{n}n)X' - y\bar{\chi}\chi X - y'\bar{\chi}\chi X' \,, \tag{2.1}$$

where μ, p and n are muon, proton and neutron fields, respectively. Throughout this paper, we will study the regime in which both mediators are much heavier than the DM particle, $m_X, m_{X'} \gg m_\chi$, and can be integrated out, leading to an effective Lagrangian

$$\mathcal{L}_{\rm eff} = -\frac{yg}{m_X^2}\bar{\chi}\chi\bar{\mu}\mu - \frac{y'g'_p}{m_{X'}^2}\bar{\chi}\chi\bar{p}p - \frac{y'g'_n}{m_{X'}^2}\bar{\chi}\chi\bar{n}n + \ldots \tag{2.2}$$

The quantum force between proton and muon, arising from the diagrams in Fig. 2.1, provides a new contribution to the Lamb shift in muonic hydrogen, resolving the proton charge radius puzzle.

A few comments are in order. For simplicity, we assumed that the mediator X (and therefore the DM) couples to muon but not to the electron. While the flavor-dependent nature of this coupling is crucial to resolving the proton charge radius puzzle, a non-zero value of electron coupling (for example, a plausible scenario in which $g_\ell \propto m_\ell$) can be introduced without altering the basic picture. Further, X' is generically expected to couple to pions and other mesons. We do not include such couplings since they would play no role in our analysis. Finally, while the interactions in Eq. (2.1) are sufficient to explain the proton charge radius and the a_μ anomaly, requiring that χ be a thermal relic with observed cosmological abundance necessitates an additional interaction involving neutrinos ν:

$$\Delta\mathcal{L}_{\rm int} = -\lambda\bar{\nu}\nu X, \tag{2.3}$$

if neutrinos are Dirac, or its Majorana counterpart. This interaction can arise from the operator $X(HL)^2$ above the weak scale, and its strength is *a priori* unrelated to the coupling of X to charged muons which arises from the operator $X(HL)\mu_R$.

We also note that if X and X' were replaced with a single mediator, coupled to both muons and quarks, the dominant new physics effect in muonic hydrogen would come from a tree-level exchange of the mediator, rather than the DM-induced quantum force. To leading order, this model would in fact be identical to that already considered in [30].

m_χ	m_X	$m_{X'}$	g	g'_p	g'_n	y	y'	λ
9 MeV	25 MeV	45 MeV	$4.5 \cdot 10^{-4}$	0.01	0.01	0.5	0.5	10^{-4}

Table 2.1: Model parameters at the benchmark point.

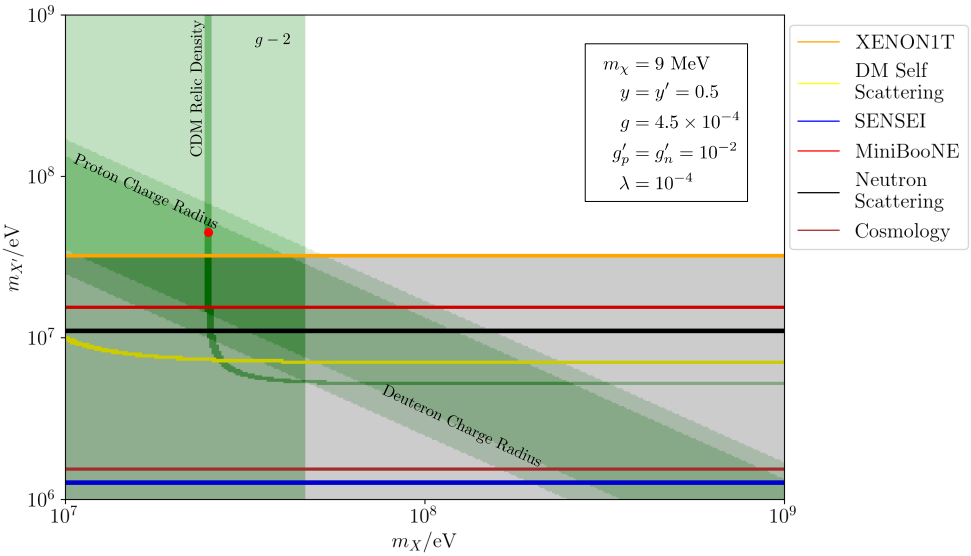


Figure 2.2: Fit to experimental data indicating non-SM physics (green) in the plane of mediator particle masses, m_X and $m_{X'}$, with the other parameters fixed to the values listed in Table 2.1. Relevant experimental and observational constraints on DM and mediator particles are also shown; the shaded areas are ruled out.

2.3 Results

Using the above model, we performed a fit to relevant experimental data and observational constraints on DM properties. The results of the fit are summarized in Figs. 2.2, 2.3 and 2.4. The model can explain the proton charge radius puzzle, the a_μ anomaly, and the observed DM relic density, while maintaining consistency with all known experimental and observational constraints. A sample benchmark point in the model parameter space which satisfies these requirements is shown in Table 2.1. Some details of the analysis are presented below.

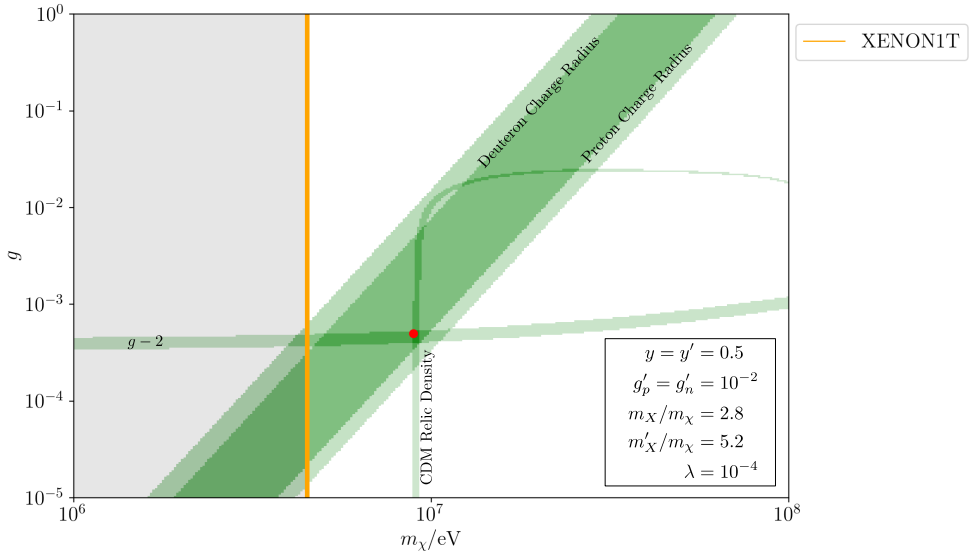


Figure 2.3: Fit to experimental data indicating non-SM physics in the plane of DM particle mass m_χ and the leptophilic mediator coupling to muons, g, with the other parameters fixed to the values listed in Table 2.1.

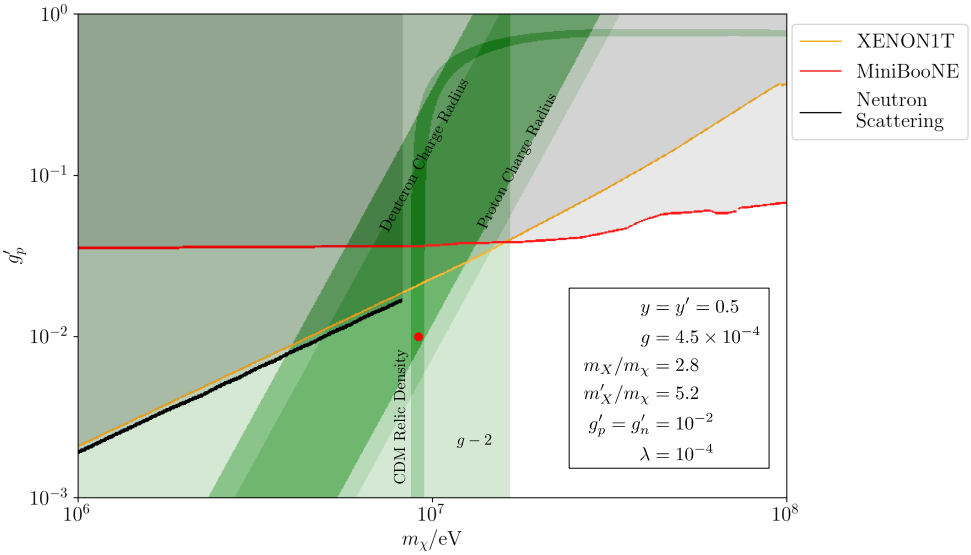


Figure 2.4: Fit to experimental data indicating non-SM physics in the plane of DM particle mass m_χ and the leptophobic mediator coupling to protons, g'_p, with the other parameters fixed to the values listed in Table 2.1. Relevant experimental and observational constraints on DM and mediator particles are also shown.

Proton Charge Radius: The fact that the proton has a finite size (with radius r_p) introduces shifts in energy levels of hydrogen-like atoms [17]. In particular, there would be a change in the energy difference between the $2S$ and $2P$ levels, *i.e.* Lamb shift ΔE_{Lamb}. By measuring ΔE_{Lamb}, one is able to deduce the value of r_p. In our model, the extra non-SM contribution to the Lamb shift in muonic hydrogen arises at one loop from the diagram in Fig. 2.1. In the non-relativistic limit, this interaction can be captured by a potential between protons and muons [32]:

$$V(r) = -\frac{3}{4\pi^3 r}\left(\frac{yg}{m_X^2}\right)\left(\frac{y'g'_p}{m_{X'}^2}\right)\frac{m_\chi^2}{r^2}K_2(2m_\chi r), \tag{2.4}$$

where K_2 is the modified Bessel function of the second kind of order 2. As a result, there is a new contribution to ΔE_{Lamb} in muonic hydrogen, given by

$$\begin{aligned}\Delta E_{\text{L}} &= \langle 2S|\, V\, |2S\rangle - \langle 2P|\, V\, |2P\rangle \\ &= -\frac{3}{8\pi^3}\frac{yy'gg'_p}{m_X^2 m_{X'}^2}\frac{m_\chi^2}{a^3}\,\mathcal{J}(x_0, a)\,.\end{aligned} \tag{2.5}$$

Here

$$\mathcal{J}(x_0, a) \;=\; \int_{x_0}^{\infty} \mathrm{d}x\, \frac{6\,(1-x)+x^2}{6x} e^{-x} K_2(2m_\chi a x), \tag{2.6}$$

a is the Bohr radius of muonic hydrogen, and $x_0 \approx (a\Lambda_{\text{QCD}})^{-1}$ is the short-distance cutoff corresponding to the breakdown of the effective field theory description in Eq. (2.1) at length scales below $\mathcal{O}(\Lambda_{\text{QCD}})$. We do not include the additional contribution to the Lamb shift from length scales below r_0, which can only be calculated within a specific UV completion of Eq. (2.1). To estimate the associated theoretical uncertainty, we vary the value of the cutoff by a factor of two around the assumed central value $(a\Lambda_{\text{QCD}})^{-1}$. This uncertainty is reflected in the width of the "proton charge radius" band in Figs. 2.2—2.4. Our fit assumes that this effect fully accounts for the "extra" Lamb shift measured in the

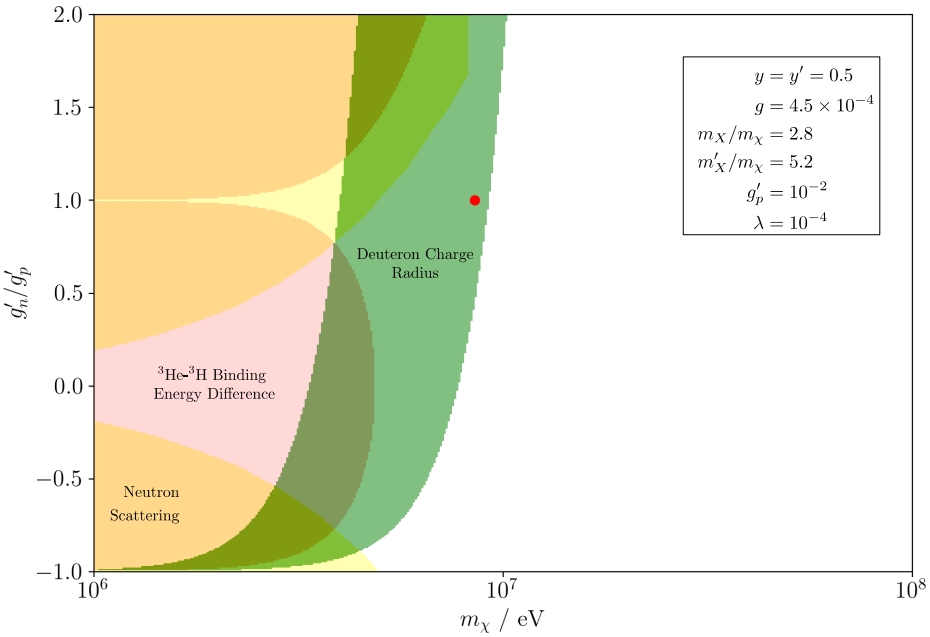


Figure 2.5: Fit to muonic deuterium data, as well as constraints from nuclear physics experiments, in the plan of the mediator coupling ratio g'_n/g'_p and the DM particle mass m_χ. The other parameters are fixed to the values listed in Table 2.1.

muonic hydrogen, compared to the baseline value inferred from electronic hydrogen and ep scattering: $\Delta E_{\rm L} = -0.307(56)$ meV. This is appropriate since in our model the electron does not couple to DM and hence is unaffected by the new force.

Muonic Deuterium: After the proton charge radius puzzle was discovered, there has been interest in performing similar experiments with other muonic atoms, in particular muonic deuterium μD. It was reported that the deuteron charge radius r_d extracted from μD shows similar discrepancy from world-averaged CODATA 2014 value [21], [20]. When comparing against spectroscopic values of r_d that involve deuterium only (*i.e.* independent of r_p), this discrepancy is reduced to 3.6σ [36], but is still statistically significant. We therefore

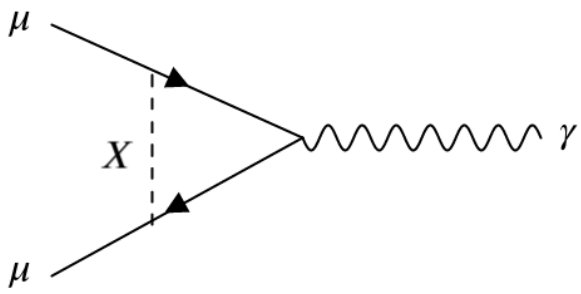


Figure 2.6: Contribution to muon anomalous magnetic moment due to the mediator particle X.

require that our model produces an extra contribution to Lamb shift in μD [21]: $\Delta E_{\rm L}^{\mu D} = -0.438(59)$ meV. The Lamb shift in μD is due to dark matter-mediated quantum force, and is given by

$$\Delta E_{\rm L}^{\mu D} = -\frac{3}{8\pi^3}\frac{yg}{m_X^2}\frac{y'\left(g'_p+g'_n\right)}{m_{X'}^2}\frac{m_\chi^2}{a_{\mu D}^3}\,\mathcal{J}(x_{0D},a_{\mu D})\,, \tag{2.7}$$

where $a_{\mu D}$ is the Bohr radius of the muonic deuterium system, r_D is the deuteron radius, and $x_{0D} = r_D/a_{\mu D}$. The shift depends on both g'_n and g'_p, while various combinations of these couplings are constrained by nuclear physics experiments (see below). As shown in Fig. 2.5, isospin-preserving coupling $g'_p = g'_n$ is consistent with both the deuteron charge radius and nuclear physics constraints.

Muon Anomalous Magnetic Moment: The leading new contribution to a_μ is given by the one-loop diagram shown in Fig. 2.6. Note that this contribution is independent of the dark matter candidate χ itself, which only enters at the two-loop level. The shift in a_μ is given by

$$\Delta a_\mu = \frac{2g^2}{(4\pi)^2}\int_0^1 \mathrm{d}x\ \frac{(1-x)^2(1+x)}{(1-x)^2+x(m_X/m_\mu)^2}. \tag{2.8}$$

In our fit we assume that this effect fully accounts for the experimental discrepancy $\Delta a_\mu = 287(80)\times 10^{-11}$, within 2 standard deviations.

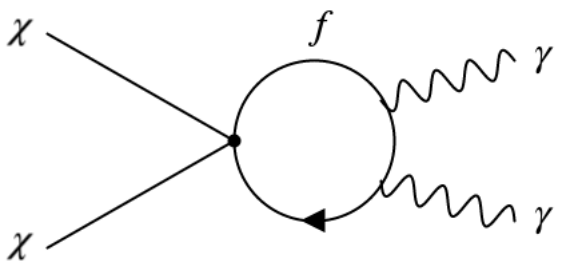


Figure 2.7: Dominant annihilation channel for χ at the time of freeze-out, if coupling to neutrinos are absent.

Relic Density: We assume that χ is a thermal relic and that it accounts for all of the observed cosmological DM abundance, $\Omega h^2 = 0.120 \pm 0.001$ [37]. With interactions in Eq. (2.1) and $m_\chi \sim \mathcal{O}(\text{MeV})$, the dominant DM annihilation channel at the time of freeze-out is $\chi\chi \to 2\gamma$, see Fig. 2.7. The leading (p-wave) contribution to the cross section in the non-relativistic limit is given by

$$\sigma_0 = \frac{3e^2}{32\pi^3} \left| m_\mu I(\tau_\mu) \frac{yg}{m_X^2} + m_p I(\tau_p) \frac{y' g'_p}{m_{X'}^2} \right|^2 . \tag{2.9}$$

Here the loop function $I(\tau_f)$ for a fermion f is defined by

$$I(\tau_f) = \int_0^1 \mathrm{d}x \int_0^{1-x} \mathrm{d}y \; \frac{1-4xy}{\tau_f - xy} \tag{2.10}$$

with $\tau_f = m_f^2/s$, where $\sqrt{s}$ is the center-of-mass energy of the scattering process. σ_0 can be used to compute the relic density of χ [38] by solving the Boltzmann equation numerically.

We find that the $\chi\chi \to 2\gamma$ cross section is too small to provide the observed relic density for model paramaters required to fit the proton charge radius and a_μ anomalies. A simple solution is to consider an additional annihilation channel, $\chi\chi \to \nu\nu$, via the interaction in Eq. (2.3). The cross section is given by $\sigma_0 = \frac{3}{4\pi}\left(\frac{y\lambda}{m_X^2}\right)^2 m_\chi^2$. Since this final state arises at tree level, it naturally dominates over the 2γ channel. With this addition, all three constraints can be satisfied simultaneously, see Figs. 2.2—2.4.

In addition to fits to the data indicating deviations from the SM, a number of constraints from data and observations consistent with the SM have to be taken into account:

Dark Matter Self-Scattering: Tree-level exchanges of mediator particles X and X' induce DM short-range self-interactions of the form

$$\mathcal{L}_{\text{self}} = -\left(\frac{y^2}{m_X^2} + \frac{y'^2}{m_{X'}^2}\right)(\bar{\chi}\chi)^2. \tag{2.11}$$

DM self-scattering cross sections at low velocities are bounded by observations of halo shapes [39, 40]: $\sigma_T/m_\chi \lesssim 1\ \text{cm}^2/\text{g}$, where σ_T is the momentum-transfer cross section defined by

$$\sigma_T \equiv \int \mathrm{d}\Omega\ \frac{\mathrm{d}\sigma_{\chi\chi\to\chi\chi}}{\mathrm{d}\Omega}(1-\cos\theta). \tag{2.12}$$

This constraint translates into an upper bound on the DM mass:

$$m_\chi \lesssim 8\pi\left(\frac{y^2}{m_X^2} + \frac{y'^2}{m_{X'}^2}\right)^{-2} \times 4600\ (\text{GeV})^{-3}. \tag{2.13}$$

Dark Matter Direct Detection: Direct detection of dark matter in the MeV mass range has been the subject of much interest recently [41, 42]. Most techniques rely on detection of DM scattering on electrons. In our model, this channel is not available, since by construction DM does not couple to electrons. However, scattering on a nucleon can occur, with cross section

$$\sigma(\chi N \to \chi N) = \frac{1}{\pi}\left(\frac{g'_N y'}{m_{X'}^2}\right)^2 \left(\frac{m_\chi m_N}{m_\chi + m_N}\right)^2, \tag{2.14}$$

where $N = n$ or p. For our benchmark point, this cross section is about $6 \times 10^{-32}\ \text{cm}^2$. This is about half an order of magnitude below the strongest current constraint from non-observation of signal due to energetic DM component generated through collisions with cosmic rays [43, 44], shown by the

XENON−1T curve in Fig. 2.4. MeV-scale dark matter can also be detected using the Migdal effect [45, 46, 47, 48, 49]. However, the recent results from SENSEI collaboration [50] are not yet sensitive enough to constrain our model.

Early-Universe Cosmology: Measurements of Cosmic Microwave Backgound (CMB) place strong constraints on possible reionization due to DM annihilations [51]. However in our model, χ annihilation proceeds in p-wave, and thus not subject to this constraint. CMB data together with the Lyman-alpha forest flux power spectrum from the Sloan Digital Sky Survey constrains the elastic scattering of DM on baryons [52, 53]. This constraint is shown by the curve labeled "cosmology" on Fig. 2.2. In addition, a scenario where DM freeze-out occurs after neutrinos decouple from the rest of the SM plasma is constrained by the CMB bound on $\Delta N_{\rm eff}$, since in this case the neutrino temperature at recombination would be raised relative to T_γ by the entropy transferred from the DM. This argument imposes a lower bound $m_\chi \gtrsim$ a few MeV [54, 55]. A similar bound is imposed by the success of Big-Bang Nucleosynthesis (BBN) [56, 57, 58].

Dark Matter Mediator Searches: In addition to direct searches for dark matter particles, there are many experiments looking for mediator particles produced at colliders or in a fixed-target setup [41, 42]. While most analyses present the results in terms of bounds on dark photons, which couple to both leptons and quarks proportional to their electric charges, the interpretation of interest to us is in terms of leptophilic or leptophobic mediators. Leptophilic mediator searches rely on their production via their interaction with electrons, making them insensitive to our model. (An exception is the recently reported NA64 search for scalars produced through their coupling to photons [59]. However,

this search does not place relevant bounds on our model, since mediator couplings to photons are loop-suppressed.) Thus we only consider bounds from leptophobic mediator searches. With sub-MeV dark matter, the most stringent bound currently comes from the MiniBooNE experiment [42, 60], which places a bound on the parameter Y related to dark matter annihilation cross section (see referenced papers for the precise definition). In our model, this parameter is given by

$$Y = \left(\frac{g'_p}{e}\right)^2 \frac{{g'_p}^2}{4\pi} \left(\frac{m_\chi}{m_{X'}}\right)^2. \tag{2.15}$$

Note that the mediator X' in our model is a scalar, while the MiniBooNE bounds were derived using a spin-1 mediator. An order-one correction to the bound may arise due to the differing kinematic acceptances and spin factors in the two cases, but we do not expect it to affect our conclusions.

