UC Irvine
UC Irvine Previously Published Works

Title
Learning to reconstruct quirky tracks

Permalink
https://escholarship.org/uc/item/1w36d3rz

Journal
Physical Review D, 114(1)

ISSN
2470-0010

Authors
Sha, Qiyu
Murnane, Daniel
Fieg, Max
et al.

Publication Date
2026-07-01

DOI
10.1103/dvzc-lgj1

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Learning to reconstruct quirky tracks

Qiyu Sha,[1,2] Daniel Murnane,[3,4] Max Fieg,[5] Shelley Tong,[5] Mark Zakharyan,[6] Yaquan Fang,[1,2] and Daniel Whiteson[5]

[1]*Institute of High Energy Physics, 19B Yuquan Road, Shijingshan District, Beijing 100049, China*
[2]*University of Chinese Academy of Sciences, 19A Yuquan Road, Shijingshan District, Beijing 100049, China*
[3]*Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark*
[4]*Scientific Data Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA*
[5]*Department of Physics and Astronomy, University of California, Irvine, California, USA*
[6]*Department of Physics, University of California, Santa Barbara, California, USA*

(Received 19 October 2025; accepted 9 June 2026; published 6 July 2026)

Analysis of data from particle physics experiments traditionally sacrifices some sensitivity to new particles for the sake of practical computability, effectively ignoring some potentially striking signatures. However, recent advances in machine learning (ML)–based tracking allow for new inroads into previously inaccessible territory, such as reconstruction of tracks that do not follow helical trajectories. This paper presents a demonstration of the capacity of ML-based tracking to reconstruct the oscillating trajectories of quirks, particles charged under an unbroken QCD-like gauge symmetry. The technique used is not specific to quirks, and opens the door to a program of searching for many kinds of nonstandard tracks.

DOI: 10.1103/dvzc-lgj1

I. INTRODUCTION

Hadronic collisions at the energy frontier have wide power to reveal or constrain new physics [1–8], such as new particles or interactions, whether expected or unexpected. In the coming years, the Large Hadron Collider (LHC) will accumulate a dataset of unprecedented luminosity, with an expected $\mathcal{L} = 300$ fb^{-1} (3 ab^{-1}) with the completion of Run 3 (Run 4+). While the current physics program is broad, the technical challenges of analyzing the high-dimensional, high-volume dataset have historically required imposing assumptions on the nature of the potential new physics, which translate to strategic sacrifices in sensitivity, especially to the unexpected. As a result, signatures of ground-breaking discoveries may be overlooked by the traditional analysis pipelines.

A crucial and challenging component of that pipeline is identifying tracks of charged particles in the inner, magnetized detector volume. Ionization leaves *hits* as particles pass through a tracking detector, but the task of assembling sets of hits to reconstruct *tracks* of particles is a challenging combinatorial problem. The task is made tractable, though still computationally costly, if one can assume that all tracks follow a helical trajectory and originate from within the beam pipe, as is expected of electrically charged particles created in an interaction at the beam crossing. This narrowed scope is sufficient for many expected signatures, but consequentially restricts the power to discover unexpected new physics. This is well understood, and significant effort has been made to relax the assumption that tracks must originate inside the beam pipe by developing dedicated software to reconstruct tracks with displaced vertices (recent results include Refs. [9–14]), as might be expected from decays of long-lived particles. Comparatively little effort [15–18] has been put toward development of algorithms capable of identifying a potentially much broader class of tracks, those which are charged under other forces and do not follow helical paths in a solenoidal magnetic field. Such striking tracks may already exist in our data, but would be largely invisible to our current algorithms.

An important challenge in reconstructing such nonhelical tracks is that traditional tracking algorithms [19–22] assume the specific parametric form of the desired track in a way that does not trivially generalize to new forms. Each new potential trajectory would therefore require its own bespoke tracking algorithm and incur its own computational expense in the analysis pipeline. However, recent innovations in machine learning (ML), such as geometric deep learning [23], have allowed for the development of ML-based tracking algorithms which are architecturally agnostic to the parametric form of the track [24], where the target class of the track is defined only by the sample used to train the algorithm. In principle,

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

this allows for rapid development of tracking pipelines which are more flexible, replacing the complex work of dedicated algorithm development with the much simpler task of preparation of a new training sample. If successful, it lays the groundwork for a more ambitious task, preparation of a generalized training sample that produces a track finder capable of reconstructing a broad class of helical and nonhelical trajectories, with true sensitivity to unexpected tracks.