Nuclear Interactions: Leading non-SM contributions to nucleon-nucleon potential are given by

$$\begin{aligned} V_{N_1 N_2} &= -\frac{g'_{N_1} g'_{N_2}}{4\pi r} e^{-m_{X'} r} \\ &\quad -\frac{3}{4\pi^3}\left(\frac{y'}{m_{X'}^2}\right)^2 g'_{N_1} g'_{N_2} \frac{m_\chi^2}{r^3} K_2(2m_\chi r). \end{aligned} \tag{2.16}$$

This extra potential can be probed at various nuclear physics experiments. It turns out that the second term in $V_{N_1 N_2}$, arising from the DM-loop exchange, is the most relevant for our analysis, since $1/r_{\rm exp} \ll m_\chi \ll m_{X'}$, where $r_{\rm exp}$ is the length scale probed by the experiments. The relevant constraints are summarized in Fig. 2.5.

The binding energy difference between ^{3}He and ^{3}H has been well-established to be caused by Coulomb force and charge asymmetry of nuclear forces [61], [62], [63]. In order not to spoil this agreement, the non-SM con-

tribution is required to be less than 30 keV [30]. It is worth noting that this contribution is proportional to $\left(g_p'^2 - g_n'^2\right)$, and vanishes in the isospin limit, see Fig. 2.5.

The charge-independence breaking (CIB) scattering length is defined as

$$\Delta a = \frac{1}{2}(a_{nn} + a_{pp} - 2a_{np}), \tag{2.17}$$

where $a_{N_1N_2}$ is the scattering length between two nucleons N_1 and N_2. Experimental and theoretical values for Δa are known to be 5.64 ± 0.60 fm [64] and 5.6 ± 0.5 fm [65] respectively. Our model gives an extra contribution

$$\delta a^{\text{th}} = -\frac{3m_N}{\pi^2}\left(\frac{y'}{m_{X'}}\right)^2 m_\chi^2 \left(g_p' - g_n'\right)^2 \log\left(\frac{m_\chi^2}{\Lambda^2}\right), \tag{2.18}$$

where m_N is the nucleon mass and Λ is the cut-off scale of the effective theory. We require $\delta a^{\text{th}} < 1.6$ fm to maintain agreement between experimental and theoretical values of Δa at 2σ level. This constraint again vanishes in the isospin limit.

Scattering lengths between cold neutrons and nuclei can be measured by different methods, such as Bragg diffraction and the transmission method [66]. In our model, this scattering length is given by Eq. (2.18) by replacing $\left(g_p' - g_n'\right)^2$ with $g_n'^2$ or $g_n' g_p'$. By comparing the scattering lengths measured by different methods, bounds can be placed on contributions from non-contact operators. The detailed analysis can be found in [67]. Neutron scattering places the most restrictive nuclear-physics bound on the model in the isospin-symmetric limit [32].

2.4 Conclusions

In this paper we presented a simple model of MeV-scale Dirac fermion dark matter χ coupled to nucleons and muons, but not electrons. The quantum force due to χ loops is responsible for resolving the proton (and deutron) charge radius puzzles, while a scalar particle introduced to mediate DM-muon interactions can account for the discrepancy between the SM prediction for the anomalous magnetic moment of the muon and the measured value. If an additional interaction between χ and neutrinos is postulated, the former can be a thermal relic responsible for all of the observed DM abundance. We verified the existence of a region in the paramater space of the model where all of the above features are obtained simultaneously, while all known experimental and observational constraints on light DM and mediators are satisfied.

New experimental data will soon be available that will test our model on various fronts. Numerous efforts, *e.g.* `MUSE` experiment [68], are under way that will hopefully clarify the status of the proton charge radius puzzle. Anomalous magnetic moment of the muon measurement will be improved by the Fermilab Muon $g-2$ experiment [69]. Experimental exploration of MeV-scale dark matter and dark sectors is an active and expanding area of research [41, 42]. While our model has some amount of freedom in parameter choices, the available parameter space is finite and it is likely that this idea will be tested conclusively by the upcoming experiments in the near future.

CHAPTER 3

DARK Z AT THE INTERNATIONAL LINEAR COLLIDER

3.1 Introduction

The discovery of the Higgs boson at the Large Hadron Collider (LHC) strongly motivates the construction of a Higgs factory: an electron-positron collider with center-of-mass energy in the 250–500 GeV range. Several proposals are currently under consideration [70, 71, 72, 73, 74, 75, 76], and hopes are high that at least one of them will be realized in the next one to two decades. While the main motivation for these colliders is a comprehensive and precise study of the Higgs, they can also search for physics Beyond the Standard Model in ways complementary to the LHC. It is this aspect of the physics program that we focus on in this paper.

Additional gauge interactions are ubiquitous in well-motivated extensions of the Standard Model (SM), such as grand unified theories and string theory. A particularly simple example is an extra U(1) gauge group. When SM fermions are charged under the extra U(1), the collider phenomenology of the corresponding gauge boson, the Z', is very well-studied. Unfortunately, the LHC sets a lower limit on the Z' mass in the multi-TeV range, precluding the possibility of its direct production at next-generation e^+e^- colliders; although the Z' may still be detected through its indirect effects. We focus on an equally well-motivated scenario that has received less attention so far: the SM fermions are *not* directly charged under the extra U(1). The new gauge boson, called a dark Z, can still interact with the SM through kinetic mixing with hypercharge [77, 78, 79, 80], as well as mass mixing induced by an extended Higgs sector [81, 82]. (If only

hypercharge mixing is present, this boson is known as a dark photon.) Such interactions are not constrained by charge quantization, and can be significantly weaker than SM gauge couplings. We show that the current constraints from the LHC and other experiments are consistent with models in which the dark Z is accessible to a direct search at a realistic 250–500 GeV e^+e^- collider. We use the International Linear Collider (ILC) design parameters as the collider benchmark in our study. Further, we extrapolate the current LHC sensitivity to the 3 ab^{-1} data set expected at the high-luminosity (HL)-LHC to show that there is a large region of parameter space where the e^+e^- collider can make the first observation of the dark Z.

If a particle consistent with a dark Z is discovered, the next step would be to measure its properties, such as mass and couplings to the SM. An e^+e^- collider can perform these measurements with unparalleled precision. Unlike the LHC, a lepton collider can separately determine the dark Z couplings to left-handed and right-handed SM fermions. These chiral couplings are crucial for discriminating between different possible models of the new particle. We use a benchmark dark-Z model as an illustrative example to show that a relatively short (one or two months) dedicated run where the e^+e^- collision center-of-mass energy is matched to the dark Z mass can measure the couplings with percent-level errors, unambiguously establish the presence of parity violation in the dark-Z couplings, and probe the underlying sources of dark-Z interactions with the SM. We show that the e^+e^- collider may distinguish between pure kinetic mixing versus models with mass mixing through exotic Higgs fields. The program of on-resonance measurements of the dark Z that we outline is a "dark" counterpart to the precision electroweak program in the 1990s at LEP and SLC. However, there are important differences that need to be properly taken into

account, such as the very small intrinsic width of the dark Z compared to the SM Z.

Before proceeding, let us make the following observation. While dark gauge bosons have attracted much attention in recent years, a vast majority of this work has been focused on mass scales in the MeV–GeV range [41, 42, 83]. In contrast, we investigate much heavier dark gauge bosons with masses around the weak scale. From a theoretical point of view, this mass is a free parameter and any scale is equally well motivated. Part of the motivation for sub-GeV dark photons comes from the observation that they can serve as a mediator between the SM and a light dark matter particle, and produce a thermal relic density of dark matter consistent with observations. The dark Z studied here cannot play this role in a minimal model. However, slightly more complicated models of the dark sector (*e.g.* involving a state directly coupled to the dark Z which decays to the stable dark matter state as in the super-WIMP mechanism [84]) can easily accommodate the observed dark matter with a weak-scale dark Z mediator. Thus, a comprehensive exploration of the dark sector has to include the possibility of weak-scale dark gauge bosons, and energy-frontier colliders offer the only way to access this regime experimentally.

The rest of the paper is organized as follows. Section 3.2 describes the theoretical model underlying our study. In Section 3.3 we discuss the current constraints on the model from precision electroweak fits and direct searches at the LHC, as well as projected reach of future searches at the HL-LHC and the ILC. The analysis of this section extends previous studies of dark photon phenomenology at TeV-scale colliders [85, 86] to include the effects of mass mixing in the dark Z model. In Section 3.4, we explore the physics potential of a short, dedicated run

of the ILC with the center-of-mass energy at the dark Z resonance. We conclude in Section 4.5, while the Appendix contains some of the details of our model and its ultraviolet completion, as well as explicit formulas for the "precision dark Z" observables.

3.2 Three-Parameter Dark Z Model

We present a model of a new spin-1 particle with both kinetic and mass mixing. The parameters are the dark Z mass $m_{A'}$, the kinetic mixing ε, and a dimensionless mass mixing parameter κ. The theory is a limit of a type-1 two-Higgs doublet model augmented with a third electroweak-neutral dark Higgs. Refs. [87, 82] present an alternative two-Higgs doublet model that realizes the same low-energy theory. While our model is less minimal, it cleanly separates the dimensionful parameters that control the dark Z mass and mass mixing. We provide additional details of the effective theory in Appendix A.1.

3.2.1 Kinetic Mixing: ε

We assume an Abelian 'dark' gauge symmetry $\mathrm{U(1)_d}$ with gauge boson A'_μ. Standard Model particles not uncharged under this symmetry. The gauge kinetic terms contain a mixing term with the hypercharge boson B_μ,

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \frac{\varepsilon}{2c_W}B_{\mu\nu}A'^{\mu\nu} - \frac{1}{4}A'_{\mu\nu}A'^{\mu\nu} \,. \tag{3.1}$$

This mixing may be generated by loops of heavy particles charged under both Abelian groups [77, 78, 79, 80]. We normalize ε by the cosine of the Weinberg angle $c_W = \cos\theta_W$ so that the kinetic mixing with the electromagnetic

field strength is ε. In what follows, we adopt the notation $s_W = \sin\theta_W$ and $t_W = \tan\theta_W$. We diagonalize the gauge kinetic term by transforming

$$B \to B + \frac{\varepsilon}{c_W}\mathcal{N}A' \qquad\qquad A' \to \mathcal{N}A' \,. \tag{3.2}$$

Henceforth we set the normalization factor $\mathcal{N}^{-2} = 1 - \varepsilon^2/c_W^2 \approx 1$ because the correction is of higher order in ε compared to the accuracy required for this study.

3.2.2 Symmetry Breaking and Mass–Mixing: κ

The $\mathrm{U}(1)_\mathrm{d}$ and electroweak symmetry are broken by separate vacuum expectation values (vevs), v_d and v_EW, as well as by a mixed vev, $v_\mathrm{mix} \ll v_\mathrm{EW}$. A benchmark UV model carries an extended Higgs sector with the following quantum numbers:

$$H_\mathrm{EW} = \square_{\frac{1}{2},0} \qquad\qquad H_\mathrm{mix} = \square_{\frac{1}{2},q_\mathrm{d}} \qquad\qquad H_\mathrm{d} = 1_{0,-q_\mathrm{d}} \,, \tag{3.3}$$

where $\square$ or 1 refer to an $\mathrm{SU}(2)_\mathrm{L}$ doublet or singlet and the subscripts refer to the $\mathrm{U}(1)_Y$ hypercharge and $\mathrm{U}(1)_\mathrm{d}$ dark charges. We assume that the extended Higgs sector potential has large enough quartic couplings relative to the dark gauge coupling that the non-Goldstone modes decouple from the dark Z dynamics; see Appendix A.2. (Ref. [88] studied the phenomenology of a model with a non-decoupled scalar.)

The vevs of the Higgs fields give produce gauge boson mass terms,

$$\mathcal{L} \supset \left[\left(\frac{1}{2}g'B - \frac{1}{2}gW^3\right)\frac{v_\mathrm{EW}}{\sqrt{2}}\right]^2 + \left[\left(\frac{1}{2}g'B - \frac{1}{2}gW^3 + g_\mathrm{d}q_\mathrm{d}A'\right)\frac{v_\mathrm{mix}}{\sqrt{2}}\right]^2 + \left[g_\mathrm{d}q_\mathrm{d}A'\frac{v_\mathrm{d}}{\sqrt{2}}\right]^2 \,. \tag{3.4}$$

The v_{mix} term is a mass mixing between the dark Z and Standard Model Z. The net order parameter of electroweak symmetry breaking is $v^2 = v_{\text{EW}}^2 + v_{\text{mix}}^2 = (246\text{ GeV})^2$. The eigenvalues of the mass-squared matrix are the dark and Standard Model Z squared masses, $m_{A'}^2$ and $m_Z^2 = (91\text{ GeV})^2$. The following parameterization of mass mixing is particularly convenient,

$$\kappa = 2q_{\text{d}}\frac{g_{\text{d}}}{\sqrt{g'^2+g^2}}\frac{v_{\text{mix}}^2}{v^2} \qquad \text{mass mixing parameter} \tag{3.5}$$

We will assume that $\kappa \ll 1$, as is in fact required by precision electroweak observables, discussed below. Moreover, we treat κ and ε as small parameters of the same order. Defining $c_\delta = \cos\delta$ and $s_\delta = \sin\delta$, the transformation to the mass eigenstates is then

$$\begin{pmatrix} A' \\ Z \end{pmatrix} = \begin{pmatrix} c_\delta & s_\delta \\ -s_\delta & c_\delta \end{pmatrix} \begin{pmatrix} A' \\ Z \end{pmatrix}_{\text{mass}} \qquad \delta \approx -\frac{m_Z^2}{m_{A'}^2 - m_Z^2}\left(t_W\varepsilon - \kappa\right) \tag{3.6}$$

3.2.3 Interactions with Fermions

The mass mixing causes the dark and Standard Model Zs to couple to linear combinations of the weak neutral and electromagnetic currents. We ignore dark Z interactions to any low-mass dark sector states. If present, these simply reduce dark Z branching ratios to visible states. Standard Model fermions are assumed to have zero charges under the U(1)$_{\text{d}}$ gauge group. The couplings of an SM fermion ψ to the SM Z and its dark counterpart have the form

$$\mathcal{L} \supset A'_\mu \bar{\psi}\gamma^\mu\left(g^{\text{V}} + g^{\text{A}}\gamma^5\right)\psi + \frac{g}{c_W}Z_\mu\bar{\psi}\gamma^\mu\left(q_Z^{\text{V}} + q_Z^{\text{A}}\gamma^5\right)\psi\,, \tag{3.7}$$

where $q_Z^{\text{V,A}}$ are the ψ weak neutral charges. In the absence of any mixing, the these charges take their Standard Model values,

$$q_{Z,\text{SM}}^{\text{V}} = \frac{1}{2}T^3 - s_W^2 Q_{\text{EM}} \qquad\qquad q_{Z,\text{SM}}^{\text{A}} = -\frac{1}{2}T^3\,, \tag{3.8}$$

with $T^3 = \pm 1/2$ according to the $\mathrm{SU(2)_L}$ weight of the left-chiral ψ and Q_{EM} is the ψ electric charge. For example, an electron has $T^3 = -1/2$ and $Q_{\mathrm{EM}} = -1$.

Modified Z couplings. The rotation to mass eigenstates (3.6) shifts the charges of fermions with respect to the Standard Model values,

$$q_Z^{\mathrm{V}} = (c_\delta + \varepsilon t_W s_\delta)\, q_{Z,\mathrm{SM}}^{\mathrm{V}} + \frac{\varepsilon e}{g} c_W s_\delta\, Q_{\mathrm{EM}} \qquad q_Z^{\mathrm{A}} = (c_\delta + \varepsilon t_W s_\delta)\, q_{Z,\mathrm{SM}}^{\mathrm{A}}\,. \tag{3.9}$$

These corrections to the Standard Model charges are second order in $\varepsilon, \kappa \ll 1$.

Dark Z couplings. The A' vector and axial couplings to ψ are readily expressed in terms of the Standard Model Z charges,

$$g^{\mathrm{V}} = \varepsilon e Q_{\mathrm{EM}} - \frac{gr}{c_W} q_{Z,\mathrm{SM}}^{V} \qquad g^{\mathrm{A}} = \frac{-gr}{c_W} q_{Z,\mathrm{SM}}^{A} \qquad r \equiv \frac{m_Z^2 \kappa + m_{A'}^2 \varepsilon t_W}{m_{A'}^2 - m_Z^2}\,. \tag{3.10}$$

These couplings arise at the linear order in $\varepsilon, \kappa \ll 1$. Unlike its coupling to the electric current, the dark Z coupling to the weak neutral current depends on the relative masses of A' and Z in the factor r.

3.2.4 Interactions with Bosons

Dark Z bosons that are at least twice as heavy as the W boson may decay to W^+W^-. At higher masses, the Zh channel is also accessible. In models with only kinetic mixing, the branching ratios for these channels are always at least an order of magnitude smaller than those for $\ell^+\ell^-$ [89]. However, the presence of mass mixing dramatically changes these results and the diboson final states can dominate, see Fig. 3.1. To the best of our knowledge, the effect of these decay modes has not been addressed in past literature.

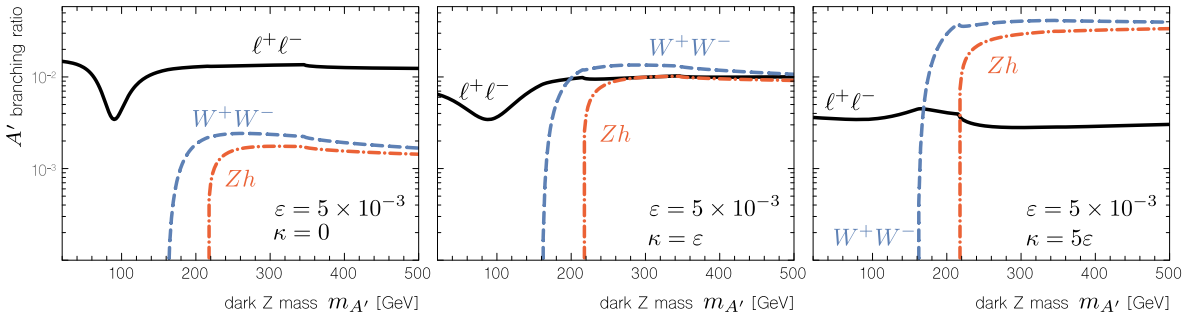


Figure 3.1: The branching ratios of A' decays into dilepton pairs $\ell^+\ell^-$ and diboson pairs W^+W^- and Zh for different values of mass mixing κ compared to kinetic mixing ε. For $\kappa \gtrsim \varepsilon$, the diboson branching ratios become large. For large $m_{A'}$, the diboson branching ratios converge as predicted by the Goldstone boson equivalence theorem.

The dark Z inherits its coupling to W^+W^- from its mixing with the W^3 component of the Standard Model Z, (3.6). This coupling is thus $-s_\delta$ times the ZW^+W^- coupling. The dramatic effect on the branching ratio for non-zero mass mixing, $\kappa \gtrsim \varepsilon$, can be understood as follows. Parametrically, $m_Z^2/m_{A'}^2 \ll 1$ in the region where the dark Z decay to gauge bosons is kinematically allowed. In the unitary gauge, the couplings of the dark Z to fermions and W^+W^- are of the same order, $\mathcal{O}(\kappa\, m_Z^2/m_{A'}^2)$. However the spin sum for W boson final state with momentum k contributes an enhancement of $k_\mu k_\nu/m_W^2 \sim (m_{A'}/m_W)^2$ for each W compared to the fermion final states. In the 't Hooft-Feynman gauge, the dark Z decay to Goldstone modes is dominant in the limit $m_{A'} \gg m_Z$. The dark Z coupling to charged Goldstones is contained in the H_{mix} kinetic term and is $\mathcal{O}(\kappa)$, without the additional suppression by $(m_Z^2/m_{A'}^2)$ compared to the fermion coupling.

The $A'Zh$ coupling depends on the mixing between the neutral CP-even Higgs-like states in H_{EW} and H_{mix}, an angle called α in the two-Higgs doublet literature. In our model, this mixing comes from integrating out the CP-even state in H_{d}. In the large $m_{A'}/m_Z$ limit, the Goldstone boson equivalence theo-

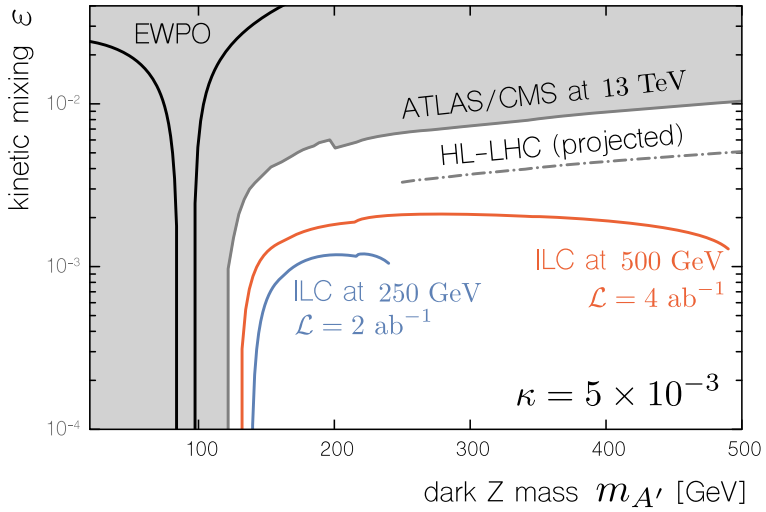


Figure 3.2: Discovery reach for a dimuon resonance search at the ILC (colored) compared to bounds from Drell–Yan production at CMS[92, 93] and ATLAS [94](gray shaded) for a benchmark value of the mass mixing, κ. The kink at 200 GeV is the transition between two separate recasted searches. We estimate the high-luminosity LHC reach by rescaling the luminosity to 3000 fb^{-1}.

rem tells us that we may replace the $W^{\pm}$ and Z final states by their eaten Goldstone bosons [90, 91]. In the limit where the heavy states are decoupled, see Appendix A.2, the eaten Goldstones and the low-energy CP-even state are components of a SU(2)$_L$ doublet and thus have the same couplings to the A'. Thus the branching ratios to W^+W^- and Zh approach each other at high dark Z mass, as shown in Fig. 3.1. There is no $A'ZZ$ coupling by SU(2)$_L$ symmetry and no Zhh coupling due to parity.

3.3 Experimental Constraints and the ILC Reach

This section explores the potential for a future e^+e^- collider, such as the ILC, to discover a dark Z. We begin by evaluating current constraints on the model

parameter space from precision electroweak observables and the LHC,[1] as well as estimating the likely reach of the HL-LHC. We then study the process $e^+e^- \to \gamma A'$, followed by the decay $A' \to \mu^+\mu^-$, at the ILC, together with the relevant Standard Model backgrounds. The results for a benchmark value of the mass mixing parameter are summarized in Fig. 3.2. The main conclusion is that the ILC can probe parameters well beyond the current bounds and the projected HL-LHC sensitivity. Our study focuses on the mass range $m_{A'} > 20$ GeV. Below this mass, the energy-frontier collider searches are less sensitive compared to fixed-target, beam-dump, and other low-energy experiments [83, 42, 41].[2]

3.3.1 Electroweak Precision Observables

The broad agreement of electroweak precision observables with their Standard Model values [99] limits the amount of mixing between new gauge bosons and the Standard Model Z. Earlier work used this to constrain a dark photon model with pure kinetic mixing [100, 85]. We generalize those analyses to the case of both kinetic and mass mixing. In our approximation, the tree-level shift in precision observables due to the dark Z must be within the measurement errors. We assume that the observables are uncorrelated so that the covariance matrix is diagonal. These approximations are sufficient for our purpose of comparing the sensitivity of these measurements to direct searches at colliders. We focus on the measurements made at the Z-pole and interpret the bounds as a constraint on the kinetic mixing $\varepsilon(m_{A'}, \kappa)$ as a function of the dark Z mass and mass mixing

[1] Deep inelastic scattering observations constrain the presence of a dark Z in a way that is agnostic to its decay modes. For visibly-decaying dark Z, the projected constraints from future electron–hadron colliders are subdominant to those from hadron colliders, both for unpolarized [95] and polarized [96] electron beams.

[2] We note in passing that experiments using the beam dumps of a high-energy e^+e^- collider such as the ILC have impressive sensitivity to lighter dark photons [97, 98].

parameter (3.5).

The leading-order predictions of the electroweak sector of the Standard Model contain three model parameters that need to be fixed by data. We establish a set of reference electroweak parameters: the weak mixing angle $s_W^2 = \sin^2\theta_W$, the electroweak vev v^2, and the Z coupling $g_Z^2 = g^2 + g'^2$. We fix the values of these parameters using the precise measurements of the fine structure constant, Fermi constant (from muon lifetime), and Z mass:

$$\hat{\alpha} = \frac{e^2}{4\pi} = \frac{g_Z^2 c_W^2 s_W^2}{4\pi} \qquad \frac{\hat{G}_F}{\sqrt{2}} = \frac{1}{2v^2} \qquad \hat{m}_Z^2 = \frac{g_Z^2 v^2}{4} \,. \tag{3.11}$$

Quantities with a caret (e.g. $\hat{\alpha}$) are measured quantities. One may invert (3.11) to obtain expressions for the reference parameters in terms of the measured inputs. In the presence of new weakly-coupled physics, the reference parameters are shifted relative to their Standard Model values. For example, $g_Z^2 = (g_Z^2)^{\rm SM}[1 + \mathcal{O}(\xi)]$, where ξ characterizes the new physics coupling. For our dark Z model, a useful definition is

$$\xi \equiv \varepsilon t_W + \kappa \ll 1 \,. \tag{3.12}$$

The order parameter of electroweak symmetry breaking, v^2, is measured from the muon lifetime and is unchanged from the Standard Model:

$$v^2 = \frac{1}{\sqrt{2}\hat{G}_F} \,. \tag{3.13}$$

In the dark Z model, the Standard Model Z boson mass is shifted, see (A.1), which then shifts the reference value of the Z coupling,

$$g_Z^2(\xi, m_{A'}^2) = 4\sqrt{2}\,\hat{G}_F \hat{m}_Z^2 \left[1 + 2\xi^2 \frac{\hat{m}_Z^2}{m_{A'}^2 - \hat{m}_Z^2} + \mathcal{O}(\xi^4)\right] \,. \tag{3.14}$$

This shift in g_Z^2, in turn, shifts the reference expression for the weak mixing angle

Observable	Expression to second order in ε, κ	
Z width	$\Gamma_Z = \sum_f^{u,d,e,\nu} N_f \Gamma_f$	$\Gamma_f \equiv \frac{m_Z}{12\pi} \frac{g^2}{\cos^2\theta_W} \left[\left(q_Z^V\right)^2 + \left(q_Z^A\right)^2 \right]_f$
hadronic cross section	$\sigma_{\text{had}} = \frac{12\pi}{m_Z^2} \frac{\Gamma_e \Gamma_{\text{had}}}{\Gamma_Z^2}$	$\Gamma_{\text{had}} \equiv 3\left(2\Gamma_u + 3\Gamma_{\text{d}}\right)$
lepton ratio	$R_\ell = \Gamma_{\text{had}}/\Gamma_e$	
quark ratios	$R_q = \Gamma_q/\Gamma_{\text{had}}$	for $q = u, d$
left–right asymmetry	$A_f = \left. \frac{2 q_Z^{\text{V}} q_Z^{\text{A}}}{\left(q_Z^{\text{V}}\right)^2 + \left(q_Z^{\text{A}}\right)^2} \right\vert_f$	for $f = e, u, d$
forward–backward asymmetry	$A_{\text{FB},f} = \frac{3}{4} A_e A_f$	for $f = e, u, d$

Table 3.1: Z-pole precision observables. The effective Z charges $q_Z^{\text{V,A}}$ are given in (3.9). N_f is the number of species for each fermion f accessible in Z decays; e.g. $N_u = 3$ colors $\times\, 2$ flavors, $N_d = 3 \times 3$, $N_e = 3$, $N_\nu = 3$. Γ_f is the associated partial width.

relative to the fine structure constant (3.11),

$$s_W^2(\xi, m_{A'}^2) = \left(1 - \hat{D}\right) \left[1 - \frac{4\hat{C}\xi^2}{\hat{D}(1-\hat{D})} \frac{\hat{m}_Z^2}{m_{A'}^2 - \hat{m}_Z^2} + \mathcal{O}(\xi^4) \right] \quad \hat{C} \equiv \frac{\pi\hat{\alpha}}{\sqrt{2}\hat{G}_F \hat{m}_Z^2} \,, \tag{3.15}$$

where $\hat{D} \equiv \sqrt{1 - 4\hat{C}}$. In our electroweak fit we work to leading non-trivial order in the dimensionless couplings, $\mathcal{O}(\xi^2)$.

We fix the mass-mixing parameter κ and find the maximum allowed kinetic mixing parameter ε as a function of the dark Z mass by performing a χ^2 fit with respect to the remaining electroweak observables [99]. For example, the W mass is related to the Z mass by $m_W^2 = c_W^2 m_Z^2$. This gives an expression for the W mass in the dark Z model,

$$m_W^2 = \frac{1}{4} g_Z^2 v^2 \left(1 - s_W^2\right) \tag{3.16}$$

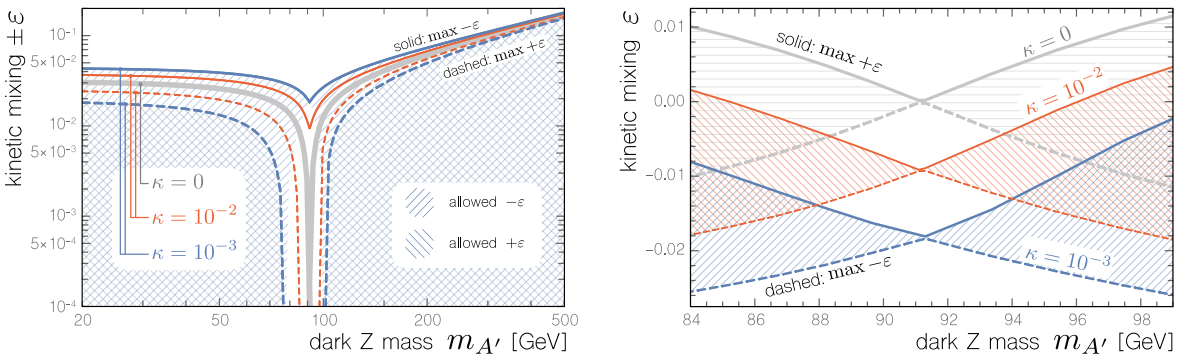


Figure 3.3: 90% confidence level constraints on the kinetic mixing parameter from electroweak precision observables as a function of dark Z mass and for representative values of the mass mixing parameter, κ. The relative sign between κ and ε matters so that bounds are asymmetric with respect to positive (solid line) and negative (dashed line) values of ε. LEFT: bounds on $|\varepsilon|$ for positive/negative values of ε shown on a log scale. RIGHT: bounds on a linear scale, showing the allowed region (hashed) near the Z-mass where mixing is large, (3.6).

where the dependence on the new physics parameters comes from (3.13)–(3.15). The rest of the fit is set by the Z-pole observables in Table 3.1. Of these, only the Z total width depends on g_Z. The remaining observables are ratios that depend on new physics parameters through the effective Z charges, (3.9), as well as the new physics dependence of s^2_W in (3.15) that appears in (3.8).

We present the 90% confidence precision electroweak constraints on the kinetic mixing in Fig. 3.3 for fixed values of the mass mixing, κ. For the case of dark photon, $\kappa = 0$, our results are consistent with Refs. [100, 85]. In the presence of mass mixing, the constraints depend on relative sign of ε and κ; see (3.9) and (3.12). The constraint becomes weaker when ε and κ have the same sign, and stronger otherwise. The dependence on the dark Z mass is the same as in the case of dark photon: the constraint is strongest for $m_{A'} \approx m_Z$, and weakens rapidly outside this region.

3.3.2 Drell–Yan Production at Hadron Colliders

The Large Hadron Collider produces on-shell dark Z bosons through the Drell–Yan process, $q\bar{q} \to A'$. To evaluate the LHC reach, we focus on the dimuon resonance search, looking for decays $A' \to \mu\bar{\mu}$. A similar reach is expected in the di-electron channel, while the quark and tau channels have much lower sensitivity due to large QCD backgrounds. In the region of parameter space where mass mixing dominates, the branching ratio to muons is suppressed and decays to W^+W^- and Zh become dominant; see Sec. 3.2.4. The LHC reach in that region can be improved by searching for a resonance in boson-pair final states. We defer this analysis to future work.

The LHC collaborations report limits on the production cross section on resonances decaying into a dilepton pair[92, 93, 94]. We recast these limits into bounds on our model parameters by computing the Drell–Yan cross section using the differential parton luminosity[3] for partons a and b,

$$\frac{d\mathcal{L}_{ab}}{d\hat{s}} = \frac{1}{s}\frac{1}{1+\delta_{ab}}\int_\tau^1 \frac{dx}{x} f_a\left(x, \sqrt{\hat{s}}\right) f_b\left(\frac{\tau}{x}, \sqrt{\hat{s}}\right) + (a \leftrightarrow b)\,, \tag{3.17}$$

where $\tau = \hat{s}/s$ is the ratio of the partonic to the proton center-of-mass collision energies-squared. The $(1+\delta_{ab})^{-1}$ factor accounts for double counting identical initial partons, though this is not relevant for Drell–Yan production of the dark Z since the quark and antiquark are distinguishable. The proton–proton cross section σ_{pp} is related to the partonic cross section $\hat{\sigma}_{ab}$ as

$$\sigma_{\text{pp}} = \sum_{ab}\int \frac{d\hat{s}}{\hat{s}}\frac{d\mathcal{L}_{ab}}{d\hat{s}}\left(\hat{s}\hat{\sigma}_{ab}\right) = \sum_f^{\text{quarks}} \frac{d\mathcal{L}_{f\bar{f}}}{d\hat{s}}\hat{\sigma}_{f\bar{f}}(m_{A'}^2)\,, \tag{3.18}$$

where the right-hand side sums over quark flavors f with associated antiquark of flavor $\bar{f}$.

[3] We use the parton luminosities $d\mathcal{L}_{a\bar{a}}/d\hat{s}$ provided by the *Mathematica* program `ManeParse` [101] and based on the nCTEQ15 parton distribution functions [102].

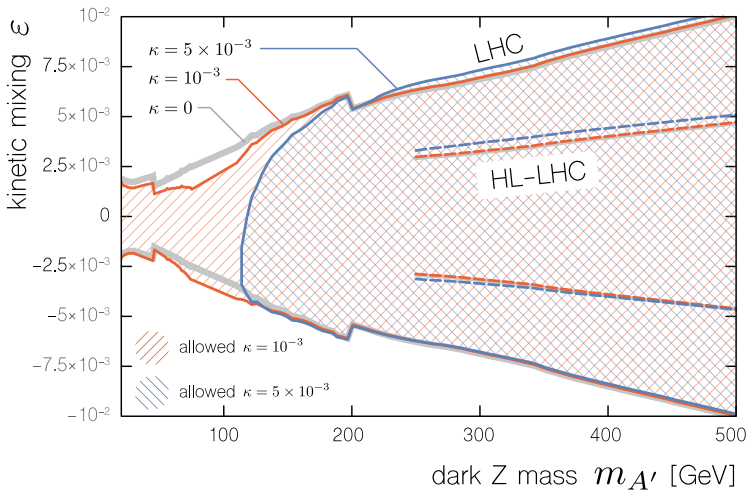


Figure 3.4: Effect of different values of κ on the LHC constraints. The vertical axis is on a linear scale to show how the relative sign of κ and ε affects the size of the coupling; see (3.10).

We ignore kinematic thresholds and use the narrow-width approximation in the partonic cross section $\hat{\sigma}$. The flavor dependence shows up in the production cross section, which is proportional to the square of the effective couplings,

$$\hat{\sigma}_{f\bar{f}}(m_{A'}^2) = \left[(g_f^{\mathrm{V}})^2 + (g_f^{\mathrm{A}})^2\right] \tilde{\sigma}(m_{A'}^2)\,. \tag{3.19}$$

The couplings g^{V} and g^{A} are given in (3.10), and the rescaled cross section $\tilde{\sigma}$ is flavor-universal.

We recast the LHC dilepton results by setting $\sigma_{\mathrm{pp}}(m_{A'}, \varepsilon, \kappa) < \sigma_{\mathrm{LHC}}$, where σ_{LHC} is the upper limit set by LHC analyses. This straightforwardly produces the bounds on the effective theory parameters ε and κ shown in Fig. 3.4. In the case of pure kinetic mixing, our results are in agreement with the bounds reported by the CMS collaboration [92]. We also extrapolate the current LHC bounds to estimate the reach of the full HL-LHC data set with 3000 fb^{-1}, assuming that statistical errors dominate so that the sensitivity to new physics cross section scales as the square root of the integrated luminosity.

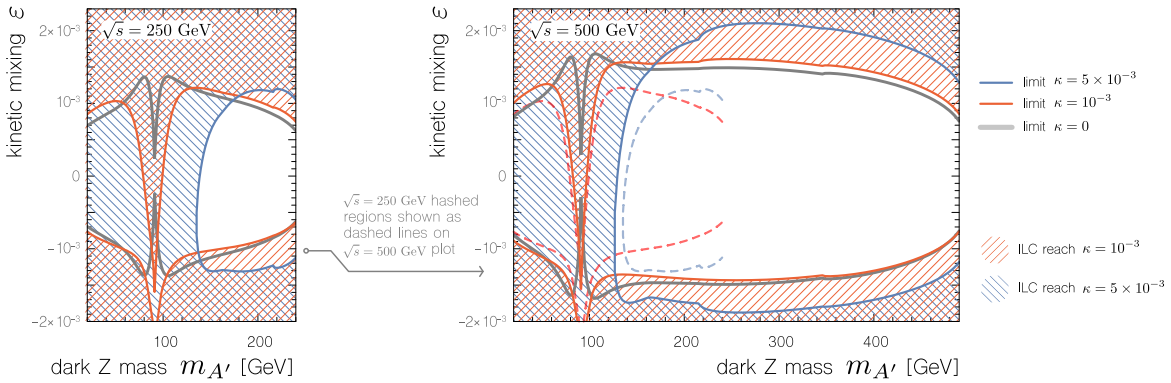


Figure 3.5: Dimuon resonance discovery reach at the ILC ($\mathcal{L}_{\rm int} = 2$ ab^{-1} at 250 GeV and 4 ab^{-1} at 500 GeV). Stronger reach corresponds to shaded regions closer to $\varepsilon = 0$. We plot both positive and negative values of the kinetic mixing, ε, for different values of the mass mixing, κ. The relative sign of κ and ε is significant, as seen in the couplings (3.10) and in the asymmetry of these plots for $\varepsilon \to -\varepsilon$. For $\kappa = 10^{-3}$ (red lines) at masses near 120 GeV, the $\varepsilon < 0$ branch has a stronger reach than $\kappa = 0$, whereas for $\varepsilon > 0$ the reach is weaker than $\kappa = 0$. For larger κ (e.g. blue lines), the mass mixing is so large that it dominates over the kinetic mixing.

3.3.3 Resonance Search at the ILC

We examine the reach for the ILC to discover the dark Z in the radiative return channel $e^+e^- \to \gamma A' \to \gamma\mu^+\mu^-$. The signal rate is strongly enhanced when the photon is collinear with the beam and hence unobservable.[4] We thus calculate the signal as on-shell $e^+e^- \to A'$ production at center-of-mass squared energy xs in the equivalent-particle approximation where the beam of energy $E = \sqrt{s}/2$ contains a distribution of lower-energy $e^\pm$ particles with energy $xE < E$ that account for the unobserved collinear photon of energy $E_\gamma = (1-x)E$ [103, 104, 105]. (See Ref. [106] for a pedagogical introduction.) At the leading order in QED perturbation theory, the probability of finding an $e^\pm$ with energy xE in the

[4] He et al. estimate the discovery reach for a model with pure kinetic mixing at a future lepton collider with cuts that require the photon to be observable [89]. This is overly conservative because a high-mass dimuon resonance is a clean signal whether or not the photon is tagged.

beam is

$$f_e(x)\,dx = \frac{\alpha}{2\pi}\frac{1+x^2}{1-x}\ln\frac{\sqrt{s}}{m_e}\,dx\,. \tag{3.20}$$

The total signal cross section for $e^+e^- \to \gamma(A' \to \mu^+\mu^-)$ is readily expressed in terms of the cross section for two-to-one production of the dark Z:

$$\sigma = \frac{\alpha}{\pi}\ln\frac{\sqrt{s}}{m_e}\int_0^1 dx\,\frac{1+x^2}{1-x}\hat{\sigma}(e^+e^- \to A')\,\mathrm{Br}(A' \to \mu^+\mu^-)\,. \tag{3.21}$$

In the narrow-width approximation, $\hat{\sigma} \propto \delta(s' - m_{A'}^2)$, where $s' = xs$ is the center-of-mass energy-squared of the "partonic" collision. The cross section for background events mediated by photons or SM Z bosons follows a similar formula, except that narrow-width approximation does not hold since we will be searching in regions where the photon or Z are off-shell (unless $m_{A'} \sim m_Z$). The dark Z appears as a narrow peak in the dimuon invariant mass distribution. We estimate the signal significance by computing $S = N_{\text{sig}}/\sqrt{N_{\text{bg}}}$, where N_{bg} is the number of background events in the mass window centered at $m_{A'}$. The intrinsic width of the dark Z is very small in the region of interest due to its small couplings. Thus the size of the mass window is controlled by muon momentum resolution. We use $\Delta(1/p_T) \sim 10^{-5}$ GeV^{-1} [107] in our estimates.

The ILC reach at 2σ confidence for a benchmark value of $\kappa = 5 \times 10^{-3}$ is shown as the colored lines in Fig. 3.2. This analysis demonstrates that the ILC can serve as the discovery machine for this type of new physics. The ILC reach significantly exceeds both the current limits and the extrapolated sensitivity of the high-luminosity LHC in the kinematically accessible regions. Fig. 3.5 presents the ILC reach for different values of the mass mixing κ. We note that for $\kappa = 0$, our results are in agreement with the previous estimates of the dark photon reach in Refs. [86, 108].

Our analysis assumes unpolarized ILC beams. Depending on the model parameters, the couplings of the dark Z can depend strongly on the lepton chirality, while the background does not have strong polarization dependence. In such models, the appropriate beam polarization may further enhance the ILC reach. However, since the chiral structure of the couplings is not known *a priori*, running with unpolarized beams is overall the preferred search strategy. If a dark Z is discovered, polarized beams can be used to measure its chiral couplings; we explore this in the next section.

We also note that, unlike the LHC, hadronic decays of the dark Z do not suffer from large backgrounds at a lepton collider, and can be included to further improve the ILC reach. We leave this analysis for future work.

3.4 Precision Measurements at the Dark Z Pole

Electron–positron colliders such as the ILC are an ideal laboratory to determine the chiral couplings of a dark Z resonance. This program is analogous to the precision measurements of the electroweak gauge bosons at LEP and SLC. To illustrate this capability, we consider the benchmark point

$$m_{A'} = 400\,\text{GeV} \qquad \varepsilon = 5 \times 10^{-3} \qquad \kappa = 5 \times 10^{-3}\,. \tag{3.22}$$

A dark Z at this benchmark would be already discovered at the HL-LHC, see Fig. 3.2. Models with slightly smaller couplings may evade the HL-LHC and be first discovered at a 500 GeV ILC. In either case, an electron–positron collider would provide the first opportunity to measure the dark Z couplings to fermions with high precision, and to determine their chiral structure which is crucial for model discrimination.

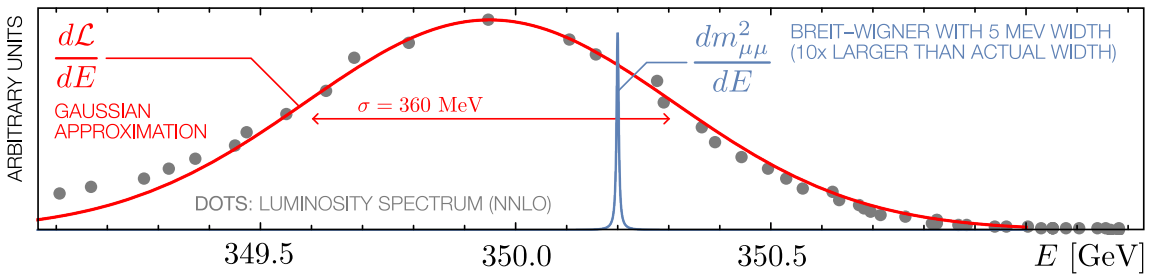


Figure 3.6: ILC beam luminosity spectrum, for nominal center-of-mass energy $\sqrt{s} = 350.5$ GeV. Dots are the simulation results from Fig. 2 of Ref. [109], while the red line is a Gaussian fit. The Gaussian width is $\sigma = 360$ MeV. We compare to a hypothetical Breit–Wigner distribution for a dark Z with mass $m_{A'} = 350.2$ GeV. We choose a dark Z width $\Gamma_{A'} = 5$ MeV; this is $10\times$ larger than the expectation for the benchmark theory so that the width is visible in the plot. Unlike typical models of $\mathcal{O}(\text{TeV})$ spin-1 gauge bosons, the luminosity spectrum spread is much broader than the dark Z width. Vertical axis normalization is arbitrary.

The mass of the dark Z will be measured experimentally at both the HL-LHC and through the radiative return production in the 500 GeV ILC run discussed above. The couplings of the dark Z may be measured efficiently in a relatively short period of time with a dedicated ILC run at a center-of-mass energy $\sqrt{s} = m_{A'}$. In this section, we estimate the potential of this run to measure the dark Z couplings to charged leptons. We study the channel $e^+e^- \to A' \to \mu^+\mu^-$ assuming flavor-universal couplings; this assumption is automatically satisfied in our underlying model. The left- and right-handed couplings are defined by

$$g_{\text{R,L}} = \frac{1}{2}\left(g^{\text{V}} \pm g^{\text{A}}\right) \qquad g_{\text{R}} = -2.1 \times 10^{-3} \qquad g_{\text{L}} = -9.3 \times 10^{-4}\,, \tag{3.23}$$

and are related to the underlying theory parameters in (3.10); we give numerical values for the benchmark point (3.22). To measure these couplings, we consider three observables:

- The number of events, N_{prod}
- The forward–backward asymmetry, A_{FB}

- The left–right asymmetry for polarized beams, A_{LR}.