In this paper, we take the first step in that direction by demonstrating that ML-based tracking algorithms can be trained to identify nonhelical tracks from a training sample, without the development of a dedicated new tracking algorithm. We train the graph neural network (GNN) machine learning algorithm Exa.TrkX [25] to learn to reconstruct the strikingly nonhelical oscillatory tracks produced by *quirk pairs*. Quirks [26] are charged under an unbroken confining QCD-like gauge symmetry but, in contrast with quarks, the quirk masses are large relative to the confinement scale, resulting in an oscillatory nonhelical motion by quirk pairs in a magnetic field. These spectacular signatures are such a departure from the classical helical tracks, given the in-flight interactions between the quirks, that even a bespoke application of the traditional algorithms has been a long-standing challenge. Previous efforts have attempted to side-step this complexity by searching for straight tracks [27], momentum imbalance [28], stopped out-of-time particles [29], or sets of hits along a plane [18], essentially ignoring the oscillatory nature of the tracks. This is reasonably effective for a limited quirk parameter range, but we show that ML-based tracking is more broadly efficient without requiring such quirk-specific simplifications. Our studies also demonstrate a capacity to generalize beyond examples in the training set. This gives hope that a similar approach could be used to construct a pipeline that can reconstruct tracks from a broader class of particles, unrestricted to helical tracks or quirks. This may pave the way for discoveries at the LHC that have evaded our detection. Complementary studies show sensitivity to quirks in the far-forward region with experiments such as FASER [30–33].

The paper is organized as follows. In Sec. II we cover the background of our study by reviewing the quirk model in a collider setting. In Sec. III, we discuss the equations of motion for the quirk pair and for the background particles. In Sec. IV, we discuss the Exa.TrkX pipeline and how it can be applied to quirk track reconstruction. In Sec. V we present our track reconstruction performance. In Sec. VI we discuss the sensitivity of our reconstruction techniques to quirks. In Sec. VII we conclude and discuss future directions.

II. THEORETICAL BACKGROUND

In the standard model (SM), quarks are charged under the unbroken $\mathrm{SU}(3)_{\mathrm{C}}$ which confines below the QCD scale, Λ_{QCD}. Notably, there exists light quarks with a mass m_q such that $m_q \ll \Lambda_{\mathrm{QCD}}$. This leads to color confinement, as quarks separated by a sufficient distance convert the energy of their interaction into quark pairs.

In contrast, if the masses were above the confinement scale, macroscopic distances could be maintained. This is the case for *quirks* [26], charged under a distinct non-Abelian, confining gauge theory $\mathrm{SU}(N)_{\mathrm{IC}}$, sometimes called *infracolor*, of which all other standard model particles are singlets. This new gauge interaction produces a confinement scale Λ_{IC}. The key difference between quirks and quarks is that for a quirk, Q, the mass $m_Q \gg \Lambda_{\mathrm{IC}}$, where for convenience, we write $\Lambda \equiv \Lambda_{\mathrm{IC}}$ for the remainder of this paper. Quirks are also typically charged under the standard model, such that they can be pair produced at colliders.

The unbroken $\mathrm{SU}(N)_{\mathrm{IC}}$ leads to an infracolor gauge string between the pair-produced quirks, as with quarks. Unlike quarks, however, as the quirks separate the breaking of the string is exponentially suppressed due to $m_Q \gg \Lambda$. The absence of string breaking results in oscillation of the pair with a characteristic length [18]

$$d_{cm} \approx 2\ \mathrm{cm}(\gamma - 1)\left(\frac{m_Q}{100\ \mathrm{GeV}}\right)\left(\frac{\mathrm{keV}}{\Lambda}\right)^2 \qquad (1)$$

in the center-of-mass frame, where γ is the Lorentz boost of the quirks at production. When reporting the quirk track reconstruction performance we will use this formula for qualitative understanding. It is important to note that since we study pair-produced quirks in association with a jet, actual oscillation lengths in the lab frame can be very different from Eq. (1) due to smaller-than-typical opening angles or different energies of each quirk.

In a collider setting, the oscillation length of produced quirks determines the potential reconstruction scenario. For $d_{cm} \lesssim 1$ mm, oscillation cannot be resolved by the detector, and the quirk pair is reconstructed as a single track. For oscillation scales much larger than the detector, $d_{cm} \gtrsim 1$ m, the quirk pair does not complete an oscillation within the detector, and so the quirk pair appears as two standard tracks. In both cases, the tracks can be described using helical parametrization, traditional tracking algorithms are effective, and existing searches are relevant [18,27–29].

However, when the oscillation length is on the order of the size of the detector, the quirk pair produces two tracks with a unique, nonhelical motion. See Fig. 1 for visualizations of example trajectories, whose features are explained in greater detail in the following section. The stark deviations from a helix make traditional tracking algorithms ineffective for finding such complex tracks.

III. MODELING

Samples of simulated events are prepared and used for training of the tracking algorithm as well as evaluation of the performance. We study quirks charged under the