The expressions for the two asymmetries are well-known and are summarized in Appendices A.3 and A.4. The number of events is given by

$$N_{\mathrm{prod}} = \int d\sqrt{s}\, \frac{d\mathcal{L}}{d\sqrt{s}} \sigma_{\mathrm{prod}}, \tag{3.24}$$

$$\sigma_{\mathrm{prod}} = \frac{\pi}{2}\big[(1-P_{e^-})(1+P_{e^+})g_L^2 + (1+P_{e^-})(1-P_{e^+})g_R^2\big]\delta\big(s-m_{A'}^2\big). \tag{3.25}$$

Here P_{e^-} and P_{e^+} denote polarizations of the electron and positron beams, respectively and $d\mathcal{L}/d\sqrt{s}$ is the differential luminosity of the colliding beams at the ILC. The somewhat unusual formula for N_{prod} is due to the fact that intrinsic width of the A' is extremely small, $\Gamma_{A'} \sim 10^{-4}$ GeV for our benchmark parameters (3.22). The beam luminosity spectrum of the ILC in the vicinity of the nominal center-of-mass energy can be modeled as a Gaussian with a width of order $\mathcal{O}\,(1\text{ GeV})$ [109], much larger than $\Gamma_{A'}$, as illustrated in Fig. 3.6. Treating the Breit–Wigner resonance as a δ-function yields (3.25).

The prediction for N_{prod} depends sensitively on the relative location of the Breit–Wigner resonance with respect to the peak of the luminosity spectrum. To obtain a precise constraint on the couplings from a measurement of N_{prod}, this relative location must be known to a precision $\ll 1$ GeV. This is a non-trivial requirement. The position of the peak of the luminosity spectrum will be known at the level of $10^{-4}E_{\mathrm{beam}} \sim 40$ MeV [110], which is sufficient. However, prior to the dedicated on-resonance run of the ILC, the mass of the A' will only be measured as the location of the peak in the dimuon invariant mass distribution, observed either at the HL-LHC or in the radiative-return sample collected at the 500 GeV ILC. In either case, the precision of the mass measurement is limited by the muon momentum resolution, which is of order GeV. A much more precise knowledge of the location of the Breit–Wigner peak is required.

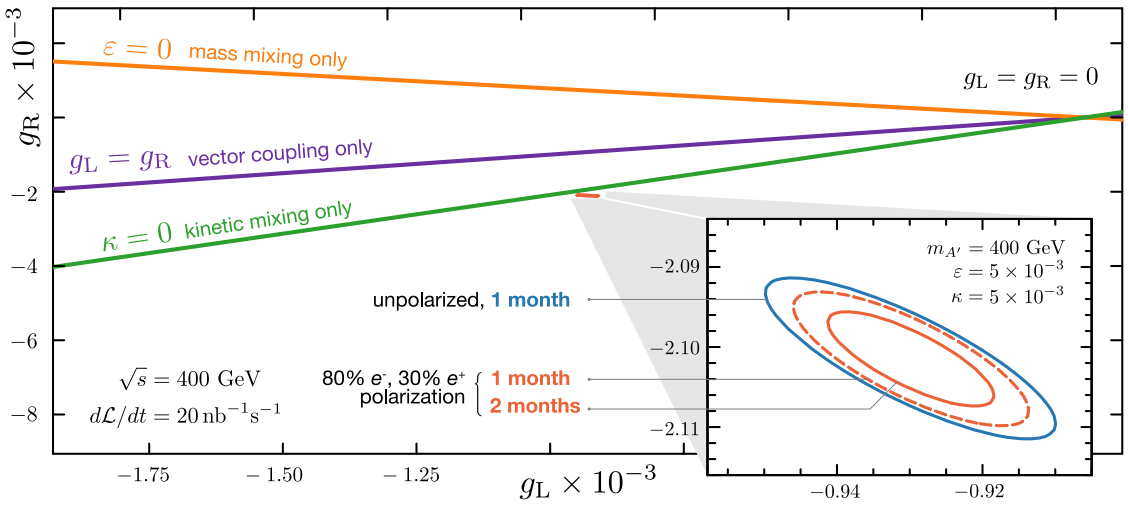


Figure 3.7: ILC measurement of the dark Z chiral couplings to leptons. The ellipses correspond to 95% confidence level, assuming that the best-fit values of the couplings match the underlying model predictions and that statistical errors dominate. The lines represent alternative models of a narrow spin-1 resonance that can be tested by this measurement.

To address this problem, we propose that the on-resonance run of the ILC should begin with an *inverse lineshape scan* in which the "peak" center-of-mass energy—the location of the peak of luminosity spectrum—is varied in small steps within the A' mass window inferred from the HL-LHC/ILC-500 data. For example, our benchmark point (3.22) motivates $\sqrt{s} = 399$–401 GeV. At each beam energy step, measure the A' production cross section. Since the luminosity spectrum is independently known, the m'_A can be inferred from the measured cross sections. We estimate that with sufficient statistics at each energy step, this method can provide an error on m'_A comparable to the peak beam energy uncertainty, ~ 40 MeV. After performing the inverse lineshape scan, the peak center-of-mass energy can be chosen to match the measured m'_A and maximize the dark-Z production rate.

Fig. 3.7 shows the expected statistical errors of the measurement of g_L and g_R

	$g_{\rm L} \times 10^5$	$g_{\rm R} \times 10^5$
Benchmark values	-93.0	-210
1 month (unpolarized)	$\pm 4.0 \pm 0.3$	$\pm 2.0 \pm 0.8$
1 month (polarized)	$\pm 3.3 \pm 0.3$	$\pm 1.7 \pm 0.7$
2 months (polarized)	$\pm 2.3 \pm 0.3$	$\pm 0.85 \pm 0.7$

Table 3.2: Chiral couplings $g_{\rm L,R}$ for the benchmark values in (3.23) and the statistical (first number) and systematic (second number) uncertainties for different ILC run times and beam polarizations. Polarized beams assume 80% and -30% polarizations in e^- and e^+ beams, respectively.

with one-month and two-month runs on the dark Z resonance. The ellipses correspond to 95% confidence level measurement, assuming that the best-fit values of the couplings match the underlying model predictions. Also shown are the predictions of three alternative theoretical models of a vector resonance: pure kinetic mixing[5] ($\kappa = 0$), mass mixing only ($\epsilon = 0$), a parity-conserving ($g_{\rm L} = g_{\rm R}$). It is clear from the figure that the measurement of the chiral couplings at the ILC can rule out all three possibilities, establishing parity violation and proving that the underlying dark Z model has both kinetic and mass mixing. In contrast, the LHC can only constrain the combination $\sqrt{g_{\rm L}^2 + g_{\rm R}^2}$, which is not sufficient to discriminate between alternative models.

The constraint on the couplings inferred from the total number of events $N_{\rm prod}$ is subject to systematic errors from the uncertainties in luminosity and peak beam energy and from the uncertainty in the measurement of the mass $m_{A'}$. The mass is measured through the inverse lineshape scan. The expected error is of the same order as that in the peak beam energy, $\delta m_{A'} \approx 40$ MeV. The expected overall luminosity uncertainty is $(\Delta\mathcal{L}/\mathcal{L}) \sim 4 \times 10^{-3}$ [111]. We estimate that these uncertainties affect the coupling determination at a level that is,

[5]The benchmark point is close to the $\kappa = 0$ line even though $\kappa \sim \varepsilon$ because the contribution of κ to the chiral couplings is suppressed by $m_Z^2/m_{A'}^2 \sim 0.05$.

at most, comparable to the statistical errors shown in Fig. 3.7. We do not include uncertainties in the shape of the beam luminosity spectrum. The measurement of the left-right asymmetry $A_{\rm LR}$ is subject to beam polarization uncertainties. ILC polarimeters can achieve a precision of $\sim 0.25\%$ [112]. The effect of this uncertainty on the coupling measurement is negligible. We collect our estimates of the statistical and combined systematic errors of the chiral coupling measurements in Table 3.2. We conclude that a dedicated run of the ILC at the dark Z pole will yield percent-level measurement of $g_{\rm L}$ and $g_{\rm R}$, providing crucial information for discrimination among alternative models of the resonance.

3.5 Conclusions

A next-generation electron–positron collider, such as the ILC, offers unique opportunities to search for and characterize new physics beyond the Standard Model. In this paper, we study a model of a new Abelian gauge boson, the dark Z. While SM fermions are not directly charged under the new gauge group, the dark Z can interact with the SM through both kinetic and mass mixing. If the dark Z mass is above 10 GeV or so, energy-frontier colliders offer the best way to search for this particle. We studied the current constraints on this particle from precision electroweak measurements and direct searches at the LHC. This extends the previous work on the dark photon model that couples to the SM only via kinetic mixing to the more general case when mass mixing is also present. We evaluate the reach of the ILC, running at 250 and 500 GeV, in the dark Z parameter space. The ILC reach covers a large part of the parameter space unconstrained by current experiments and not accessible at the HL-LHC; see Fig. 3.2. Thus, the ILC can serve a discovery machine in this new physics sce-

nario. If a dark Z is discovered, the ILC offers a unique opportunity to precisely measure its couplings to the SM fermions, including their chiral structure. We demonstrated that a short dedicated run at the center-of-mass energy matching the dark Z mass will achieve percent-level measurement of the chiral couplings of the dark Z to leptons. Such a measurement can play a crucial role in model discrimination; this is illustrated in Fig. 3.7.

The idea of using on-resonance electron-positron collisions to measure the properties of a massive vector boson is, of course, not new: LEP-1 and SLC/SLD used it to understand the Standard Model Z with unprecedented precision. A new twist in the case of dark Z is its extremely small intrinsic width, which is much smaller than the width of the beam luminosity spectrum as well as the energy/momentum resolution of the detector for all visible decay products. In this situation, precise measurement of the mass and couplings of the resonance hinges on understanding the beam luminosity spectrum. Note that the small intrinsic width of the dark Z is directly related to the smallness of its couplings to SM fermions, which is in turn necessary for a sub-TeV resonance to avoid precision electroweak and LHC constraints. Independently of the underlying model, any resonance compatible with current data and kinematically accessible at the ILC will necessarily be very narrow. This motivates further studies of on-resonance production of such narrow states at the ILC and other proposed lepton colliders.

CHAPTER 4

ANOMALY DETECTION IN THE PRESENCE OF IRRELEVANT FEATURES

4.1 Introduction

Experiments at high-energy colliders, such as the Large Hadron Collider (LHC), continue to be the primary source of information about the nature of physics at the microscopic scales. A major task of the current and future experiments is to search for deviations from the Standard Model (SM) of particle physics. Traditionally, such searches are performed by assuming a particular model for physics beyond the Standard Model (BSM), and optimizing the event selection and statistical analysis to obtain maximum sensitivity to the new physics signal in the presence of the SM background. Increasingly, these are supplemented with data-driven methods which minimize model-dependent assumptions about the structure of deviations from the SM, with machine-learning (ML) based approaches the primary driver of such searches [113, 114].

The enormous size and complexity of the data sets collected by collider experiments currently preclude conducting a search for "anything that doesn't look like the SM" in the full data set at once. Even if it were possible in principle, the dependence of collider analyses on complex simulations to interpret measured signals would make such an approach extremely sensitive to mismodelling errors at all stages of the simulation chain. Instead, recent work focuses on a simpler task of anomaly detection when localized with respect to a particular variable [115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125]. A well-studied benchmark example, starting with the work of [115, 118], is a search

for a dijet resonance, in which the signal jets are produced by a boosted resonance decay which is imprinted in non-trivial jet substructure. ML techniques allow for searches of anomalous events with such topology, without making strong model-dependent assumptions about the new physics model that gives rise to this signal. The original algorithm used the Classification Without Labels (CWoLa) approach [126]. In this approach, events are divided into signal and side-band regions based on the invariant mass m_{JJ}. A neural-network (NN) classifier is trained to discriminate between events from the signal and side-band regions. This classifier is then applied to search for anomalous events in the signal region. This approach requires that the features distinguishing signal and background be uncorrelated with m_{JJ}, which is not always the case in real-world applications. To circumvent this problem, algorithms such as ANODE [119] and CATHODE [122] were developed to detect anomalies based on probability density estimation.

A serious issue that can hinder practical applications of the ML-driven anomaly detection methods is the rapid deterioration of performance with growing dimensionality of data space. Typically, collider data contains some observables (or *features*) that are relevant for discriminating signal and background, and a number of observables whose distribution is very similar in the signal and background samples. In a true model-agnostic search, one rarely has the privilege of knowing what features are important beforehand, and inevitably many of the included features can be *irrelevant*. It has been observed that the existing algorithms for anomaly detection, in the context described above, lose their discriminating power very rapidly as even a small number of irrelevant features are added to the input vectors [127]. In this paper, we will present approaches that address this issue within both the classifier-based and

probability-density-based approaches to anomaly detection.

The algorithms on which we focus here are based on Boosted Decision Trees (BDTs), rather than neural networks. BDTs tend to outperform neural networks on tabular data, where they can take advantage of the preferred basis implied by the input features [128, 129]. Additionally, given that we are working with meaningful inputs (i.e. high level features), BDTs generally require much less data preprocessing and computational cost compared to neural networks.

The rest of the paper is organized as follows. In Sec. 4.2, we describe the "signal" and "background" data sets that are used in our analysis, and specify how we model the extraneous irrelevant features. In Sec. 4.3, we present a BDT-based classifier which uses the CWoLa approach to aid anomaly detection. We show that before irrelevant features are added, the BDT algorithm achieves performance similar to that of NN-based classifiers. However unlike the NN, the BDT performance does not deteriorate significantly when irrelevant features are present. In Sec. 4.4, we show how the BDT can be used as a probability density estimator, providing a powerful tool for anomaly detection even when relevant features are correlated with m_{JJ}. Furthermore, this algorithm is also robust in the presence of irrelevant features. Sec. 4.5 contains our conclusions. Technical details related to tuning of hyperparameters of the BDT algorithms are presented in Apps. B.1 and B.2, while a case study of our methods' performance on a dataset with mutually dependent irrelevant features is discussed in App. B.3.

In all plots in this paper, the curves showing performance of neural network anomaly-detection tools are generated using code provided at `https://github.com/HEPML-AnomalyDetection/CATHODE`.

4.2 Dataset

The signal and background events used in this study are from the LHC Olympics 2020 R&D dataset [130]. In particular, the SM background corresponds to QCD dijet events while the anomalous signal we want to detect is produced by the decay $W' \to X(\to qq)Y(\to qq)$. Here W', X and Y are hypothetical new bosons with masses $3.5\,\mathrm{TeV}$, $500\,\mathrm{GeV}$ and $100\,\mathrm{GeV}$ respectively. All events are produced using the `Pythia8` [131] and `Delphes 3.4.1` [132] Monte Carlo generators, and jets in each event are identified using `FastJet` [133] using anti-k_T clustering with $R = 1$.

The training (plus validation) set is constructed by combining 1000 randomly selected signal events with a sample of $1\,000\,000$ background events. For evaluation purposes, a separate test set is constructed by having $20\,000$ signal events and $40\,000$ background events, all of which lie inside the signal region (defined below). This test set is not used during training.

The physically motivated relevant features are based on the two highest p_T jets. They include

- m_{JJ}: invariant mass of the two jets, which will be the resonant feature.
- m_{J_1}: invariant mass of the lighter jet.
- Δm_J: absolute mass difference between the two jets' invariant masses.
- $\tau_{21}^{J_1}$, $\tau_{21}^{J_2}$: n-subjettiness ratios [134, 135] of the two jets, defined by $\tau_{21} \equiv \tau_2/\tau_1$.

Following [118, 119, 122], we define the signal region (SR) by $m_{JJ} \in [3.3, 3.7]$ TeV, and the sideband region (SB) by $m_{JJ} \notin [3.3, 3.7]$ TeV. Additionally, for

the CWoLa method, we also define a *short side-band* (SSB) region, which extends to both sides of the SR by $200\,\mathrm{GeV}$: $m_{JJ} \in ([3.1, 3.3] \cup [3.7, 3.9])\ \mathrm{TeV}$. These definitions will be used throughout the rest of the paper.

However, for a model-agnostic search, one would not know *a priori* that the observables above are the only features of interest and would likely not have any principled way of excluding additional superfluous features. To simulate such a scenario, we artificially augment the original dataset with features drawn from Gaussian distributions, which will be considered as our irrelevant features. We vary the number of such irrelevant features and examine how much effect they have on anomaly detection performance. Specifically, we study the cases of 4 and 16 irrelevant features. The first case represents the situation where the dataset is a roughly equal mix of relevant and irrelevant features, while the second case provides an example of a dataset dominated by irrelevant features.

4.2.1 What Do We Mean by Irrelevant?

Even though the notion of an irrelevant feature is intuitively clear, it is necessary for us to define it more precisely. We provide here two possible characterizations of ignorable irrelevant features, each suited to the respective anomaly detection method considered in the text.[1] Here "ignorable" means that the feature should not matter in the limit of an infinite amount of data.

- CWoLa method:

  A feature y is irrelevant if $p(m_{JJ} \in \text{SR}|y) = p(m_{JJ} \in \text{SR})$.

[1] The general definition of irrelevancy has been explored in [136]. The conditions stated below are less general, but suffice in the context of anomaly detection.

- Probability density estimation-based method:

  A feature y is irrelevant if it is statistically independent of m_{JJ} and the auxiliary (relevant) features: $p(m_{JJ}, x_1, \ldots, x_K, y) = p(m_{JJ}, x_1, \ldots, x_K)\, p(y)$. This must hold in both the background and signal samples. Moreover, $p(y)$ in the signal and background samples must be identical.

In this paper, we will explore the performance of anomaly detection algorithms as a function of the number of irrelevant features N. As a baseline model, throughout this paper we assume that the irrelevant features y_i are distributed according to a direct product of Gaussians:

$$p(y_i) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-y_i^2/2}. \tag{4.1}$$

A vector of N features drawn from this distribution is then tacked onto each event in the LHCO 2020 dataset described above, with no distinction made between signal and background events. The features y_i in the resulting dataset satisfy both of the irrelevancy definitions above.

Within our baseline model, the y_i's are mutually statistically independent among themselves. This feature is not generic, and is not expected to always hold in realistic physics scenarios. In App. B.3 we show that our anomaly detection algorithms continue to perform well when the irrelevant features are mutually dependent.

4.2.2 Performance Metric

As is standard, we shall present performance comparisons between NNs and our proposed methods in terms of the *significance improvement characteristic* (SIC)

curve, which is obtained by plotting the significance improvement,

$$\text{SIC} = \frac{\epsilon_S}{\sqrt{\epsilon_B}}\,, \tag{4.2}$$

against ϵ_S. Here ϵ_S is the fraction of correctly identified signal events (true positive rate), and ϵ_B is the fraction of background events incorrectly identified as signals (false positive rate). It should be emphasized here that SIC is a meaningful metric only when the analysis is statistics-limited and not systematics-limited, and the sample is background-dominated; we shall assume that this is the case.

4.3 CWoLa on a Tree: Classifier BDTs

In this section we compare the performance of BDT-based and NN-based CWoLa methods in the presence of irrelevant features.[2]

The CWoLa hunting [115, 118] method attempts to construct the Neyman–Pearson optimal discriminator [137] between a signal and a background where the signal is assumed to be dominantly present in the SR. The key observation underlying this is that if the SSB and SR have different admixtures of signal and background, then the optimal signal-background discriminator is monotonically related to the optimal SSB-SR classifier and finding one produces the other, provided that the auxiliary features $\vec{x}$ are independent of the resonant mass m_{JJ} for the background. While this is a theoretical guarantee, finding an optimal SSB-SR classifier can be difficult in practice. This is because at very low

[2] In [127], similar comparisons are made in the context of *idealized anomaly detection*, in which perfect understanding of background is assumed. This included a more detailed, physical model of irrelevant features, while we consider a more realistic measurement scenario. We hence view the two studies as naturally complimentary.

S/B ratio, SSB and SR events largely overlap in feature space with very similar distributions, and most modern machine learning models are flexible enough to mistake local fluctuations for actual excess of signal events (*i.e.*, over-fitting). This situation is particularly exacerbated in the presence of irrelevant features, because they provide additional sources of statistical fluctuations in a higher dimensional space.

The above consideration do not actually select for a method of approximating the SSB-SR classifier. In studies involving the CWoLa hunting method, the classifier typically consists of a fully-connected feedforward neural network. However, it is well-known that neural networks do not fare well with irrelevant inputs, and this is especially so when they are applied in the CWoLa setting for reasons above.

On the other hand, tree-based models are known to be innately robust against irrelevant features [138, 128], an observation usually attributed to the way they are constructed — for most tree-based models they are built by performing cuts in feature space to greedily minimize metrics such as information gain, meaning that they already have some degree of internal feature selection built in. Here we capitalize on this empirical observation and apply BDT-based CWoLa to a more realistic setting where inevitably there will be a lot of irrelevant features.

In what follows, we use `xgboost` as a reference BDT model to compare with a fully-connected feed-forward NN. `xgboost` is chosen since it is widely considered as (one of) the state-of-the-art gradient boosting tree algorithms in terms of speed and accuracy. For detailed descriptions of the `xgboost` algorithm, refer to [139].

n_estimators	max_depth	eta	alpha	lambda	subsample
292	9	6.2×10^{-3}	50	74	0.75

Table 4.1: One set of `xgboost` hyperparameters found for the dataset without any irrelevant features. We use default values for other hyperparameters. Refer to App. B.1 or [139] for a more detailed discussion of these hyperparameters.

4.3.1 Training Procedures

For training the CWoLa classifiers[3], we select from the raw training set events for which $m_{JJ} \in [3.1, 3.9]$ TeV. This results in roughly $250\,000$ training events with about 760 signal events in the SR, which corresponds to $S/B \approx 0.6\%$ and $S/\sqrt{B} \approx 2.2$. Classifiers are then trained to differentiate between SR and SSB labels.

The NN-based classifier is constructed by a fully-connected feed-forward neural network with 3 hidden layers, each of which has 64 neurons. Rectified Linear Unit (ReLU) activation function is used. The network is trained for 100 epochs with binary cross-entropy loss using the `Adam` optimizer [140] and learning rate set to 10^{-3}. During training, only half of the dataset constructed above is used for actual training while the other half is used for validation purposes. In particular, the 10 epochs with the lowest validation error are used to construct an ensemble of 10 classifiers.

For the `xgboost`-based classifier, we employ a 10-fold cross-validation so that the entire training set is utilized during actual training. Specifically, we use the cross-validation process to tune `xgboost`'s hyperparameters to the dataset without irrelevant features[4]. Details of this procedure can be found in App. B.1.

[3]The data handling and training procedures are the same as in [122]. Here we summarize them for the sake of completeness.

[4]This is necessary because the default hyperparameters are far too aggressive and lead to severe overfitting.

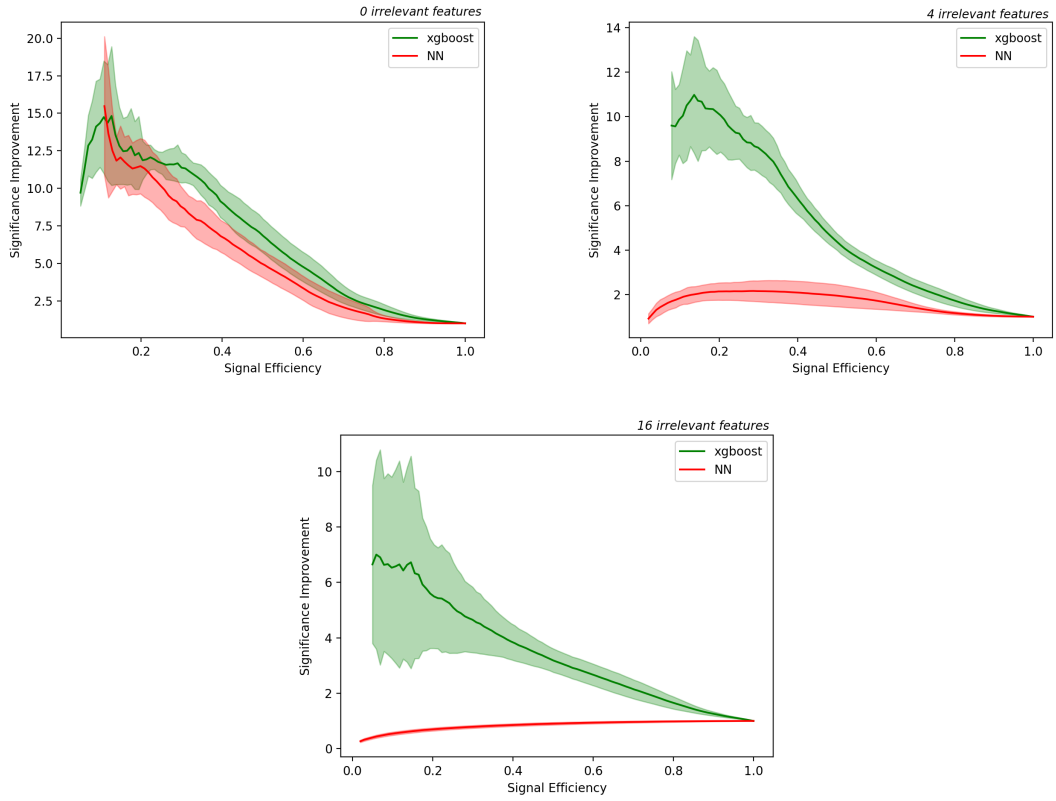


Figure 4.1: Performance comparisons between BDT-based (green) and NN-based (red) CWoLa methods for dataset augmented with 0, 4, and 16 irrelevant Gaussian features. The same `xgboost` hyperparameters are used to train in all 3 cases. The solid lines represent average SIC value across a classifier ensemble defined in the text, and the bands refer to 1 standard deviation of SIC. It is important to note that for neural networks, the bands correspond to variability for a *fixed* set of hyperparameters, while for `xgboost` they correspond to variability across *different* hyperparameters found by Bayesian optimization. Clearly, `xgboost` is far more robust against the inclusion of irrelevant features than neural networks.

Since the hyperparameter optimization procedure is stochastic, we find 10 independent sets of hyperparameters, each of which is used to train a separate classifier and they together form an ensemble of 10 classifiers. The same set of hyperparameters is also used to train dataset augmented with irrelevant features. In Tab. 4.1, we show one set of hyperparameters found.

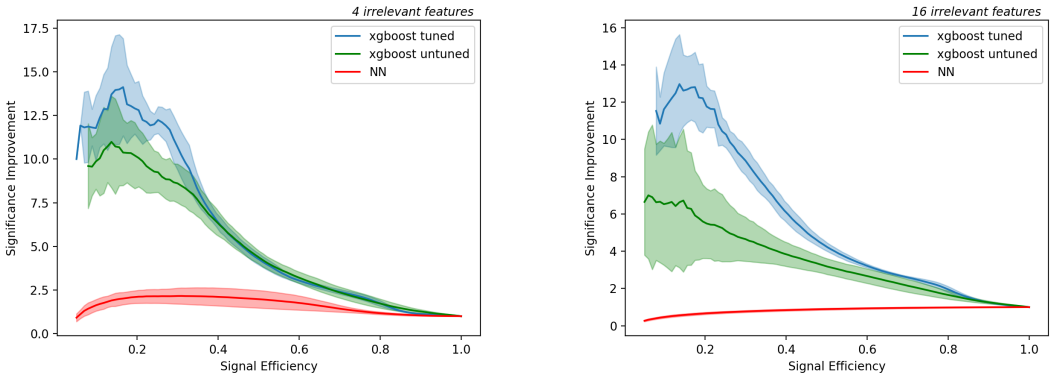


Figure 4.2: SIC curves for the `xgboost` classifier with hyperparameters optimized for the dataset with no irrelevant features (green), and the same classifier with hyperparameters re-optimized each time more irrelevant features are added (blue). With proper tuning, much of the original performance can be recovered even when the dataset has a large fraction of irrelevant features. NN-based classifier's performance on the same datasets (red) is provided to help guide the eye.

4.3.2 Performance Comparison

The performances of `xgboost`-based and NN-based CWoLa are shown in Fig. 4.1. Both `xgboost`-based and conventional CWoLa perform similarly in the absence of irrelevant features. However, when irrelevant features are present, the performance degradation of the neural network is much more severe than that of the BDT. In particular, in the regime of large number of irrelevant features (relative to number of relevant ones), the neural network-based CWoLa method becomes essentially ineffective. On the other hand, while BDTs-based CWoLa also suffers from the presence of irrelevant features, it is far more resilient. In particular, even with 16 noisy features, the classifier can still attain an average maximum significance improvement of around 7.

In the plots in Fig. 4.1, the BDT hyperparameters are optimized once, using the dataset with no irrelevant features, and then kept fixed as the classifier is

applied to the datasets with 4 and 16 irrelevant features. This performance can be further improved by dedicated hyperparameter optimization each time more irrelevant features are added. The performances of `xgboost` on the augmented dataset when the hyperparameters are properly tuned are shown in Fig. 4.2. Impressively, much of the model's original performance in the absence of irrelevant features can be recovered without too much of computational burden (relative to neural networks). This shows the overall superiority of using BDTs when the input data is of tabular form in the context of CWoLa hunting.

Another added bonus of using a tree-based classifier is that there exists a naturally defined and easily computable notion of *feature importance* [138]. Recall how a tree-based model is constructed: cuts along different feature directions are selected so as to greedily minimize the loss function. Hence, for each feature one can compute how much it contributes to the overall decrease in loss. This adds a layer of interpretability to the model which can potentially be used to shed light on what features are more relevant in discerning signal from background.[5]

We can use this notion of feature importance to understand why `xgboost` is so much more robust compared to neural networks. In Fig. 4.3, we show box plots of feature importance values as given by the 10 different classifiers in the ensemble in the case of having 16 irrelevant features. Strikingly, the model clearly utilizes the relevant features much more than the irrelevant ones, corroborating with our intuition that tree-based models by nature perform a certain degree of internal feature selection. This is likely the reason why the `xgboost`-based CWoLa shows such favorable results.

[5]It is important to emphasize that this is meaningful only when the features are mostly uncorrelated from each other. If not, it becomes difficult to isolate the effect of each individual feature.

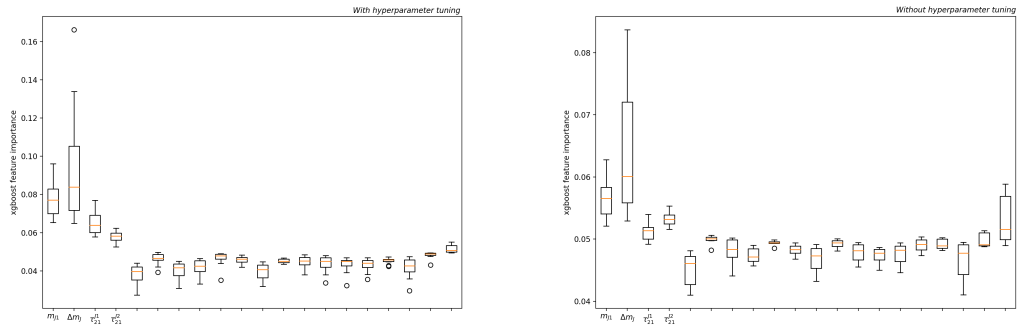


Figure 4.3: Box plots of feature importance values by 10 independent `xgboost` BDT classifiers applied to dataset with 16 irrelevant features. The BDT hyperparameters are optimized on the same dataset (left panel) or on the dataset with no irrelevant features (right panel). The four labeled boxed on the left side of each plot correspond to relevant features, while the 16 unlabelled boxes correspond to the artificially introduced Gaussian noise features. The relevant features in the original dataset are found by the method to be more important for signal/background discrimination compared to the Gaussian noise, and such a difference is even more pronounced with proper hyperparameter tuning.

In conclusion, even a naive direct application of BDT algorithms to CWoLa method can significantly increase its robustness to irrelevant features compared to NN-based CWoLa.

4.4 Probability Density Estimation with BDTs

Even though the CWoLa method can achieve significant sensitivity improvement in anomaly detection, its success hinges on the independence of the auxiliary features $\vec{x}$ with m_{JJ} under the background hypothesis, which is quite a strong assumption and does not hold in general. When this assumption is sufficiently violated, CWoLa performance drops drastically [119, 122].

Since this is a strong assumption that does not always hold in physical analyses scenarios of interest, various methods have been proposed to circumvent

it. In particular, we examine the anomaly detection with density estimation (ANODE) method [119], which was originally implemented using normalizing flows. While this is not the state-of-the-art anomaly detection method, it is chosen since the lessons learned here can be easily transferred to other similar density-estimation-based methods.

Unlike CWoLa, the ANODE method tries to estimate the two probability densities: $p(\vec{x}|m)$ of the full data set, and $p(\vec{x}|m, \text{bkgd})$ of the background only (estimated from the sideband regions and extrapolated in the signal region). Then, the likelihood ratio

$$R = \frac{p(\vec{x}|m)}{p(\vec{x}|m, \text{bkgd})} \tag{4.3}$$

is computed in the signal region. This ratio can be shown to be optimal (in the Neyman–Pearson sense) without any need of additional assumptions.

In other words, the ANODE method mainly consists of two steps:

- Estimate the full density $p(\vec{x}|m)$ directly from data,
- Estimate the background density $p(\vec{x}|m, \text{bkgd})$ by interpolating from the SB regions into the SR region.

Note that a hidden assumption here is that the auxiliary features have smooth distributions over the SR in the background, for otherwise there would be no reason to believe that interpolation would give a sensible background estimate. This is often true in practice given that the SR is rather small.

Below we explain how the same steps can be achieved using boosted trees.

4.4.1 Boosted Density Estimation Trees

Motivated by the success of using BDTs with the CWoLa method, here we examine the possibility of applying them to density estimation. Specifically, we follow the tree density estimation algorithm presented in [141], which we describe briefly here. For details, please refer to the original literature.

Conceptually, the BDT density estimation algorithm is very similar to that of normalizing flows [142] — they both model the transformation between the target density and some base density as a composition of simple, bijective maps. Importantly, each composition is thought of as a round of boosting just as in the traditional algorithm.

The major difference between the two is that in the case of BDT, the transformations are built from cuts in the feature space (selected so that they locally minimize the KL divergence between the empirical distribution of the transformed data and the base distribution, which is uniform in our case) with Jacobians admitting closed-form evaluations, whereas for normalizing flows they are typically parameterized by neural networks [142]. The density estimated by the corresponding tree is then a leaf-wise constant function. After each round of boosting, one can define and compute the difference between the learned density and the target density, which is used as the target density for next round. This procedure is recursively performed until some termination condition is satisfied.

Copula

When estimating probability densities, it is often helpful to separate the task of estimating marginal densities and from the task of estimating the dependence structure between variables. This can be achieved explicitly by Sklar's theorem [143], which states, as part of the theorem, that any multivariate probability density $p(x_1, \ldots, x_d)$ (satisfying some very mild conditions) can be represented in the following form:

$$\begin{aligned} p(x_1, \ldots, x_d) &= c\big(F_1(x_1), \ldots, F_d(x_d)\big) f_1(x_1) \cdots f_d(x_d) \\ &\equiv \tilde{c}(x_1, \ldots, x_d) f_1(x_1) \cdots f_d(x_d)\,, \end{aligned} \tag{4.4}$$

where the f_i's are the marginal densities, the F_i's are the corresponding cumulative distribution function (CDF), and c is the so-called copula density function. This copula function completely encapsulates the information about dependence structure among variables.

The input data considered in this paper consists of the dijet mass m_{JJ}, the auxiliary features $x_1, \ldots, x_K$, and the additional features $y_1 \ldots y_N$ which contain no information relevant for anomaly detection. Moreover, we assume that irrelevant features are statistically independent of $(m_{JJ}, x_1, \ldots, x_K)$. With this assumption, the copula decomposition takes the form

$$\begin{aligned} p(m_{JJ}, x_1, \ldots, x_K, y_1, \ldots, y_N) = \tilde{c}(m_{JJ}, x_1, \ldots, x_K)\, \tilde{c}(y_1, \ldots, y_N) \\ \times f_1(x_1) \cdots f_K(x_K) g_1(y_1) \cdots g_N(y_N)\,. \end{aligned} \tag{4.5}$$

Furthermore, if the irrelevant features are mutually independent among themselves, as in Eq. (4.1), the corresponding copula function is trivial,

$$\tilde{c}(y_1, \ldots, y_N) = 1. \tag{4.6}$$

Then, the likelihood ratio in Eq. (4.3) takes the form

$$R = \frac{\tilde{c}(m_{JJ}, x_1, \ldots, x_K)}{\tilde{c}(m_{JJ}, x_1, \ldots, x_K|\text{bkgd})} \prod_{k=1}^{K} \frac{f(x_k)}{f(x_k|\text{bkgd})} \,. \tag{4.7}$$

Note that by using the copula decomposition, the dependence on the irrelevant features in R drops out in both the marginal and the copula densities. The cancellation in the marginal density ratio is easy to ensure in practice since it simply relies on univariate density estimation. As for the copula density, the model needs to be able to learn that it is independent of $(y_1 \ldots y_N)$. This is where the tree-structure shines — similar to the supervised case, the tree model should be able to learn to not cut along the irrelevant directions, since they do not contribute much towards the decrease in KL divergence when estimating the copula density.

In view of the discussion above, we follow the basic two-stage strategy suggested in [141]: we first fit models to the marginal variables, and then we use the learned CDF to transform them to the copula space on which we estimate the corresponding copula density. The final learned density is given by Eq. (4.4). Note that neither the copula factorization, Eq. (4.5), nor the mutual independence of the irrelevant features, Eq. (4.6), are hardwired into our algorithm. Rather, these features are efficiently learned by the BDT from the structure of the training data. The high quality of the trained tree model contributes to robustness of the anomaly detection algorithm in the presence of irrelevant features. At the same time, the underlying BDT has sufficient flexibility to remain useful when the structure of the input data is more complex. This is evidenced by the example with mutually dependent irrelevant features considered in App. B.3.

4.4.2 Interpolation

Once the probability density in the SB region is estimated, the next step is to interpolate it into the SR. Unlike the NN, a tree-based density estimator does not automatically provide such an interpolation, and it needs to be implemented by hand. This represents an additional step in the algorithm, but has an inherent advantage of being controllable, in contrast to a black box-like interpolation performed by the NN. As a baseline, we employ a naive linear interpolation:

$$p(\vec{x}|m) = p(\vec{x}|m_L) + \frac{p(\vec{x}|m_R) - p(\vec{x}|m_L)}{m_R - m_L}(m - m_L), \quad m \in (m_L, m_R)\,, \tag{4.8}$$

where $m \equiv m_{JJ}$; m_L and m_R are the lower and upper boundaries of the signal region in m_{JJ}; and the vector $\vec{x}$ includes both auxiliary (relevant) and irrelevant features. While more elaborate methods of interpolation exist [120, 123, 144], this simple form is chosen here for the following reasons:

- Under the assumption that the SR is sufficiently small and that the SB is not significantly signal-contaminated, we expect linear interpolation to give reasonable results (there are however some subtleties, see Sec. 4.4.5).
- More importantly, in Eq. (4.8), the interpolated density is explicitly linear in the learned density. In the ideal case that the irrelevant features' densities factorize from the learned density, this property ensures that dependence on irrelevant variables will be cancelled out in the construction of the likelihood ratio. As we shall see below, this linearity property is important in ensuring robustness.

	n_estimators	max_depth	lr	gamma
marginal	100	10	0.1	0.3
copula	2500	50	0.1	0.3

Table 4.2: Hyperparameters used to estimate the marginal and copula densities with the tree-based algorithm in [141]. Please refer to the original literature or App. B.2 for meanings of these parameters.

4.4.3 Training and Evaluation Procedures

The NN-based density estimator we use in our comparison is a masked-autoregressive flow (MAF). The training procedure for a MAF is the same as in [122], and we refer readers to the original paper for details.

The training of the tree-based density estimator is done by feeding the algorithm the entire training dataset ($560\,000$ events), with hyperparameters listed in Tab. 4.2. The performance of the BDT density estimation algorithm is fairly insensitive to the choice of hyperparameters as long as the resulting model is sufficiently expressive, due to the fact that we are estimating the density from the background sample for which we have large statistics. (This is in contrast with the CWoLa case, where the BDT needs to learn a small difference between the distributions of auxiliary features in the signal and side-band regions.) The preprocessing of [122] is not necessary in this case, since trees are invariant under monotonic transformations. The trained model is then used to evaluate $p(\vec{x}|m)$, $p(\vec{x}|m_L)$, and $p(\vec{x}|m_R)$. The latter two are used to estimate the background density in the signal region according to Eq. (4.8).

The performance of each method is evaluated on a separate test set consisting of $20\,000$ signal events and $60\,000$ background events, all of which lie in the SR. In particular, we train the tree-based model for 10 random training-validation-test splits to cross-validate its variance. The same is not done for neu-

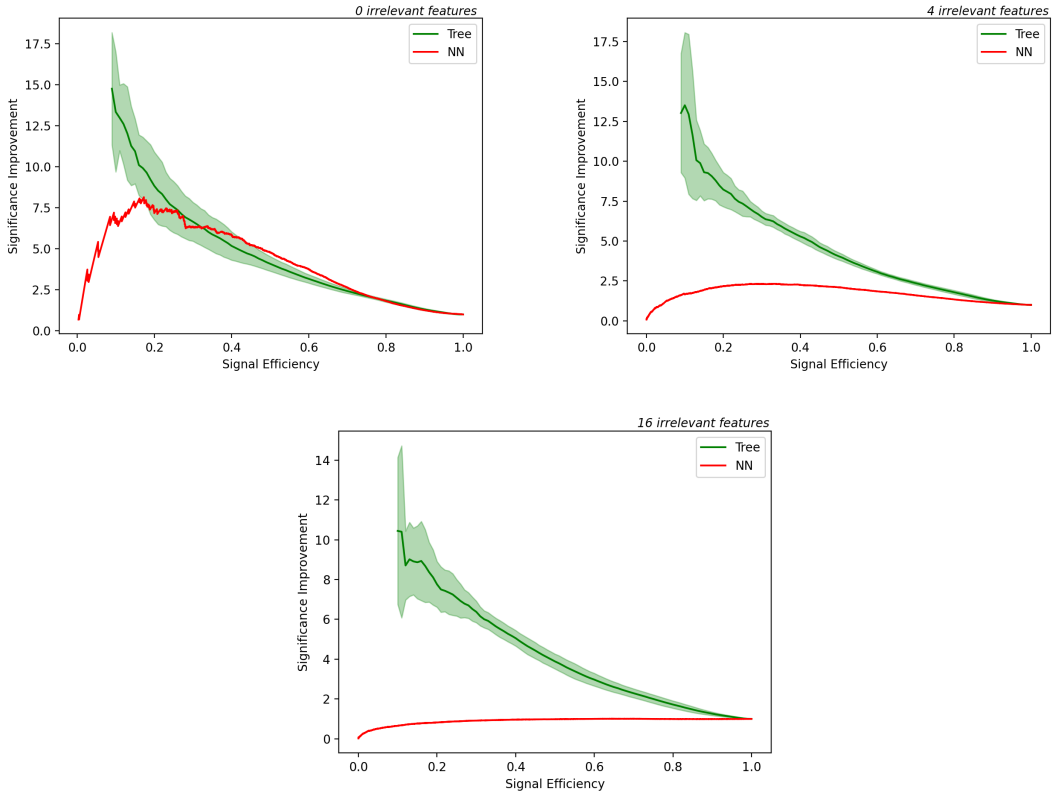


Figure 4.4: Performance comparison between the BDT implementation of density estimator (green) and the NN-based implementation found in [141] (red), for the original LHCO dataset. The error bands show 1 standard deviation of significance improvement across 10 random training-validation-test splits. The BDT implementation shows superior performance both with and without inclusion of irrelevant features. In particular, even with 16 irrelevant features added, the BDT only shows a small level of degradation.

ral networks due to their high computational costs. Comparisons are shown in the next section.

4.4.4 Performance Comparison

In Fig. 4.4, we show the performance comparisons between MAF- and tree-based density estimation algorithms. Without irrelevant features, the tree-based

algorithm already provides a significant improvement over the NN in the low signal efficiency region. Furthermore, just as in the case of CWoLa, the MAF-based algorithm suffers from severe performance degradation as the number of irrelevant features increases. In the case where 16 irrelevant features are added, the method is essentially no different from a random classifier. On the other hand, the tree-based algorithm is remarkably robust, showing almost no degradation of performance with up to 16 irrelevant features.

As an additional note, the success of the tree-based density estimation algorithm also shows that a simple linear interpolation for background estimation is very effective, at least for the LHCO dataset. In Fig. 4.5 we show a scatter plot of the data-to-background density ratio against m_{jj} for both signal and background events. It can be clearly seen that the simple linear interpolation is effective in estimating the background density for the LCHO dataset. We believe this is evidence that more considerations should go into studying the interpolation method instead of relying on a black box like NNs.

4.4.5 Correleated Auxiliary Features

The primary motivation for density estimation methods is to address situations where the CWoLa assumption of statistical independence between the auxiliary features $\vec{x}$ and m_{JJ} does not hold. However, in the example considered above, $\vec{x}$ and m_{JJ} were independent to a large degree. In this section we explore how our strategies proposed above perform when $\vec{x}$ and m_{JJ} are not independent.

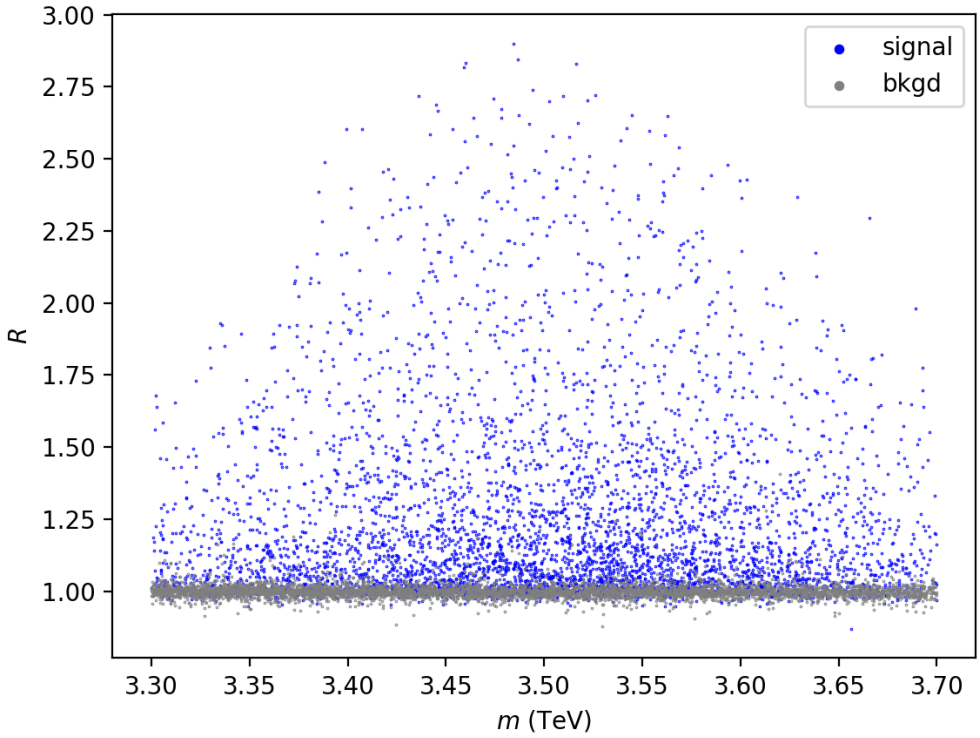


Figure 4.5: Scatter plot of the density ratio R defined in Eq. (4.3) against m_{jj} for 5000 signal events and 5000 background events in the test set (signal region). The plot shows that a naive linear interpolation is sufficient for the LHCO dataset.