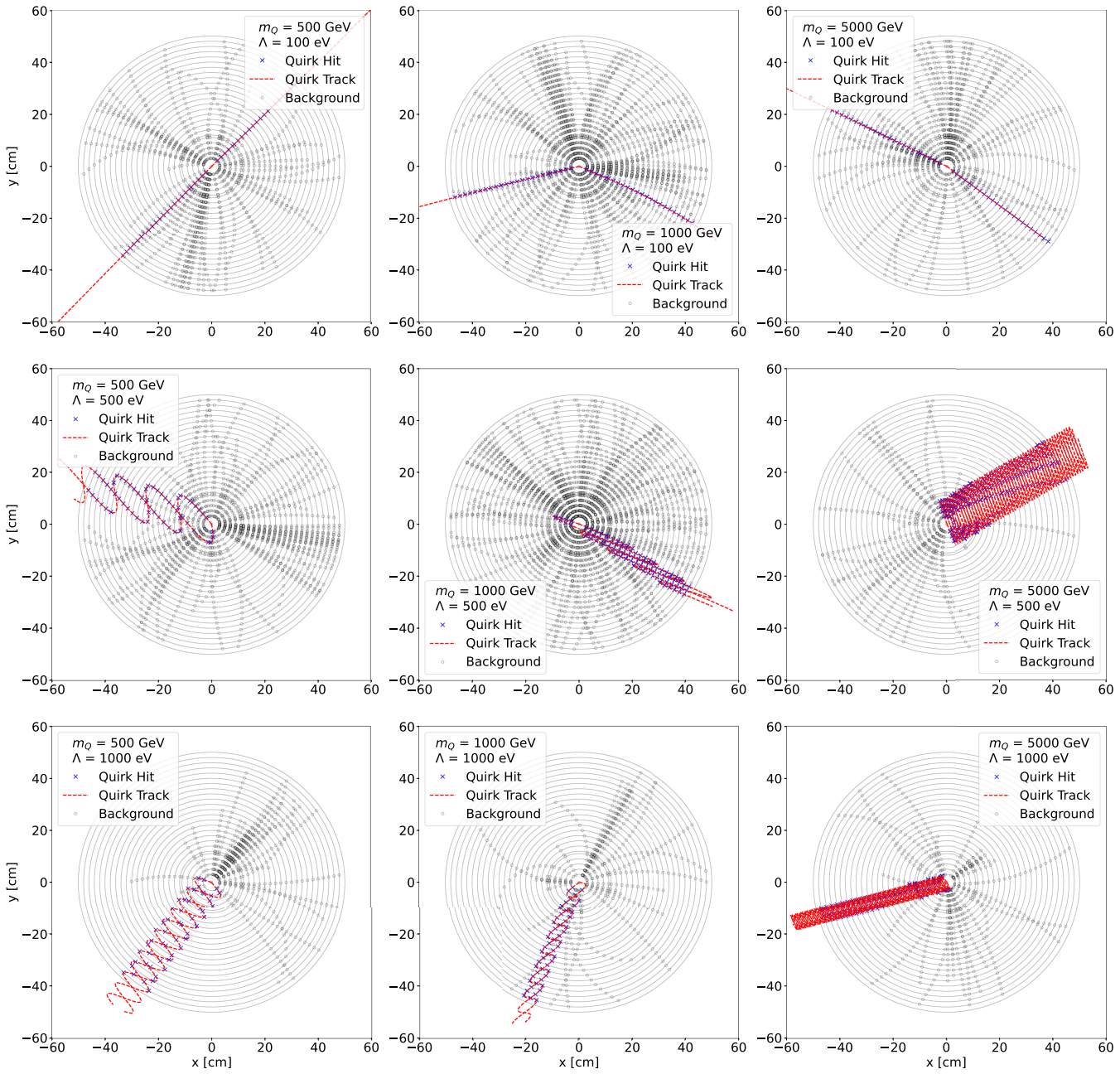


FIG. 1. Example trajectories of quirk pairs in simulated events with a simplified detector geometry described in the text. The top row shows example trajectories for $\Lambda = 100$ eV, and increases in mass from left to right, with $m_Q = 500$, 1000, 5000 GeV, all with small oscillation distances. The center (bottom) row shows examples for the same mass range but for $\Lambda = 500(1000)$ eV, both with visible oscillations. Hits from SM background tracks (see Sec. V D for details) are in gray, and from quirks are in blue. The true quirk trajectory is also shown, in red.

standard model $\mathrm{U}(1)_{\mathrm{EM}}$ (with charge ± 1) such that they can be pair produced through, e.g., an s-channel Drell-Yan process. Samples of simulated quirk signal and SM background are generated in pp collisions at $\sqrt{s} = 13$ TeV with MadGraph5_aMC [34]. Nonquirk particles are showered and hadronized with Pythia8 [35] and a selection of $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$ is applied on all charged particles, including quirks. The trajectories of particles and their interaction with the detector are then modeled with a custom package described below.

The leading-order Drell-Yan production mechanism results in a quirk pair with zero transverse momentum, but each quirk can have nonzero transverse momentum. For certain regions of the (m_Q, Λ) parameter space with oscillation lengths $\sim$1 cm, the resultant motion is that the quirks travel in the beam direction [31] and only leave hits