Specifically, we artificially introduce dependence between $\vec{x}$ and m_{JJ} via[6]

$$m_{J_1} \to m_{J_1} + \log m_{JJ}, \tag{4.9}$$

$$\Delta m_J \to \Delta m_J + \log m_{JJ}, \tag{4.10}$$

where all the masses are measured in units of TeV.

In this case, we immediately see a difficulty with our proposed interpolation method. When $\vec{x}$ and m_{JJ} are strongly dependent, the support of $p(m, \vec{x})$ can be of arbitrary shape in general, but the interpolation in Eq. (4.8) implicitly as-

[6]We consider a non-polynomial dependence on m_{JJ} instead of a linear one, as in [119, 122], because our decorrelation scheme below will be able to completely undo linear correlation, thus making the comparison not very useful.

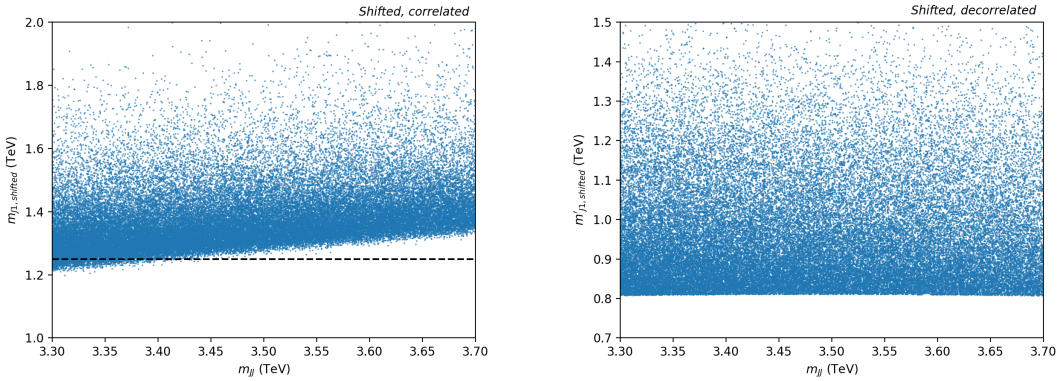


Figure 4.6: Scatter plot of m_{J1} against m_{JJ} for the log-shifted dataset over the signal region. Before decorrelation (left panel), the plot clearly shows why a naive linear interpolation should fail — the interpolation over the dashed line crosses the support of data density, which would cause a sharp change in the interpolated density. After decorrelation (right panel), we see that the support is roughly axis-parallel, and we expect that a simple linear interpolation should suffice.

sumes that for a fixed $\vec{x}$, $p(\vec{x}|m)$ does not vary too much as a function of m_{JJ} across the SR.[7] This is illustrated in the left panel of Fig. 4.6, where a naive linear interpolation over the dashed line would result in an abrupt and unphysical drop in the interpolated density. This situation can be handled automatically by NNs since they are able to perform more global interpolations, but we need to be more careful when implementing the interpolation by hand.

It is clear from the above discussion that the quality of linear interpolation Eq. (4.8) requires that $\vec{x}$ and m_{JJ} be roughly independent *over the SR*. To achieve this, we perform the following simple "decorrelation" procedure. For each fea-

[7]Note that this is different from the CWoLa assumption since we only require weak dependence over the SR. In general this is easier to attain.

ture x, consider the following transformation[8]:

$$m \to m\,, \tag{4.11}$$

$$x \to f(x, m)\,, \tag{4.12}$$

where f is such that the transformation is bijective so that no information carried in x is lost. We can then search for f such that the dependence between x and m_{JJ} within the SR, as measured by distance correlation, is minimized. In particular, we consider f belonging to a family of functions of the form

$$f(x, m) = a_0(m) + a_1(m)x. \tag{4.13}$$

Since m_{JJ} lies within the SR which we assume to be small, we further parameterize the coefficients a_0 and a_1 as[9]

$$a_0(m) = \alpha m + \beta m^2, \quad a_1(m) = 1 + \gamma m + \delta m^2\,. \tag{4.14}$$

To summarize, we search for values of α, β, γ and δ that minimize the corresponding distance correlation. This minimization is performed using the L-BFGS method [145] implemented in `SciPy` [146].

The right panel of Fig. 4.6 shows the scatter plot of m_{J1} against m_{JJ} over the SR after our decorrelation procedure. Visually we can observe that the support of data density is now parallel to the m_{JJ}-axis, and numerically we can achieve a distance correlation of order 10^{-4} between auxiliary features and m_{JJ} over the SR. This signals the success of our decorrelation scheme.

With decorrelation carried out, the rest of the algorithm remains the same as in the previous section. In Fig. 4.7 we compare the SIC curves of the NN-based

[8]We do not transform m_{JJ} since this is a privileged variable under the localized-signal assumption.

[9]Without loss of generality, we can take the constant term in a_0 to be zero and the constant term in a_1 to be 1, since distance correlation remains invariant under such a choice.

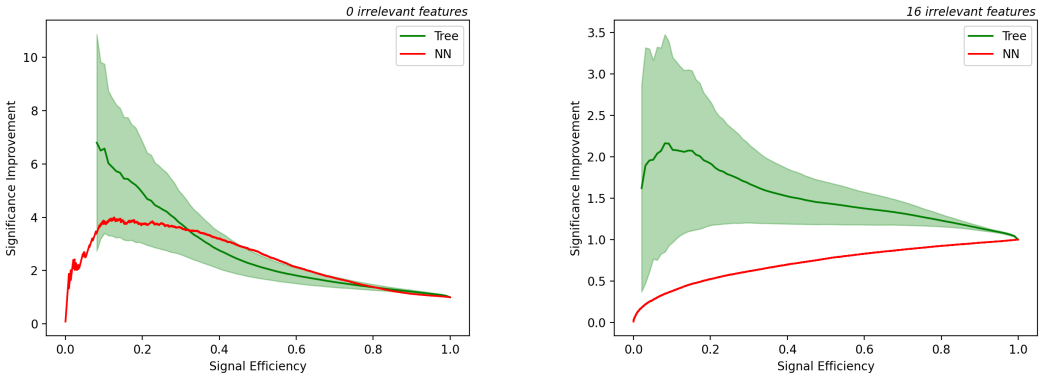


Figure 4.7: The SIC curves for NN-based (red) and tree-based (green) density-estimation algorithms applied to the log-shifted dataset. The cases with 0 (left panel) and 16 (right panel) irrelevant features are shown for comparison. The error bands are defined in the same way as in Fig. 4.4.

and tree-based algorithms applied to the log-shifted dataset. We observe that the tree-level algorithm still greatly outperforms the NN-based method when irrelevant features are added. At the same time, we also note that the performance of the tree-based algorithm is not as robust with respect to addition of irrelevant features as in the unshifted case (see Sec. 4.4.4). This is likely due to the decorrelation procedure above, which by chance will find non-zero $(\alpha, \beta, \gamma, \delta)$ such that the in-sample distance correlation between m_{JJ} and the transformed irrelevant feature is minimized. This effect can in principle be mitigated by more rigorous cross-validation technique, but we do not pursue this point here.

While the simple approach to decorrelation and interpolation taken in this paper is effective, it may be seen as somewhat *ad hoc*. Many more elaborate interpolation methods exist in the literature (*e.g.*, Gaussian process regression, high dimensional splines [138]), which may further improve the performance and robustness of our algorithm. We leave the exploration of such methods for future work.

4.5 Discussion and Conclusions

In this work, we have presented two tree-based approaches to detect anomalies in the presence of irrelevant features. Anomaly detection methods are already starting to be used in LHC analyses, with searches based on CWoLa hunting at ATLAS already released [147]. Since BDT-based methods are already used in experimental analyses, we hope that our methods would be readily able to be adopted and calibrated for experimental use. We first considered a CWoLa-inspired method, and showed that boosted decision trees are more robust to irrelevant features compared to neural networks. By exploiting the inherent feature selection of decision trees, the BDT-based classifier maintained good performance even with the addition of significantly more irrelevant than discriminating auxiliary features.

In analogy to density estimation methods like ANODE, we proposed using tree-based models paired with a copula transformation and interpolation step. By estimating the marginal and copula densities separately, irrelevant features can be factorized out of the likelihood ratio assuming their mutual independence. Even when this is not the case, we observe that the resulting reduction in significance improvement still leaves the tree-based approach much less sensitive to the presence of these features. Our results demonstrated the promising performance of the tree-based density estimator compared to normalizing flows, especially in higher dimensionality with many irrelevant features. The tree-based model allows for a simple and effective linear interpolation scheme for estimating the background density.

Recently, [127] also explored the use of BDTs for anomaly detection in high-

energy collider analyses. This study includes a larger and more physical set of irrelevant features, while also finding increasingly improved performance and greater stability as irrelevant features are added during training. However, it assumes that a perfect sample of the background is available and does not deal with the extrapolation of such a model into a resonant region, as we do. We thus view our results as complementary and together making a compelling case for the application of tree-based methods to anomaly detection.

Overall, tree-based methods seem well-suited for anomaly detection tasks when operating on high-level observables with potential irrelevant features. These naturally lend themselves to presentation as tabular data. The techniques presented here could find useful application in collider searches and other physics analyses aiming to be robust against the embedding of low-dimensional signals in high-dimensional feature spaces. More advanced interpolation schemes than what we consider here might improve the performance and stability of the density-based approach, while exploring other tree-based algorithms like Bayesian Additive Regression Trees might improve the overall fidelity of the learned functions. We leave these possibilities to future work.

APPENDIX A

APPENDIX TO CHAPTER 3

A.1 Simplified Model and Feynman Rules

Our effective dark Z model is parameterized by the dark Z mass, $m_{A'}$, its kinetic mixing to hypercharge, ε, and a dimensionless Z–dark Z mass mixing parameter, κ. These are related to a benchmark ultraviolet model in Section 3.2. We use the shorthand $s_W = \sin\theta_W$, and similarly for cosine, tangent, and the massive gauge boson mixing angle δ, (3.6).

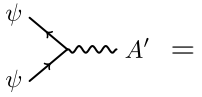
Mass Spectrum. The mass spectrum comes from diagonalizing the mass terms (3.4). Assuming $m_{A'} > m_Z$, the Z boson mass to second order in the new physics parameters is

$$m_Z^2 = M_Z(v)^2 \left[1 - \frac{2M_Z(v)^2}{m_{A'}^2 - M_Z(v)^2}(\varepsilon t_W + \kappa)^2\right] \qquad M_Z(v)^2 \equiv \frac{g^2 v^2}{4c_W^2}\,, \tag{A.1}$$

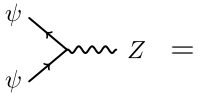
where $v^2 = (246\text{ GeV})^2$ is the order parameter for electroweak symmetry breaking. For completeness, the dark Z mass is related to the benchmark ultraviolet model parameters to leading order in the new physics parameters by

$$m_{A'}^2 = M_{A'}^2 + \frac{2M_Z(v)^4}{M_{A'}^2 - M_Z(v)^2}(\varepsilon t_W + \kappa)^2 \qquad M_{A'}^2 = g_{\rm d}^2 v_{\rm d}^2 + \frac{\kappa}{2q_{\rm d}}\frac{g g_{\rm d}}{c_W} v^2\,. \tag{A.2}$$

Feynman Rules. The interactions with Standard Model fermions ψ are described by



$$= i\gamma^\mu \left(g^{\rm V} + g^{\rm A}\gamma^5\right)$$



$$= i\frac{g}{c_W}\gamma^\mu \left(q_Z^{\rm V} + q_Z^{\rm A}\gamma^5\right). \tag{A.3}$$

The effective couplings and charges to leading non-trivial order in ε and κ are

$$g^{\mathrm{V}} = \varepsilon e Q_{\mathrm{EM}} - \frac{g}{c_W} \frac{m_Z^2 \kappa + m_{A'}^2 \varepsilon t_W}{m_{A'}^2 - m_Z^2} \left(\frac{1}{2} T^3 - s_W^2 Q_{\mathrm{EM}} \right) \tag{A.4}$$

$$g^{\mathrm{A}} = \frac{g}{c_W} \frac{m_Z^2 \kappa + m_{A'}^2 \varepsilon t_W}{m_{A'}^2 - m_Z^2} \frac{1}{2} T^3 \tag{A.5}$$

$$q_Z^{\mathrm{V}} = \quad (c_\delta + \varepsilon t_W s_\delta) \left(\frac{1}{2} T^3 - s_W^2 Q_{\mathrm{EM}} \right) + \frac{\varepsilon e}{g} c_W s_\delta \, Q_{\mathrm{EM}} \tag{A.6}$$

$$q_Z^{\mathrm{A}} = -\left(c_\delta + \varepsilon t_W s_\delta\right) \frac{1}{2} T^3 \, , \tag{A.7}$$

where $T^3 = \pm 1/2$ according to the $\mathrm{SU}(2)_{\mathrm{L}}$ weight of the left-chiral ψ and Q_{EM} is the ψ electric charge. Additional interactions between the A', Higgs, and electroweak gauge bosons are not used in this study. The effects are well approximated using the Goldstone boson equivalence theorem, as discussed in Section 3.2.4. See Ref. [89] for explicit expressions in the pure kinetic mixing case.

A.2 UV Model and Decoupling Limit

In Section 3.2.2 we derive our three-parameter model for mass and kinetic mixing as the low-energy theory of an extended Higgs sector. This appendix addresses some model-building aspects of the UV model and connects to known results in the two-Higgs doublet literature.

Decoupling pseudo-Goldstone modes. There are a total of seven Goldstone modes. Four are eaten by the massive Standard Model gauge bosons and the dark Z. We decouple the remaining three by adding a trilinear potential that

explicitly breaks the global symmetries of the Higgs sector:

$$V \supset \mu' H_{\text{EW}}^{\dagger} H_{\text{mix}} H_{\text{d}} + \text{h.c.} \,. \tag{A.8}$$

The scale μ' is assumed to be large compared to the masses in the three-parameter model. The uneaten Goldstones are an orthogonal linear combination to the eaten combination; see Appendix A of Ref. [148].

Decoupling the dark Higgs . We may tune the H_{d} potential such that that CP-even state has the largest mass in the UV theory, we define this to be Λ. When integrating out this mode, the trilinear coupling (A.8) generates four-point interactions between the Higgs doublets with a quartic coupling of order $(\mu'/\Lambda)^2$. In a given UV model, the perturbativity of the lower-energy two-Higgs doublet theory requires this coupling to be smaller than $\sim 4\pi$.

The decoupling limit of the two-Higgs doublet model. Upon integrating out H_d, our Higgs sector is a CP-conserving two-Higgs doublet model. Gunion and Haber examine the viability of a decoupling limit these models in [149]. In their notation, $\Phi_1 = H_{\text{EW}}$ and $\Phi_2 = H_{\text{mix}}$ and the parameters of their general potential are:

$$m_{12}^2 = \frac{\mu' v_d}{\sqrt{2}} \qquad \lambda_{4,5} = \frac{\mu'^2}{\Lambda^2} \qquad \lambda_{3,6,7} = 0 \,. \tag{A.9}$$

The diagonal quadratic and quartic terms are set to produce symmetry-breaking potentials for each of $\Phi_{1,2}$ that yield the appropriate vacuum expectation values, v_{EW} and v_{mix}, with mixing angle $t_\beta = v_{\text{mix}}/v_{\text{EW}} \ll 1$. The mixing angle β encodes the same information as our mass mixing parameter, κ, (3.5). This potential is manifestly stable in all field directions. The masses for the CP-odd and charged

Higgses—the uneaten would-be Goldstones–are

$$m_A^2 = m_{H^\pm}^2 = \frac{\mu' v_d}{\sqrt{2} s_\beta c_\beta} - \frac{v^2 \mu'^2}{\Lambda^2} \gg v^2 \,. \tag{A.10}$$

The mass matrix between CP-even Higgses is, to leading order in $s_\beta \ll 1$,

$$\mathcal{M}^2 = \begin{pmatrix} \lambda_{\rm EW} v^2 & \left(\frac{\mu'^2}{\Lambda^2} - m_A^2\right) s_\beta \\ \left(\frac{\mu'^2}{\Lambda^2} - m_A^2\right) s_\beta & m_A^2 \end{pmatrix} \tag{A.11}$$

where $\lambda_{\rm EW}$ is the quartic coupling for $\Phi_1 = H_{\rm EW}$. The rotation to diagonalize the CP-even mass matrix is α. The lighter eigenstate is identified with the 125 GeV Higgs boson. The decoupling limit corresponds to the case where an entire $\mathrm{SU(2)_L}$ doublet is heavy. It corresponds to the limit $(\beta - \alpha) \to \pi/2$, or $m_A^2 \gg |\lambda_{\rm EW}| v^2$. This is readily realized in our model.

Alternative UV completions. The minimal model in Refs. [87, 82] avoids any extraneous would-be Goldstones. Another alternative is to introduce a Stuckelberg [150, 151] mass for the Abelian hidden sector.

A.3 Unpolarized Observables

The angular dependence of the differential $e^+e^- \to A' \to \mu^+\mu^-$ cross section with spin-averaged $e^\pm$ beams is

$$\frac{d\sigma}{d\Omega} = \frac{f(s)}{4} \left[\left(g_{\rm L}^4 + g_{\rm R}^4\right) (1 + \cos\theta)^2 + 2 g_{\rm L}^2 g_{\rm R}^2 (1 - \cos\theta)^2 \right] \,, \tag{A.12}$$

where $f(s)$ is a kinematic factor independent of the scattering angle of the outgoing muon with respect to the electron beam, $\cos\theta$. The forward and backward and cross sections are

$$\sigma_{\rm F} \equiv \int_0^1 d\cos\theta \, \frac{d\sigma}{d\cos\theta} \qquad\qquad \sigma_{\rm B} \equiv \int_{-1}^0 d\cos\theta \frac{d\sigma}{d\cos\theta} \,. \tag{A.13}$$

The forward–backward asymmetry is expressed in terms of a coupling asymmetry A_e,

$$A_{\rm FB}^0 \equiv \frac{\sigma_{\rm F} - \sigma_{\rm B}}{\sigma_{\rm F} + \sigma_{\rm B}} = \frac{3}{4} A_e^2 \qquad\qquad A_e \equiv \frac{g_{\rm L}^2 - g_{\rm R}^2}{g_{\rm L}^2 + g_{\rm R}^2} = \frac{2 q_{\rm A} q_{\rm V}}{q_{\rm A}^2 + q_{\rm V}^2} \,, \tag{A.14}$$

where the superscript on $A_{\rm FB}$ indicates an unpolarized asymmetry.

A.4 Polarized Beams and Observables

We review the collider phenomenology of polarized beams following the conventions in Ref. [152]. Define the longitudinal electron beam polarization to be $P = (R - L)/(R + L)$ where R (L) is the number of right (left) helicity particles in the beam. Similarly, let $\bar{P}$ be the analogous quantity for the positron beam. The polarized cross section $\sigma_{P\bar{P}}$ is related to the pure helicity cross sections $\sigma_{\rm R\bar{L}}$ and $\sigma_{\rm L\bar{R}}$ as

$$\sigma_{P\bar{P}} = \frac{1}{4}(1+P)(1-\bar{P})\sigma_{\rm R\bar{L}} + \frac{1}{4}(1-P)(1+\bar{P})\sigma_{\rm L\bar{R}} \,, \tag{A.15}$$

where we used $\sigma_{\rm L\bar{L}} = \sigma_{\rm R\bar{R}}$ in the limit of massless fermions coupled to a a vector boson. It is convenient to write this in terms of the symmetric (anti-symmetric) combinations

$$\sigma_{\rm (LR)} = \sigma_{\rm L\bar{R}} + \sigma_{\rm R\bar{L}} = 4\sigma \qquad\qquad \sigma_{\rm [LR]} = \sigma_{\rm L\bar{R}} - \sigma_{\rm R\bar{L}} \,, \tag{A.16}$$

where we note that the spin-averaged (unpolarized) cross section, σ, is one fourth of $\sigma_{\rm (LR)}$.[1] This gives a compact expression for the polarized cross section

[1] The unpolarized cross section may be written in terms of helicity cross sections as

$$\frac{d\sigma}{d\Omega} = \frac{1}{4}\left(\frac{d\sigma}{d\Omega}\bigg|_{\rm RL\to RL} + \frac{d\sigma}{d\Omega}\bigg|_{\rm LR\to LR} + \frac{d\sigma}{d\Omega}\bigg|_{\rm RL\to LR} + \frac{d\sigma}{d\Omega}\bigg|_{\rm LR\to RL} \right) .$$

in terms of an effective luminosity $\mathcal{L}_{\text{eff}}$ and an effective polarization P_{eff}

$$\sigma_{P\bar{P}} \equiv 2\sigma \frac{\mathcal{L}_{\text{eff}}}{\mathcal{L}} \left[1 - P_{\text{eff}} A_{\text{LR}}\right] \qquad \frac{\mathcal{L}_{\text{eff}}}{\mathcal{L}} \equiv \frac{1 - P\bar{P}}{2} \qquad P_{\text{eff}} \equiv \frac{P - \bar{P}}{1 - P\bar{P}} \,. \tag{A.17}$$

The effective polarization is zero for unpolarized beams and ± 1 for perfectly polarized beams.