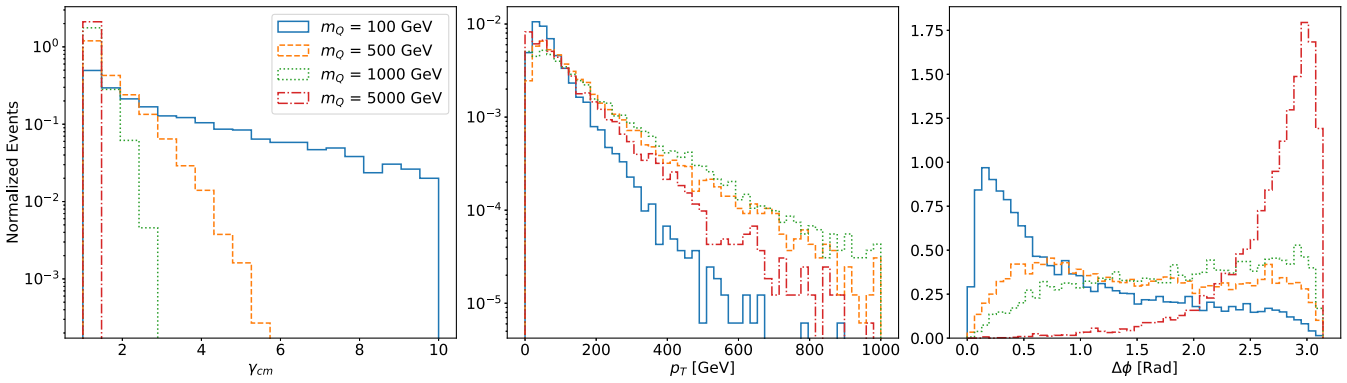


FIG. 2. Distributions of kinematic quantities for simulated quirk pairs, for several mass values, independent of Λ. Shown are the γ factor, transverse momentum p_{T}, and opening angle $\Delta\phi$.

in the innermost layers of the detector. To generate visible tracks, we produce the quirk pair in association with a jet, giving the quirk pair a net transverse momentum. We generate samples of simulated quirks in the mass range $m_Q \in [100, 5000]$ GeV. Figure 2 shows the distributions of the simulated γ factor of the quirk pair, γ_{cm}, quirk pair transverse momentum, and opening angles in this range. Here, the opening angle refers to the initial angle between the quirks within the plane of their production. These quantities broadly describe the oscillation calculated for quirk trajectories. In general, as the opening angle, $\Delta\phi$, approaches π, and for quirks produced with large momenta, the resultant quirk motion will have the largest separation. An additional feature of the initial quirk production which has important consequences for the quirk trajectory is the case where $p_{Q1} \gg p_{Q2}$ (or vice versa). In this case Q_1 will drag Q_2 in the direction of $\vec{p}_1$, which will result in a complicated motion and exchange of momenta between each quirk.

In pp collisions at $\sqrt{s} = 13$ TeV, the production cross section for two quirks and a jet $\sigma_{QQj} \approx 1$ fb for $m_Q = 500$ GeV, which is sufficient to expect a significant number of quirks in the current LHC dataset. For the purposes of this study we will explore quirk masses up to 5 TeV—for the larger masses in this range, the cross section is too small for appreciable production at the LHC with current planned luminosities. Nevertheless, we explore these heavier cases to assess the strength of our reconstruction technique which may be useful at future colliders or for models with higher production cross sections. Larger cross sections can be achieved for quirks with lower p_{T}, with associated reconstruction challenges, or quirks charged under $\mathrm{SU}(3)_{\mathrm{C}}$.

Instead of following a helical path, quirk pairs follow the motion of two objects subjected to a central force. Given that their angular momentum is approximately conserved, their trajectories are largely confined to a single plane [18]. Solutions to the equations of motion assume an extra central force between the two quirks in addition to the Lorentz force that arises from the external electric and magnetic fields. Following Ref. [18], the system is modeled using a Nambu-Goto action with massive end points [36], leading to the equations of motion for a quirk given by the force,

$$\frac{d(m\gamma\vec{v})}{dt} = -\Lambda^2\left(\frac{\sqrt{1-v_\perp^2}}{v_\parallel}\vec{v}_\parallel + \frac{v_\parallel}{\sqrt{1-v_\perp^2}}\vec{v}_\perp\right) + \vec{F}_{\mathrm{ext}}, \qquad (2)$$

with a similar equation for the second quirk. Here, m is the mass of the quirk, $\vec{v}_\perp$ and $\vec{v}_\parallel$ are the perpendicular and parallel velocities of the quirk with respect to the infracolor string, and F_{ext} is the external force acting on the quirks, which for us will only include the affect of the external magnetic field on the (anti-)quirk with charge $q = \pm 1$.

To assess the track reconstruction performance, we assume two simplified detector configurations for an LHC-style detector. The simplest is an eight-layer [18] configuration with four tracking layers in the inner detector between $r = 3.1$ cm and $r = 12.25$ cm covering $|z| < 40$ cm and four layers in the outer detector, between $r = 29.9$ cm and $r = 100$ cm, covering $|z| < 75$ cm. In a second configuration, we consider a configuration where there are 25 tracking layers between $r = 3.1$ cm and $r = 53$ cm with the same z-range as the eight-layer detector. We calculate particle trajectories and their intersection with a detector layer is recorded as a hit. No noise hits are included in these first studies, though hit location uncertainties are assessed in Sec. V H.

Figure 1 shows selected examples of quirk trajectories for $\Lambda = 100$, 500, 1000 eV (top row) and $m_Q = 500$ GeV, 1 TeV, and 5 TeV. The oscillation length depends strongly on the mass and Λ, as expected from Eq. (1). As quirks become heavier, they naturally receive less kinetic energy which results in smaller oscillation lengths. For heavy quirks with large Λ, the oscillations can be quite rapid. Both very long and very short oscillation lengths are similar

enough to traditional helical tracks that standard algorithms can have reasonable efficiency. The focus of this work is on the intermediate regime, where the quirk separation is approximately a few centimeters, making strikingly nonhelical tracks.

For the production of background tracks arising from the associated jet production, we collect all stable charged particles on the detector length scale ($e^{\pm}, \mu^{\pm}, \pi^{\pm}, K^{\pm}$) with $p_T > 100$ MeV and $|\eta| < 2.5$ from Pythia. We model their propagation through the detector as charged particles in a magnetic field, neglecting matter interactions. We find that background tracks contribute many more hits as compared to quirks. For $m_Q \lesssim$ TeV we find that background hits outnumber quirk hits by a factor of $\mathcal{O}(100)$, with lighter quirks being more outnumbered by background than heavier quirks with the same Λ. For the heaviest quirks we consider, $m_Q = 5$ TeV, which is nearly at the kinematic threshold, there is significantly less energy available for background particles resulting in comparatively fewer background tracks, although background still dominates by a factor of ≈ 20.[1]

IV. TRACKING

Track reconstruction broadly consists of track *finding*, which proposes a set of track candidates given hits in a detector, and track *fitting*, which aims to predict the particle momentum and production vector of the matching track candidate. Reconstructed tracks are used for many downstream reconstruction and analysis tasks, including trigger decisions, jet tagging, and signal-background discrimination.

Most traditional algorithms make assumptions about the characteristics of the particles being reconstructed. This particle hypothesis typically includes the assumption that the particle track is approximately helical and that the particle has ionization and interaction properties typical of known particles. In this way, algorithms used in collider tracking such as the Kalman filter (KF) propagate a helical (or otherwise modeled) state vector and update it hit-by-hit as part of track finding. Machine learning approaches, such as graph neural networks, need not explicitly model this state and therefore need no such assumption. The characteristics of the particle tracks are learned by providing examples of target tracks, which may therefore be of any shape.

There are a variety of ways ML can be used to solve the track-finding problem (see [38–40] for thorough reviews). It can be treated as an object detection problem and one can perform instance segmentation with object condensation [41–43]. One can frame reconstruction as a generative process, where track candidates are recurrently produced by some autoregressive process [44]. The combinatorial Kalman filter (CKF) technique can also be combined with techniques for extending track seeds using ML [45], or using reinforcement learning to choose a next hit [46,47].

A successful approach has been to take advantage of a natural description of tracks as a set of paths in a directed graph, and apply a GNN. The Exa.TrkX pipeline is built on this GNN paradigm, and uses the physics intuition of a track being a chain of sequentially connected hits [25,48,49]. This and similar approaches have been shown to give high track-finding efficiency and stable training on a diverse set of HEP experiments [50–61]. This pipeline contains three stages: (i) graph construction, where 3D spacepoints are connected into a graph of spacepoint-to-spacepoint edge connections, (ii) edge classification, where edges are scored with a confidence of being either "true" (a connection between two sequential hits from a signal particle) or "fake" (any other kind of edge), and (iii) graph segmentation, where low-scoring edges are removed, and collections of connected hits are used to produce track candidates, using an algorithm called Connected Components [62,63].

The Exa.TrkX pipeline can be applied nearly out-of-the-box to nonhelical tracking, as in the case of quirk tracks. Some alterations to the definition of "true edges" generalize the pipeline to train correctly for this broader class of tracks. The most significant alteration is in the assumptions about the ordering of the hits. As many tracking detectors (as in ATLAS and the open detector used in the TrackML challenge [64–66]) do not include reliable timing information, the ordering of hits in tracks within Exa.TrkX is defined heuristically using increasing distance from the production vertex of the track. This is reliable for high-momentum helical tracks that are unlikely to scatter back toward the interaction point. However, quirk tracks with a non-negligible opening angle are able to oscillate away from and toward the detector origin. Relying on the assumption that hits are in order according to their distance from the production vertex can lead to misinterpretation of the true track shape in training. In this work, we access the underlying simulation data to define the time-ordering of hits in order to correctly construct true edges for training, though such information is not required for testing.

To construct graphs, we train a *metric learning* model, which attempts to embed sequential hits in a particle track into nearby positions in a latent space, while hits from other particles should be distant in the latent space. As this is purely data driven, tracks need not be helical, and any smooth particle trajectory may be learned by such a nonlinear embedding. All that is required is a definition of pairwise similarity, as defined earlier. In our training analyses, we split our dataset into training, validation, and testing sets with 2k, 400, and 400 samples, respectively. We note that we have compared these results to a training dataset with 4k events and did not find much improvement

[1]The data for the quirk and SM charged particle trajectories can be found in [37].

and an order of magnitude more events would require significantly higher computing resources. As we will discuss, we focus on certain regions of quirk parameter space to maximize track reconstruction, and it is likely that larger training samples would be useful for parameter space that lead to more exotic tracks. We use 150 training epochs and a learning rate of 0.001 in the metric learning for constructing the tracking graph. We then train a GNN to classify the edges as either "true" or "false." In this step, we modify the epochs and learning rate for better efficiency and to avoid overfitting by checking the loss and efficiency plots, using early stopping and decreasing the learning rate. Also, we define the "score cut" to select our target candidates. This value is set to 0.9 in most analyses and is optimized in some cases.