The left–right asymmetry for perfectly polarized beams is $A_{\text{LR}}^{\text{perfect}} = \sigma_{[LR]}/\sigma_{(LR)}$. For partially polarized beams, we assume that one beam is mostly left-helicity while the other is mostly-right helicity. The cross section the mostly-left helicity electron configuration, σ_{-+}, is related to that of the mostly-right helicity electron configuration, σ_{+-}, by the replacements $P \to -P$ and $\bar{P} \to -\bar{P}$. The left–right asymmetry for the partially polarized beams is

$$A_{\text{LR}} = \frac{1}{P_{\text{eff}}} \frac{\sigma_{[-+]}}{\sigma_{(-+)}} = \frac{1}{P_{\text{eff}}} \left(\frac{|P| + |\bar{P}|}{1 + |P\bar{P}|} \right) \frac{\sigma_{[\text{LR}]}}{\sigma_{(\text{LR})}} = \frac{1}{P_{\text{eff}}} \left(\frac{|P| + |\bar{P}|}{1 + |P\bar{P}|} \right) A_e \,, \tag{A.18}$$

where $\sigma_{(-+)}$ and $\sigma_{[-+]}$ are the symmetric and antisymmetric combinations of the partially polarized cross sections analogous to (A.16). These are related to the helicity cross sections by

$$\sigma_{(-+)} = \frac{1}{2} \left(1 + |P\bar{P}|\right) \sigma_{(\text{LR})} \qquad \sigma_{[-+]} = \frac{1}{2} \left(|P| + |\bar{P}|\right) \sigma_{[\text{LR}]} \,. \tag{A.19}$$

One may measure the forward–backward asymmetry for polarized beams. In this case, (A.12) is replaced with

$$\frac{d\sigma}{d\Omega} = \frac{f(s)}{4} \left[g_L^4 (1 - P)(1 + \bar{P}) + g_R^4 (1 + P)(1 - \bar{P}) \right] (1 + \cos\theta)^2 \tag{A.20}$$

$$+ \frac{f(s)}{4} g_L^2 g_R^2 \left[(1 + P)(1 - \bar{P}) + (1 - P)(1 + \bar{P}) \right] (1 - \cos\theta)^2 \,. \tag{A.21}$$

The resulting forward-backward asymmetry for partially polarized beams is

$$A_{\text{FB}} = \frac{3}{4} \frac{(g_{\text{R}}^2 - g_{\text{L}}^2)^2 + P_{\text{eff}}(g_{\text{R}}^4 - g_{\text{L}}^4)}{(g_{\text{R}}^2 + g_{\text{L}}^2)^2 + P_{\text{eff}}(g_{\text{R}}^4 - g_{\text{L}}^4)} = \frac{3}{4} \frac{(A_e - P_{\text{eff}}) A_e}{1 - P_{\text{eff}} A_e} \,. \tag{A.22}$$

APPENDIX B

APPENDIX TO CHAPTER 4

B.1 Hyperparameter Tuning for `xgboost`

Here we provide details regarding hyperparameter tuning for the `xgboost` model used in CWoLa method. The hyperparameters we choose to optimize are as follows[1]:

- `n_estimators`: this controls the number of boosting rounds
- `max_depth`: this controls how complex the base tree learner is by limiting how deep each tree can be
- `eta`: this controls how much each tree contributes in building the ensemble
- `alpha`: L_1 regularizer on weights of the model
- `lambda`: L_2 regularizer on weights of the model

`xgboost` has a lot of other parameters, but here we choose to focus on these few because (i) `n_estimators`, `max_depth` and `eta` are known to have the most impact on the model's performance, and (ii) `alpha` and `lambda` explicitly control the model's weights, and therefore they have a direct impact on how much the model will overfit, which is exactly our concern here. In addition, we (arbitrarily) fix the `subsample` parameter, which measures how much of the training data is used in fitting each individual tree, to be 0.75. In principle one can also include it in the hyperparameter search but our empirical results

[1] More details about these hyperparameters can be found in [139]

show that the final performance is not very dependent on its exact value. We use default values for all other hyperparameters.

To search for the optimal hyperparameters, we perform Bayesian optimization on the 10-fold cross-validation score, which we define to be the true positive rate of SR-SB labels at a fixed false positive rate of 10^{-3}. Specifically, the Bayesian optimization is carried out using the `gp_minimize` function in the `scikit-optimize` library, with default settings except we reduce the number of calls to 30 in order to save time.

B.2 Hyperparameters of Boosted Density Estimation Tree Algorithm

Here we describe briefly meanings of some of the hyperparameters used when training on the LHCO dataset. Please refer to [141] for details.

- `n_estimators`: number of boosting rounds
- `max_depth`: maximum depth each base tree learner can grow to
- `lr`: global shrinkage parameter that helps smooth out density learned during each boosting round. When it is equal to 0, each tree returns the uniform base distribution (no learning); when it is equal to 1, each tree returns the empirical distribution (most aggressive learning).
- `gamma`: amount of node-specific shrinkage. When it is 0, only the global learning rate `lr` is used; when it is a positive real number, the amount of shrinkage for each node grows as its volume in feature space decreases.

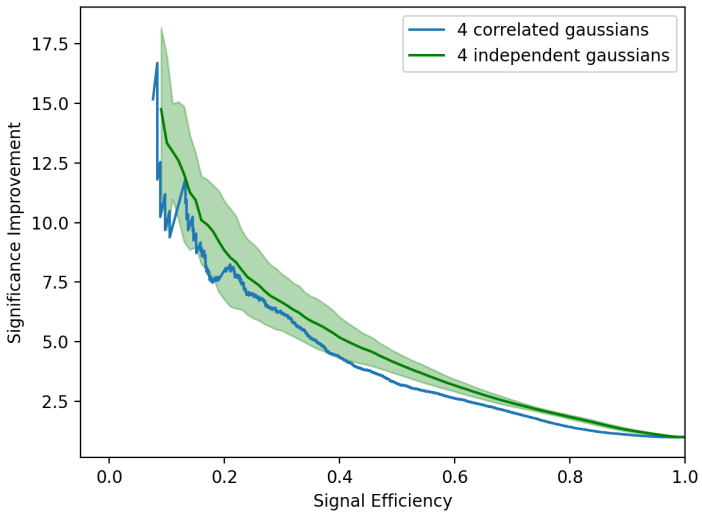


Figure B.1: SIC curve of the tree-based density estimation algorithm with mutually correlated irrelevant features (blue), compared to the baseline case of mutually independent irrelevant features studied in Sec. 4.4 (green).

B.3 Mutually Dependent Irrelevant Features

In the baseline model used throughout this paper, the irrelevant features enjoy the extra property that they are mutually independent, see Eq. (4.1). While this extra property has no bearing on the CWoLa hunting method, it does affect our use of copula in Sec. 4.4: If the irrelevant features $\vec{y}$ are mutually independent, the copula density c becomes independent of $\vec{y}$. While such independence was not hardwired into our algorithm, it can potentially make the copula density easier to learn, and one might wonder how robust the algorithm is if irrelevant features are mutually dependent.

To test this, we rotate the original irrelevant features by a random matrix A:

$$\vec{y}\,' = A\vec{y}\,. \tag{B.1}$$

Here A is constructed by independently sampling each of its elements from the

standard normal distribution. The elements of the rotated irrelevant feature vector $\vec{y}\,'$ are now mutually dependent. We then apply the tree-based density estimation algorithm described in Sec. 4.4 to the dataset $(m_{JJ}, \vec{x}, \vec{y}\,')$, where $\vec{x}$ are the relevant auxiliary features.

The resulting SIC curve is shown in Fig. B.1. We can see that our method's performance is very similar to the case considered in the main text, demonstrating that the method's performance is not reliant on the factorization property of Eq. (4.1). In other words, the BDT is able to learn the non-trivial copula function involving irrelevant features well enough to not cause any degradation in the overall performance. In future work, it would be interesting to further test this aspect of the algorithm in realistic physical applications of anomaly detection.

BIBLIOGRAPHY

[1] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Phys. Rev. Lett.* 81 (8 Aug. 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[2] Seth Koren. "New Approaches to the Hierarchy Problem and their Signatures from Microscopic to Cosmic Scales". PhD thesis. UC, Santa Barbara (main), 2020. arXiv: 2009.11870 [hep-ph].

[3] Jihn E. Kim and Gianpaolo Carosi. "Axions and the Strong CP Problem". In: *Rev. Mod. Phys.* 82 (2010). [Erratum: Rev.Mod.Phys. 91, 049902 (2019)], pp. 557–602. DOI: 10.1103/RevModPhys.82.557. arXiv: 0807.3125 [hep-ph].

[4] Glennys R. Farrar. "A Stable Sexaquark: Overview and Discovery Strategies". In: (Jan. 2022). arXiv: 2201.01334 [hep-ph].

[5] Riccardo Scarpa. "Modified newtonian dynamics, an introductory review". In: *AIP Conf. Proc.* 822.1 (2006). Ed. by Eric J. Lerner and Jose B. Almeida, pp. 253–265. DOI: 10.1063/1.2189141. arXiv: astro-ph/0601478.

[6] Marco Cirelli, Alessandro Strumia, and Jure Zupan. "Dark Matter". In: (June 2024). arXiv: 2406.01705 [hep-ph].

[7] Vera C. Rubin and Jr. Ford W. Kent. "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions". In: 159 (Feb. 1970), p. 379. DOI: 10.1086/150317.

[8] Yoshiaki Sofue and Vera Rubin. "Rotation curves of spiral galaxies". In: *Ann. Rev. Astron. Astrophys.* 39 (2001), pp. 137–174. DOI: 10.1146/annurev.astro.39.1.137. arXiv: astro-ph/0010594.

[9] R. D. Peccei and Helen R. Quinn. "Constraints imposed by CP conservation in the presence of pseudoparticles". In: *Phys. Rev. D* 16 (6 Sept. 1977), pp. 1791–1797. DOI: 10.1103/PhysRevD.16.1791. URL: https://link.aps.org/doi/10.1103/PhysRevD.16.1791.

[10] Michela Paganini. "Machine Learning Algorithms for b-Jet Tagging at the ATLAS Experiment". In: *J. Phys. Conf. Ser.* 1085.4 (2018), p. 042031. DOI: 10.1088/1742-6596/1085/4/042031. arXiv: 1711.08811 [hep-ex].

[11] Vinicius Mikuni and Florencia Canelli. "Unsupervised clustering for collider physics". In: *Phys. Rev. D* 103.9 (2021), p. 092007. DOI: 10.1103/PhysRevD.103.092007. arXiv: 2010.07106 [physics.data-an].

[12] Sydney Otten et al. "Event Generation and Statistical Sampling for Physics with Deep Generative Models and a Density Information Buffer". In: *Nature Commun.* 12.1 (2021), p. 2985. DOI: 10.1038/s41467-021-22616-z. arXiv: 1901.00875 [hep-ph].

[13] G. Peter Lepage. "A New Algorithm for Adaptive Multidimensional Integration". In: *J. Comput. Phys.* 27 (1978), p. 192. DOI: 10.1016/0021-9991(78)90004-9.

[14] Matthew D. Klimek and Maxim Perelstein. "Neural Network-Based Approach to Phase Space Integration". In: *SciPost Phys.* 9 (2020), p. 053. DOI: 10.21468/SciPostPhys.9.4.053. arXiv: 1810.11509 [hep-ph].

[15] Christina Gao, Joshua Isaacson, and Claudius Krause. "i-flow: High-dimensional Integration and Sampling with Normalizing Flows". In: *Mach. Learn. Sci. Tech.* 1.4 (2020), p. 045023. DOI: 10.1088/2632-2153/abab62. arXiv: 2001.05486 [physics.comp-ph].

[16] Mathias Backes et al. "How to GAN Event Unweighting". In: *SciPost Phys.* 10.4 (2021), p. 089. DOI: 10.21468/SciPostPhys.10.4.089. arXiv: 2012.07873 [hep-ph].

[17] Randolf Pohl et al. "The size of the proton". In: *Nature* 466 (2010), pp. 213–216. DOI: 10.1038/nature09250.

[18] Aldo Antognini et al. "Proton Structure from the Measurement of $2S - 2P$ Transition Frequencies of Muonic Hydrogen". In: *Science* 339 (2013), pp. 417–420. DOI: 10.1126/science.1230016.

[19] Peter J. Mohr, Barry N. Taylor, and David B. Newell. "CODATA Recommended Values of the Fundamental Physical Constants: 2006". In: *Rev. Mod. Phys.* 80 (2008), pp. 633–730. DOI: 10.1103/RevModPhys.80.633. arXiv: 0801.0028 [physics.atom-ph].

[20] Peter J. Mohr, David B. Newell, and Barry N. Taylor. "CODATA Recommended Values of the Fundamental Physical Constants: 2014". In: *Rev. Mod. Phys.* 88.3 (2016), p. 035009. DOI: 10.1103/RevModPhys.88.035009. arXiv: 1507.07956 [physics.atom-ph].

[21] R. Pohl *et al.* "Laser spectroscopy of muonic deuterium". In: *Science* 353.6300 (2016), pp. 669–673. ISSN: 0036-8075. DOI: 10.1126/science.aaf2468. eprint: https://science.sciencemag.org/content/353/6300/669.full.pdf. URL: https://science.sciencemag.org/content/353/6300/669.

[22] W. Xiong et al. “A small proton charge radius from an electron–proton scattering experiment”. In: *Nature* 575.7781 (2019), pp. 147–150. DOI: 10.1038/s41586-019-1721-2.

[23] Axel Beyer et al. “The Rydberg constant and proton size from atomic hydrogen”. In: *Science* 358.6359 (2017), pp. 79–85. DOI: 10.1126/science.aah6677.

[24] N. Bezginov et al. “A measurement of the atomic hydrogen Lamb shift and the proton charge radius”. In: *Science* 365.6457 (2019), pp. 1007–1012. DOI: 10.1126/science.aau7807.

[25] Hélène Fleurbaey et al. “New Measurement of the 1S-3S Transition Frequency of Hydrogen: Contribution to the Proton Charge Radius Puzzle”. In: *Phys. Rev. Lett.* 120.18 (2018), p. 183001. DOI: 10.1103/PhysRevLett.120.183001. arXiv: 1801.08816 [physics.atom-ph].

[26] David Tucker-Smith and Itay Yavin. “Muonic hydrogen and MeV forces”. In: *Phys. Rev. D* 83 (2011), p. 101702. DOI: 10.1103/PhysRevD.83.101702. arXiv: 1011.4922 [hep-ph].

[27] Brian Batell, David McKeen, and Maxim Pospelov. “New Parity-Violating Muonic Forces and the Proton Charge Radius”. In: *Phys. Rev. Lett.* 107 (2011), p. 011803. DOI: 10.1103/PhysRevLett.107.011803. arXiv: 1103.0721 [hep-ph].

[28] Carl E. Carlson and Benjamin C. Rislow. “New Physics and the Proton Radius Problem”. In: *Phys. Rev. D* 86 (2012), p. 035013. DOI: 10.1103/PhysRevD.86.035013. arXiv: 1206.3587 [hep-ph].

[29] Savely G. Karshenboim, David McKeen, and Maxim Pospelov. "Constraints on muon-specific dark forces". In: *Phys. Rev. D* 90.7 (2014). [Addendum: Phys.Rev.D 90, 079905 (2014)], p. 073004. DOI: 10.1103/PhysRevD.90.073004. arXiv: 1401.6154 [hep-ph].

[30] Yu-Sheng Liu, David McKeen, and Gerald A. Miller. "Electrophobic Scalar Boson and Muonic Puzzles". In: *Phys. Rev. Lett.* 117.10 (2016), p. 101801. DOI: 10.1103/PhysRevLett.117.101801. arXiv: 1605.04612 [hep-ph].

[31] Sylvain Fichet. "Quantum Forces from Dark Matter and Where to Find Them". In: *Phys. Rev. Lett.* 120.13 (2018), p. 131801. DOI: 10.1103/PhysRevLett.120.131801. arXiv: 1705.10331 [hep-ph].

[32] Philippe Brax, Sylvain Fichet, and Guillaume Pignol. "Bounding Quantum Dark Forces". In: *Phys. Rev.* D97.11 (2018), p. 115034. DOI: 10.1103/PhysRevD.97.115034. arXiv: 1710.00850 [hep-ph].

[33] G.W. Bennett et al. "Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL". In: *Phys. Rev. D* 73 (2006), p. 072003. DOI: 10.1103/PhysRevD.73.072003. arXiv: hep-ex/0602035.

[34] Michel Davier et al. "Reevaluation of the Hadronic Contributions to the". In: *Eur. Phys. J. C* 71 (2011). [Erratum: Eur.Phys.J.C 72, 1874 (2012)], p. 1515. DOI: 10.1140/epjc/s10052-012-1874-8. arXiv: 1010.4180 [hep-ph].

[35] Kaoru Hagiwara et al. "re-evaluated using new precise data". In: *J. Phys. G* 38 (2011), p. 085003. DOI: 10.1088/0954-3899/38/8/085003. arXiv: 1105.3149 [hep-ph].

[36] Randolf Pohl et al. “Deuteron charge radius and Rydberg constant from spectroscopy data in atomic deuterium”. In: *Metrologia* 54.2 (Mar. 2017), pp. L1–L10. DOI: 10.1088/1681-7575/aa4e59. URL: https://doi.org/10.1088%5C%2F1681-7575%5C%2Faa4e59.

[37] N. Aghanim et al. “Planck 2018 results. VI. Cosmological parameters”. In: (2018). arXiv: 1807.06209 [astro-ph.CO].

[38] Edward W. Kolb and Michael S. Turner. “The Early Universe”. In: *Front. Phys.* 69 (1990), pp. 1–547.

[39] Alexander Kusenko and Leslie J. Rosenberg. *Snowmass-2013 Cosmic Frontier 3 (CF3) Working Group Summary: Non-WIMP dark matter*. 2013. arXiv: 1310.8642 [hep-ph].

[40] Jesús Zavala, Mark Vogelsberger, and Matthew G. Walker. “Constraining self-interacting dark matter with the Milky Way?s dwarf spheroidals”. In: *Monthly Notices of the Royal Astronomical Society: Letters* 431.1 (Feb. 2013), L20?L24. ISSN: 1745-3925. DOI: 10.1093/mnrasl/sls053. URL: http://dx.doi.org/10.1093/mnrasl/sls053.

[41] Jim Alexander et al. In: *Dark Sectors 2016 Workshop: Community Report*. Aug. 2016. arXiv: 1608.08632 [hep-ph].

[42] Marco Battaglieri et al. “US Cosmic Visions: New Ideas in Dark Matter 2017: Community Report”. In: *U.S. Cosmic Visions: New Ideas in Dark Matter*. July 2017. arXiv: 1707.04591 [hep-ph].

[43] Torsten Bringmann and Maxim Pospelov. “Novel direct detection constraints on light dark matter”. In: *Phys. Rev. Lett.* 122.17 (2019), p. 171801. DOI: 10.1103/PhysRevLett.122.171801. arXiv: 1810.10543 [hep-ph].

[44] Christopher Cappiello and John F. Beacom. "Strong New Limits on Light Dark Matter from Neutrino Experiments". In: *Phys. Rev. D* 100.10 (2019), p. 103011. DOI: 10.1103/PhysRevD.100.103011. arXiv: 1906.11283 [hep-ph].

[45] Masahiro Ibe et al. "Migdal Effect in Dark Matter Direct Detection Experiments". In: *JHEP* 03 (2018), p. 194. DOI: 10.1007/JHEP03(2018)194. arXiv: 1707.07258 [hep-ph].

[46] Matthew J. Dolan, Felix Kahlhoefer, and Christopher McCabe. "Directly detecting sub-GeV dark matter with electrons from nuclear scattering". In: *Phys. Rev. Lett.* 121.10 (2018), p. 101801. DOI: 10.1103/PhysRevLett.121.101801. arXiv: 1711.09906 [hep-ph].

[47] Nicole F. Bell et al. "Migdal effect and photon bremsstrahlung in effective field theories of dark matter direct detection and coherent elastic neutrino-nucleus scattering". In: *Phys. Rev. D* 101.1 (2020), p. 015012. DOI: 10.1103/PhysRevD.101.015012. arXiv: 1905.00046 [hep-ph].

[48] Daniel Baxter, Yonatan Kahn, and Gordan Krnjaic. "Electron Ionization via Dark Matter-Electron Scattering and the Migdal Effect". In: *Phys. Rev. D* 101.7 (2020), p. 076014. DOI: 10.1103/PhysRevD.101.076014. arXiv: 1908.00012 [hep-ph].

[49] Rouven Essig et al. "Relation between the Migdal Effect and Dark Matter-Electron Scattering in Isolated Atoms and Semiconductors". In: *Phys. Rev. Lett.* 124.2 (2020), p. 021801. DOI: 10.1103/PhysRevLett.124.021801. arXiv: 1908.10881 [hep-ph].

[50] Liron Barak et al. *SENSEI: Direct-Detection Results on sub-GeV Dark Matter from a New Skipper-CCD*. 2020. arXiv: 2004.11378 [astro-ph.CO].

[51] Tracy R. Slatyer. "Indirect dark matter signatures in the cosmic dark ages. I. Generalizing the bound on s-wave dark matter annihilation from Planck results". In: *Phys. Rev. D* 93.2 (2016), p. 023527. DOI: 10.1103/PhysRevD.93.023527. arXiv: 1506.03811 [hep-ph].

[52] Kimberly K. Boddy and Vera Gluscevic. "First Cosmological Constraint on the Effective Theory of Dark Matter-Proton Interactions". In: *Phys. Rev. D* 98.8 (2018), p. 083510. DOI: 10.1103/PhysRevD.98.083510. arXiv: 1801.08609 [astro-ph.CO].

[53] Weishuang Linda Xu, Cora Dvorkin, and Andrew Chael. "Probing sub-GeV Dark Matter-Baryon Scattering with Cosmological Observables". In: *Phys. Rev. D* 97.10 (2018), p. 103530. DOI: 10.1103/PhysRevD.97.103530. arXiv: 1802.06788 [astro-ph.CO].

[54] Chiu Man Ho and Robert J. Scherrer. "Limits on MeV Dark Matter from the Effective Number of Neutrinos". In: *Phys. Rev. D* 87.2 (2013), p. 023505. DOI: 10.1103/PhysRevD.87.023505. arXiv: 1208.4347 [astro-ph.CO].

[55] Celine Boehm, Matthew J. Dolan, and Christopher McCabe. "Increasing Neff with particles in thermal equilibrium with neutrinos". In: *JCAP* 12 (2012), p. 027. DOI: 10.1088/1475-7516/2012/12/027. arXiv: 1207.0497 [astro-ph.CO].

[56] Edward W. Kolb, Michael S. Turner, and Terrence P. Walker. "The Effect of Interacting Particles on Primordial Nucleosynthesis". In: *Phys. Rev. D* 34 (1986), p. 2197. DOI: 10.1103/PhysRevD.34.2197.

[57] Marco Hufnagel, Kai Schmidt-Hoberg, and Sebastian Wild. "BBN constraints on MeV-scale dark sectors. Part I. Sterile decays". In: *JCAP* 02

(2018), p. 044. DOI: 10.1088/1475-7516/2018/02/044. arXiv: 1712.03972 [hep-ph].

[58] Kenneth M. Nollett and Gary Steigman. "BBN And The CMB Constrain Neutrino Coupled Light WIMPs". In: *Phys. Rev. D* 91.8 (2015), p. 083505. DOI: 10.1103/PhysRevD.91.083505. arXiv: 1411.6005 [astro-ph.CO].

[59] D. Banerjee et al. "Search for Axionlike and Scalar Particles with the NA64 Experiment". In: *Phys. Rev. Lett.* 125.8 (2020), p. 081801. DOI: 10.1103/PhysRevLett.125.081801. arXiv: 2005.02710 [hep-ex].

[60] A.A. Aguilar-Arevalo et al. "Dark Matter Search in Nucleon, Pion, and Electron Channels from a Proton Beam Dump with MiniBooNE". In: *Phys. Rev. D* 98.11 (2018), p. 112004. DOI: 10.1103/PhysRevD.98.112004. arXiv: 1807.06137 [hep-ex].

[61] J. L. Friar and B. F. Gibson. "Coulomb energies in S-shell nuclei and hypernuclei". In: *Phys. Rev. C* 18 (2 Aug. 1978), pp. 908–913. DOI: 10.1103/PhysRevC.18.908. URL: https://link.aps.org/doi/10.1103/PhysRevC.18.908.