V. TRACK RECONSTRUCTION PERFORMANCE

Performance of the tracking algorithm is measured in simulated samples described above. Performance is measured by tracking efficiency, the fraction of *reconstructable* particles that are *reconstructed*. Reconstructable particles are those with at least seven (22) true hits in the eight-layer (25-layer) geometry. A particle is *reconstructed* if it is double-majority matched[2] to a *matchable* track candidate: one with at least five (15) hits in the eight-layer (25-layer) geometry.

Below, we demonstrate the challenges of finding quirks using helical track finders, study the power of ML tracking to find quirk tracks, and explore the ability of the pipeline to generalize from one set of quirk parameters to another. Rather than provide an exhaustive set of results for every configuration discussed below, we begin with a single characteristic choice of quirk model parameters, our *benchmark sample*, with $m_Q = 500$ GeV and $\Lambda = 500$ eV, before presenting a broader study across all parameters. We choose this benchmark as it displays an oscillation length in the region of interest $\mathcal{O}(\mathrm{cm})$.

A. Reconstructing quirk tracks using SM tracks

We wish to compare the performance of a traditional helix-finding algorithm to a dedicated ML quirk-finding approach. However, while powerful, the CKF is costly to tune on a new geometry. As such, for this study we take a GNN trained on helical SM tracks as the baseline, as it achieves similar physics performance as non-ML algorithms [67]. We take this GNN as a proxy for a general "helix-finding algorithm," which could also be a hand-tuned CKF or Hough transform, for example.

The SM-trained GNN-based tracker excels at finding SM tracks. For our benchmark sample, in the eight-layer geometry, SM tracks are found with 97.9% efficiency, while quirks are mostly missed, with 10.2% efficiency, using the definition of track efficiency above. For the 25-layer geometry, we find similar performance.

These results are typical for other m_Q, Λ in the regime of interest, where the oscillation length is centimeters. As expected, if the oscillation length is much longer, these trackers succeed in reconstructing the tracks due to the quirk track in the longer oscillation being similar to the SM tracks. See Tables II and III in the Appendix for details.

B. Well-behaved quirks

Tracks of quirks are substantially more complex than helical tracks, and over the range of parameters can display a broad range of behavior that can be challenging for reconstruction algorithms. As seen in Fig. 1, tracks can cross individual layers multiple times due to oscillations. Before tackling these more difficult scenarios, we focus on the subset of quirk tracks identified as *well behaved*, those whose number of hits do not exceed the number of layers by a factor of 3. This eliminates, for the moment, the most challenging quirk tracks, those which leave a very large number of hits as they travel in and out of detector layers, but retains tracks with clearly nonhelical trajectories.

In the benchmark sample with eight layers, training on a pure sample of well-behaved quirks produces a track finder with efficiency of 91.5% in events with only well-behaved quirks and no background tracks. This is a highly idealized scenario, but gives a rough upper bound on the performance of the pipeline and demonstrates its capacity to learn to find nonhelical tracks when they comprise its training set, without background. Figure 3 shows an example of reconstructed and truth tracks in this scenario.

C. All quirks

Quirks are an interesting and challenging problem precisely because they are not always well behaved, and performance on these not-well-behaved tracks is also important for broad sensitivity. When the training sample is expanded to include all reconstructable tracks, the efficiency in our benchmark scenario with eight detector layers is 56.3%, as compared to 10.2% for the tracker trained on SM tracks. Figure 4 presents an example of reconstructed and truth tracks.

The efficiency with respect to p_T, opening angle, and the number of true hits are shown in Fig. 5, for all quirks as well as the well-behaved subset. The efficiency falls as the number of hits increases, reflecting the challenge of reconstructing tracks with complex in-and-out trajectories. There is also a soft dependence on opening angle and quirk track p_T, due to tracks with wider opening angles or higher momentum being more likely to exhibit complex oscillations with in-and-out behavior. The efficiency shown in Fig. 5 illustrates that the best performance is obtained on well-behaved quirks. There is thus a trade-off in choosing a

[2]Double-majority matching requires that at least 50% of the reconstructed track hits belong to a matching truth track, and that at least 50% of the hits in that truth track are found in the reconstructed track.

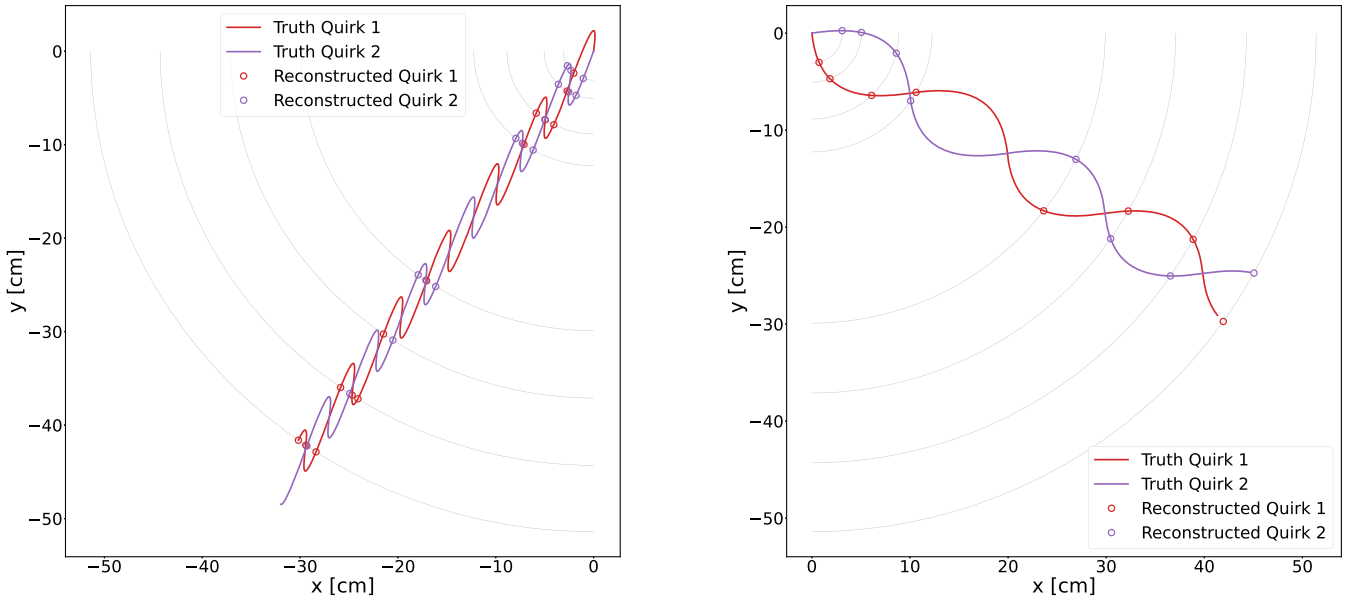


FIG. 3. Example truth and reconstructed well-behaved quirk tracks with ($m_Q = 500$ GeV, $\Lambda = 500$ eV).

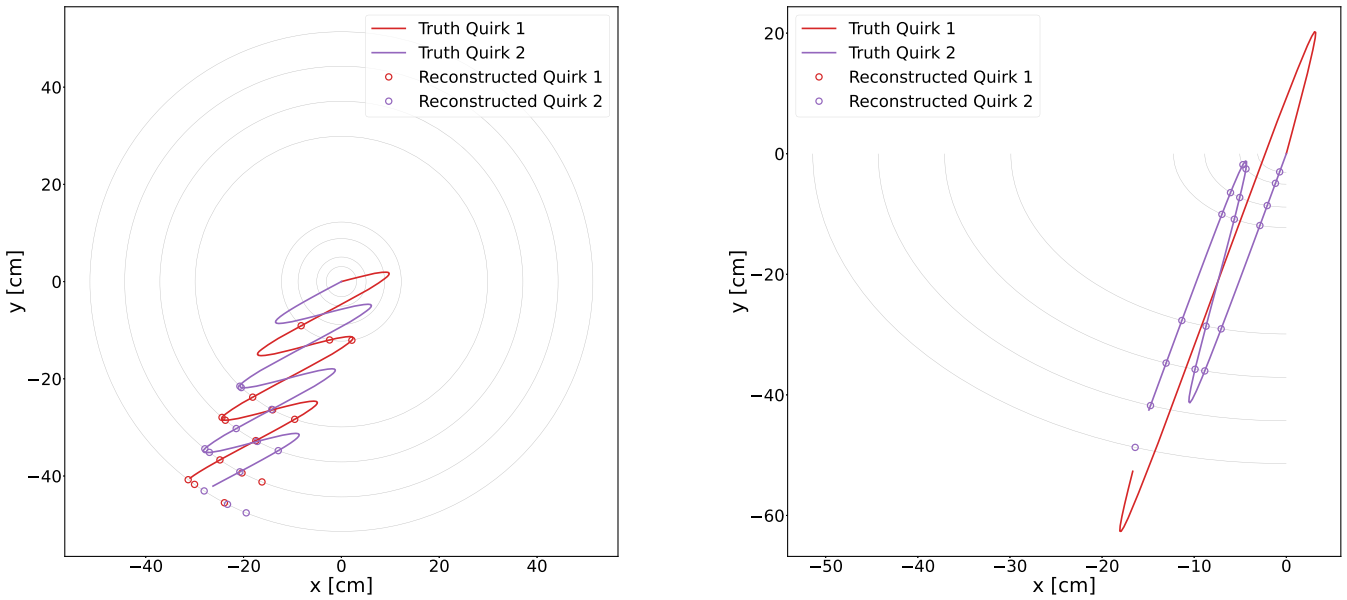


FIG. 4. Examples of truth and reconstructed tracks for not-well-behaved quirks with $m_Q = 500$ GeV, $\Lambda = 500$ eV.

model configuration that has some acceptable discovery potential given computational constraints. Here we choose that working point to be those well-behaved quirks defined above.