[62] Volker Koch and Gerald A. Miller. "Six quark cluster effects and binding energy differences between mirror nuclei". In: *Phys. Rev. C* 31 (2 Feb. 1985), pp. 602–612. DOI: 10.1103/PhysRevC.31.602. URL: https://link.aps.org/doi/10.1103/PhysRevC.31.602.

[63] G.A. Miller, B.M.K. Nefkens, and I. ?laus. "Charge symmetry, quarks and mesons". In: *Physics Reports* 194.1 (1990), pp. 1–116. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(90)90102-8. URL:

http://www.sciencedirect.com/science/article/pii/0370157390901028.

[64] T.E.O. Ericson and G.A. Miller. "Charge dependence of nuclear forces". In: *Physics Letters B* 132.1 (1983), pp. 32–38. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)90216-2. URL: http://www.sciencedirect.com/science/article/pii/0370269383902162.

[65] R Machleidt and I Slaus. "The nucleon-nucleon interaction". In: *Journal of Physics G: Nuclear and Particle Physics* 27.5 (Mar. 2001), R69–R108. DOI: 10.1088/0954-3899/27/5/201. URL: https://doi.org/10.1088%2F0954-3899%2F27%2F5%2F201.

[66] L. Koester, H. Rauch, and E. Seymann. "Neutron scattering lengths: A survey of experimental data and methods". In: *Atomic Data and Nuclear Data Tables* 49.1 (1991), pp. 65–120. ISSN: 0092-640X. DOI: https://doi.org/10.1016/0092-640X(91)90012-S. URL: http://www.sciencedirect.com/science/article/pii/0092640X9190012S.

[67] V.V. Nesvizhevsky, G. Pignol, and K.V. Protasov. "Neutron scattering and extra short range interactions". In: *Phys. Rev. D* 77 (2008), p. 034020. DOI: 10.1103/PhysRevD.77.034020. arXiv: 0711.2298 [hep-ph].

[68] R. Gilman et al. "Technical Design Report for the Paul Scherrer Institute Experiment R-12-01.1: Studying the Proton "Radius" Puzzle with μp Elastic Scattering". In: (Sept. 2017). arXiv: 1709.09753 [physics.ins-det].

[69] J. Grange et al. “Muon (g-2) Technical Design Report”. In: (Jan. 2015). arXiv: 1501.06858 [physics.ins-det].

[70] “The International Linear Collider Technical Design Report - Volume 2: Physics”. In: (June 2013). Ed. by Howard Baer et al. arXiv: 1306.6352 [hep-ph].

[71] Keisuke Fujii et al. “Physics Case for the 250 GeV Stage of the International Linear Collider”. In: (Oct. 2017). arXiv: 1710.07621 [hep-ex].

[72] “Physics and Detectors at CLIC: CLIC Conceptual Design Report”. In: (Feb. 2012). Ed. by Lucie Linssen et al. DOI: 10.5170/CERN-2012-003. arXiv: 1202.5940 [physics.ins-det].

[73] T. K. Charles et al. “The Compact Linear Collider (CLIC) - 2018 Summary Report”. In: 2/2018 (Dec. 2018). Ed. by P. N. Burrows et al. DOI: 10.23731/CYRM-2018-002. arXiv: 1812.06018 [physics.acc-ph].

[74] Mingyi Dong et al. “CEPC Conceptual Design Report: Volume 2 - Physics & Detector”. In: (Nov. 2018). Ed. by João Barreiro Guimarães da Costa et al. arXiv: 1811.10545 [hep-ex].

[75] A. Abada et al. “FCC Physics Opportunities: Future Circular Collider Conceptual Design Report Volume 1”. In: *Eur. Phys. J. C* 79.6 (2019), p. 474. DOI: 10.1140/epjc/s10052-019-6904-3.

[76] A. Abada et al. “FCC-ee: The Lepton Collider: Future Circular Collider Conceptual Design Report Volume 2”. In: *Eur. Phys. J. ST* 228.2 (2019), pp. 261–623. DOI: 10.1140/epjst/e2019-900045-4.

[77] I.Yu. Kobzarev, L.B. Okun, and I.Ya. Pomeranchuk. “On the Possibility of Experimental Observation of Mirror Particles”. In: *Sov. J. Nucl. Phys.* 3.6 (1966), pp. 837–841.

[78] L.B. Okun. "Limits of Electrodynamics: Paraphotons?" In: *Sov. Phys. JETP* 56 (1982), p. 502.

[79] Bob Holdom. "Two U(1)'s and Epsilon Charge Shifts". In: *Phys. Lett. B* 166 (1986), pp. 196–198. DOI: 10.1016/0370-2693(86)91377-8.

[80] Bob Holdom. "Searching for ϵ Charges and a New U(1)". In: *Phys. Lett. B* 178 (1986), pp. 65–70. DOI: 10.1016/0370-2693(86)90470-3.

[81] Robert Foot and Xiao-Gang He. "Comment on Z Z-Prime Mixing in Extended Gauge Theories". In: *Phys. Lett. B* 267 (1991), pp. 509–512. DOI: 10.1016/0370-2693(91)90901-2.

[82] Hooman Davoudiasl, Hye-Sung Lee, and William J. Marciano. "'Dark' Z Implications for Parity Violation, Rare Meson Decays, and Higgs Physics". In: *Phys. Rev. D* 85 (2012), p. 115019. DOI: 10.1103/PhysRevD.85.115019. arXiv: 1203.2947 [hep-ph].

[83] Marco Fabbrichesi, Emidio Gabrielli, and Gaia Lanfranchi. "The Dark Photon". In: (May 2020). arXiv: 2005.01515 [hep-ph].

[84] Jonathan L. Feng, Arvind Rajaraman, and Fumihiro Takayama. "Superweakly interacting massive particles". In: *Phys. Rev. Lett.* 91 (2003), p. 011302. DOI: 10.1103/PhysRevLett.91.011302. arXiv: hep-ph/0302215.

[85] David Curtin et al. "Illuminating Dark Photons with High-Energy Colliders". In: *JHEP* 02 (2015), p. 157. DOI: 10.1007/JHEP02(2015)157. arXiv: 1412.0018 [hep-ph].

[86] Marek Karliner et al. "Radiative return capabilities of a high-energy, high-luminosity e^+e^- collider". In: *Phys. Rev. D* 92.3 (2015), p. 035010.

DOI: 10.1103/PhysRevD.92.035010. arXiv: 1503.07209 [hep-ph].

[87] Shrihari Gopalakrishna, Sunghoon Jung, and James D. Wells. "Higgs Boson Decays to Four Fermions Through an Abelian Hidden Sector". In: *Phys. Rev. D* 78 (2008), p. 055002. DOI: 10.1103/PhysRevD.78.055002. arXiv: 0801.3456 [hep-ph].

[88] Jia Liu, Xiao-Ping Wang, and Felix Yu. "A Tale of Two Portals: Testing Light, Hidden New Physics at Future e^+e^- Colliders". In: *JHEP* 06 (2017), p. 077. DOI: 10.1007/JHEP06(2017)077. arXiv: 1704.00730 [hep-ph].

[89] Min He et al. "Search for a heavy dark photon at future e^+e^- colliders". In: *JHEP* 03 (2018), p. 139. DOI: 10.1007/JHEP03(2018)139. arXiv: 1712.09095 [hep-ph].

[90] John M. Cornwall, David N. Levin, and George Tiktopoulos. "Derivation of Gauge Invariance from High-Energy Unitarity Bounds on the s Matrix". In: *Phys. Rev. D* 10 (1974). [Erratum: Phys.Rev.D 11, 972 (1975)], p. 1145. DOI: 10.1103/PhysRevD.10.1145.

[91] Michael S. Chanowitz and Mary K. Gaillard. "The TeV Physics of Strongly Interacting W's and Z's". In: *Nucl. Phys. B* 261 (1985), pp. 379–431. DOI: 10.1016/0550-3213(85)90580-2.

[92] Albert M Sirunyan et al. "Search for a Narrow Resonance Lighter than 200 GeV Decaying to a Pair of Muons in Proton-Proton Collisions at $\sqrt{s} =$ TeV". In: *Phys. Rev. Lett.* 124.13 (2020), p. 131802. DOI: 10.1103/PhysRevLett.124.131802. arXiv: 1912.04776 [hep-ex].

[93] Albert M Sirunyan et al. "Search for resonant and nonresonant new phenomena in high-mass dilepton final states at $\sqrt{s}$ = 13 TeV". In: *JHEP* 07 (2021), p. 208. DOI: 10.1007/JHEP07(2021)208. arXiv: 2103.02708 [hep-ex].

[94] Georges Aad et al. "Search for high-mass dilepton resonances using 139 fb^{-1} of pp collision data collected at $\sqrt{s}$ =13 TeV with the ATLAS detector". In: *Phys. Lett. B* 796 (2019), pp. 68–87. DOI: 10.1016/j.physletb.2019.07.016. arXiv: 1903.06248 [hep-ex].

[95] Graham D. Kribs, David McKeen, and Nirmal Raj. "Breaking up the Proton: An Affair with Dark Forces". In: *Phys. Rev. Lett.* 126.1 (2021), p. 011801. DOI: 10.1103/PhysRevLett.126.011801. arXiv: 2007.15655 [hep-ph].

[96] Bin Yan. "Probing the dark photon via polarized DIS scattering at the HERA and EIC". In: (Mar. 2022). arXiv: 2203.01510 [hep-ph].

[97] Shinya Kanemura, Takeo Moroi, and Tomohiko Tanabe. "Beam dump experiment at future electron–positron colliders". In: *Phys. Lett. B* 751 (2015), pp. 25–28. DOI: 10.1016/j.physletb.2015.10.002. arXiv: 1507.02809 [hep-ph].

[98] Kento Asai et al. "New physics searches at the ILC positron and electron beam dumps". In: *JHEP* 09 (2021), p. 183. DOI: 10.1007/JHEP09(2021)183. arXiv: 2105.13768 [hep-ph].

[99] P.A. Zyla et al. "Review of Particle Physics". In: *PTEP* 2020.8 (2020). See Section 10, *Electroweak Model and Constraints on New Physics*, p. 083C01. DOI: 10.1093/ptep/ptaa104.

[100] Anson Hook, Eder Izaguirre, and Jay G. Wacker. "Model Independent Bounds on Kinetic Mixing". In: *Adv. High Energy Phys.* 2011 (2011), p. 859762. DOI: 10.1155/2011/859762. arXiv: 1006.0973 [hep-ph].

[101] D. B. Clark, E. Godat, and F. I. Olness. "ManeParse : A Mathematica reader for Parton Distribution Functions". In: *Comput. Phys. Commun.* 216 (2017), pp. 126–137. DOI: 10.1016/j.cpc.2017.03.004. arXiv: 1605.08012 [hep-ph].

[102] K. Kovarik et al. "nCTEQ15 - Global analysis of nuclear parton distributions with uncertainties in the CTEQ framework". In: *Phys. Rev. D* 93.8 (2016), p. 085037. DOI: 10.1103/PhysRevD.93.085037. arXiv: 1509.00792 [hep-ph].

[103] C. F. von Weizsacker. "Radiation emitted in collisions of very fast electrons". In: *Z. Phys.* 88 (1934), pp. 612–625. DOI: 10.1007/BF01333110.

[104] E. J. Williams. "Correlation of certain collision problems with radiation theory". In: *Kong. Dan. Vid. Sel. Mat. Fys. Med.* 13N4.4 (1935), pp. 1–50.

[105] Min-Shih Chen and Peter M. Zerwas. "Equivalent-Particle Approximations in electron and Photon Processes of Higher Order QED". In: *Phys. Rev. D* 12 (1975), p. 187. DOI: 10.1103/PhysRevD.12.187.

[106] Maxim Perelstein. "Introduction to Collider Physics". In: *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics of the Large and the Small*. 2011, pp. 421–486. DOI: 10.1142/9789814327183_0008. arXiv: 1002.0274 [hep-ph].

[107] Halina Abramowicz et al. “The International Linear Collider Technical Design Report - Volume 4: Detectors”. In: (June 2013). Ed. by Ties Behnke et al. arXiv: 1306.6329 [physics.ins-det].

[108] Carl Mikael Berggren. “New physics searches with the ILD detector at the ILC”. In: *PoS* PANIC2021 (2022), p. 123. DOI: 10.22323/1.380.0123. arXiv: 2111.09928 [hep-ex].

[109] Frank Simon. “Perspectives for Top Quark Physics at the (I)LC”. In: *7th International Workshop on Top Quark Physics*. Nov. 2014. arXiv: 1411.7517 [hep-ex].

[110] N. Muchnoi, H. J. Schreiber, and M. Viti. “ILC Beam Energy Measurement by means of Laser Compton Backscattering”. In: *Nucl. Instrum. Meth. A* 607 (2009), pp. 340–366. DOI: 10.1016/j.nima.2009.05.145. arXiv: 0812.0925 [physics.ins-det].

[111] I. Božović Jelisavčić et al. “Luminosity measurement at ILC”. In: *JINST* 8 (2013), P08012. DOI: 10.1088/1748-0221/8/08/P08012. arXiv: 1304.4082 [physics.acc-ph].

[112] S. Boogert et al. “Polarimeters and Energy Spectrometers for the ILC Beam Delivery System”. In: *JINST* 4 (2009), P10015. DOI: 10.1088/1748-0221/4/10/P10015. arXiv: 0904.0122 [physics.ins-det].

[113] Gregor Kasieczka et al. “The LHC Olympics 2020 a community challenge for anomaly detection in high energy physics”. In: *Rept. Prog. Phys.* 84.12 (2021), p. 124201. DOI: 10.1088/1361-6633/ac36b9. arXiv: 2101.08320 [hep-ph].

[114] Thea Aarrestad et al. “The Dark Machines Anomaly Score Challenge: Benchmark Data and Model Independent Event Classification for the Large Hadron Collider”. In: *SciPost Phys.* 12.1 (2022), p. 043. DOI: 10.21468/SciPostPhys.12.1.043. arXiv: 2105.14027 [hep-ph].

[115] Jack H. Collins, Kiel Howe, and Benjamin Nachman. “Anomaly Detection for Resonant New Physics with Machine Learning”. In: *Phys. Rev. Lett.* 121.24 (2018), p. 241803. DOI: 10.1103/PhysRevLett.121.241803. arXiv: 1805.02664 [hep-ph].

[116] Theo Heimel et al. “QCD or What?” In: *SciPost Phys.* 6.3 (2019), p. 030. DOI: 10.21468/SciPostPhys.6.3.030. arXiv: 1808.08979 [hep-ph].

[117] Marco Farina, Yuichiro Nakai, and David Shih. “Searching for New Physics with Deep Autoencoders”. In: *Phys. Rev. D* 101.7 (2020), p. 075021. DOI: 10.1103/PhysRevD.101.075021. arXiv: 1808.08992 [hep-ph].

[118] Jack H. Collins, Kiel Howe, and Benjamin Nachman. “Extending the search for new resonances with machine learning”. In: *Phys. Rev. D* 99.1 (2019), p. 014038. DOI: 10.1103/PhysRevD.99.014038. arXiv: 1902.02634 [hep-ph].

[119] Benjamin Nachman and David Shih. “Anomaly Detection with Density Estimation”. In: *Phys. Rev. D* 101 (2020), p. 075042. DOI: 10.1103/PhysRevD.101.075042. arXiv: 2001.04990 [hep-ph].

[120] Anders Andreassen, Benjamin Nachman, and David Shih. “Simulation Assisted Likelihood-free Anomaly Detection”. In: *Phys. Rev. D* 101.9

(2020), p. 095004. DOI: 10.1103/PhysRevD.101.095004. arXiv: 2001.05001 [hep-ph].

[121] Kees Benkendorfer, Luc Le Pottier, and Benjamin Nachman. "Simulation-assisted decorrelation for resonant anomaly detection". In: *Phys. Rev. D* 104.3 (2021), p. 035003. DOI: 10.1103/PhysRevD.104.035003. arXiv: 2009.02205 [hep-ph].

[122] Anna Hallin et al. "Classifying anomalies through outer density estimation". In: *Phys. Rev. D* 106.5 (2022), p. 055006. DOI: 10.1103/PhysRevD.106.055006. arXiv: 2109.00546 [hep-ph].

[123] John Andrew Raine et al. "CURTAINs for your sliding window: Constructing unobserved regions by transforming adjacent intervals". In: *Front. Big Data* 6 (2023), p. 899345. DOI: 10.3389/fdata.2023.899345. arXiv: 2203.09470 [hep-ph].

[124] Anna Hallin et al. "Resonant anomaly detection without background sculpting". In: *Phys. Rev. D* 107.11 (2023), p. 114012. DOI: 10.1103/PhysRevD.107.114012. arXiv: 2210.14924 [hep-ph].

[125] Tobias Golling et al. "Flow-enhanced transportation for anomaly detection". In: *Phys. Rev. D* 107.9 (2023), p. 096025. DOI: 10.1103/PhysRevD.107.096025. arXiv: 2212.11285 [hep-ph].

[126] Eric M. Metodiev, Benjamin Nachman, and Jesse Thaler. "Classification without labels: Learning from mixed samples in high energy physics". In: *JHEP* 10 (2017), p. 174. DOI: 10.1007/JHEP10(2017)174. arXiv: 1708.02949 [hep-ph].

[127] Thorben Finke et al. “Back to the Roots: Tree-Based Algorithms for Weakly Supervised Anomaly Detection”. In: (Sept. 2023). arXiv: 2309.13111 [hep-ph].

[128] Leo Grinsztajn, Edouard Oyallon, and Gael Varoquaux. “Why do tree-based models still outperform deep learning on typical tabular data?” In: *Advances in Neural Information Processing Systems*. Ed. by S. Koyejo et al. Vol. 35. Curran Associates, Inc., 2022, pp. 507–520. URL: https://proceedings.neurips.cc/paper_files/paper/2022/file/0378c7692da36807bdec87ab043cdadc-Paper-Datasets_and_Benchmarks.pdf.

[129] Vadim Borisov et al. “Deep Neural Networks and Tabular Data: A Survey”. In: *CoRR* abs/2110.01889 (2021). arXiv: 2110.01889. URL: https://arxiv.org/abs/2110.01889.

[130] Gregor Kasieczka, Ben Nachman, and David Shih. *R&D Dataset for LHC Olympics 2020 Anomaly Detection Challenge*. Version v5. Apr. 2019. DOI: 10.5281/zenodo.6466204. URL: https://doi.org/10.5281/zenodo.6466204.

[131] Christian Bierlich et al. “A comprehensive guide to the physics and usage of PYTHIA 8.3”. In: (Mar. 2022). DOI: 10.21468/SciPostPhysCodeb.8. arXiv: 2203.11601 [hep-ph].

[132] J. de Favereau et al. “DELPHES 3, A modular framework for fast simulation of a generic collider experiment”. In: *JHEP* 02 (2014), p. 057. DOI: 10.1007/JHEP02(2014)057. arXiv: 1307.6346 [hep-ex].

[133] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet User Manual". In: *Eur. Phys. J. C* 72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[134] Jesse Thaler and Ken Van Tilburg. "Identifying Boosted Objects with N-subjettiness". In: *JHEP* 03 (2011), p. 015. DOI: 10.1007/JHEP03(2011)015. arXiv: 1011.2268 [hep-ph].

[135] Jesse Thaler and Ken Van Tilburg. "Maximizing Boosted Top Identification by Minimizing N-subjettiness". In: *JHEP* 02 (2012), p. 093. DOI: 10.1007/JHEP02(2012)093. arXiv: 1108.2701 [hep-ph].

[136] George John, Ron Kohavi, and Karl Pfleger. "Irrelevant Features and the Subset Selection Problem". In: *In International Conference on Machine Learning* 129 (Feb. 1998). DOI: 10.1016/B978-1-55860-335-6.50023-4.

[137] Jerzy Neyman and Egon Sharpe Pearson. "On the Problem of the Most Efficient Tests of Statistical Hypotheses". In: *Phil. Trans. Roy. Soc. Lond. A* 231.694-706 (1933), pp. 289–337. DOI: 10.1098/rsta.1933.0009.

[138] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer Series in Statistics. New York, NY, USA: Springer New York Inc., 2001.

[139] Tianqi Chen and Carlos Guestrin. "XGBoost". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, Aug. 2016. DOI: 10.1145/2939672.2939785. URL: https://doi.org/10.1145%2F2939672.2939785.

[140] Diederik Kingma and Jimmy Ba. "Adam: A Method for Stochastic Optimization". In: *International Conference on Learning Representations (ICLR)*. San Diega, CA, USA, 2015. arXiv: 1412.6980 [cs.LG].

[141] Naoki Awaya and Li Ma. *Unsupervised tree boosting for learning probability distributions*. 2023. arXiv: 2101.11083 [stat.ME].

[142] George Papamakarios et al. "Normalizing Flows for Probabilistic Modeling and Inference". In: *Journal of Machine Learning Research* 22.57 (2021), pp. 1–64. URL: http://jmlr.org/papers/v22/19-1028.html.

[143] Abe Sklar. "Fonctions de répartition à n dimensions et leurs marges". In: *Publications de l'Institut de Statistique de l'Université de Paris* 8 (1959), pp. 229–231.

[144] Debajyoti Sengupta et al. "CURTAINs Flows for Flows: Constructing Unobserved Regions with Maximum Likelihood Estimation". In: (May 2023). arXiv: 2305.04646 [hep-ph].

[145] D. C. Liu and J. Nocedal. "On the limited memory BFGS method for large scale optimization". In: *Math. Programming* 45.3, (Ser. B) (1989), pp. 503–528.

[146] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[147] Georges Aad et al. "Dijet resonance search with weak supervision using $\sqrt{s} = 13$ TeV pp collisions in the ATLAS detector". In: *Phys. Rev. Lett.* 125.13 (2020), p. 131801. DOI: 10.1103/PhysRevLett.125.131801. arXiv: 2005.02983 [hep-ex].

[148] Ian Chaffey and Philip Tanedo. "Vector self-interacting dark matter". In: *Phys. Rev. D* 101.7 (2020), p. 075005. DOI: 10.1103/PhysRevD.101.075005. arXiv: 1907.10217 [hep-ph].

[149] John F. Gunion and Howard E. Haber. "The CP conserving two Higgs doublet model: The Approach to the decoupling limit". In: *Phys. Rev. D* 67 (2003), p. 075019. DOI: 10.1103/PhysRevD.67.075019. arXiv: hep-ph/0207010.

[150] E. C. G. Stueckelberg. "Interaction Energy in Electrodynamics and in the Field Theory of Nuclear Forces". In: *Helv. Phys. Acta* 11 (1938), pp. 225–244. DOI: 10.5169/seals-110852.

[151] Daniel Feldman, Zuowei Liu, and Pran Nath. "The Stueckelberg Z-Prime Extension with Kinetic Mixing and Milli-Charged Dark Matter from the Hidden Sector". In: *Phys. Rev. D* 75 (2007), p. 115001. DOI: 10.1103/PhysRevD.75.115001. arXiv: hep-ph/0702123.

[152] G. Moortgat-Pick et al. "The Role of polarized positrons and electrons in revealing fundamental interactions at the linear collider". In: *Phys. Rept.* 460 (2008), pp. 131–243. DOI: 10.1016/j.physrep.2007.12.003. arXiv: hep-ph/0507011.