The performance on the full set of quirk tracks is a significant improvement relative to the efficiency of the SM-trained algorithm. However, due to the complexity of reconstruction of the all-quirks dataset, we leave the broader challenge of optimizing the all-quirks reconstruction, and indeed any parametrizable path, to future work, and base our analysis on well-behaved quirks in the following sections. The well-behaved quirks contain sufficient complexity to represent a novel tracking challenge and a large enough fraction of the all-tracks dataset to provide sensitivity to quirks at the LHC.

D. Finding quirks among the background

Realistic applications of nonhelical tracking require identifying such tracks among many background helical tracks from SM processes. To assess the performance in this environment, we use a training sample which includes

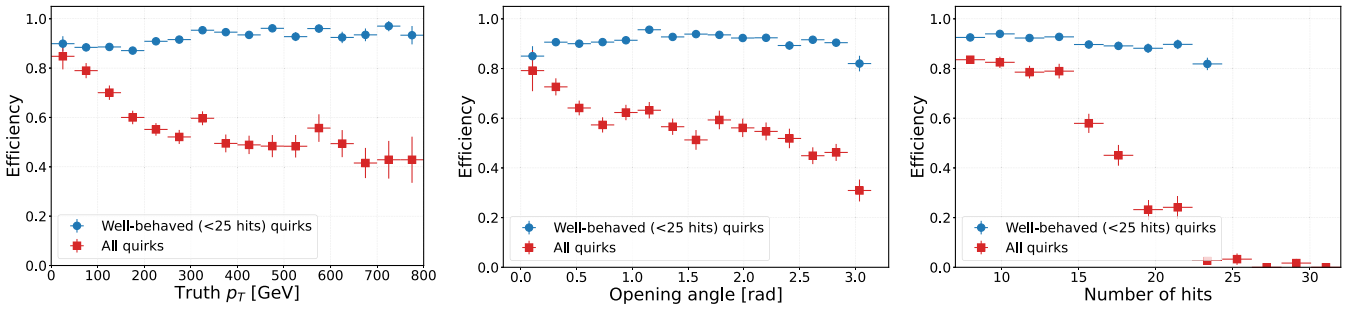


FIG. 5. Tracking performance for quirks without background for $m_Q = 500$ GeV, $\Lambda = 500$ eV in the eight-layer geometry. Shown is efficiency versus track p_T (left), quirk opening angle (center), and number of true hits (right) for a model trained on well-behaved quirks (those with 24 or fewer hits) searching for well-behaved quirks (blue) and one trained on all quirks and searching for all quirks (red). The error bars indicate statistical errors.

tracker hits from both quirk and SM particles in each event.[3] In graph construction and edge classification, only quirk hit pairs are treated as true. All other pairs, including pairs of hits from the same SM particle, are treated as fake in training, and models are encouraged to *not* reconstruct them. Previous work has explored other strategies with such background tracks, such as treating them as true or masking them from the loss function [50]. In this work, our goal is quirk finding at the highest possible background rejection rate, rather than the reconstruction of all true tracks, and this motivates the strategy, as described. The efficiency in this scenario is 71.9% for all quirk tracks in our benchmark sample with eight detector layers, somewhat less efficient than the much simpler task of background-free tracking, which scores 91.5%. The fake rate from track finding, defined as one minus the ratio of matched track candidates to all track candidates, is globally 5% for the ($m_Q = 500$ GeV, $\Lambda = 500$ eV) working point, with efficiency given in Table I. Across the quirk parameter space, fake rates increase with Λ and are $\mathcal{O}(1)\%$ for $\Lambda = 100$ eV and can reach as high as 50% for $\Lambda = 5000$ eV, but do not vary significantly with m_Q, however, we note that this number can be misleading for the overall performance, as it describes fake rates before fitting. Our prefit fake rate is considerably higher than the typical fake rate of $O(10^{-4})$ obtained from the expected Run 4 ATLAS tracking chain [68] but, as we demonstrate in Sec. VI, quirk and SM tracks can be distinguished after fitting by using both the fit quality and the opening angle between the track pair.

Most of the fakes measured here are the result of combined quirk tracks not being matched according to the strict double-majority rule, rather than background hits being built into completely fictitious tracks. This suggests that a more sophisticated track segmentation algorithm than Connected Components would benefit these merged or

[3]A reminder that the SM particles are those produced in the hard-scatter jet, giving a typical number of hits $\approx 10^3$–10^4. We do not include pileup in these studies.

TABLE I. Efficiency and fake rate to reconstruct quirk tracks for various values of quirk mass (m_Q) and confinement scale (Λ) based on a 25-layer and *well-behaved* setting. We also calculate the mean Lorentz factor of the quirks ($\bar{\gamma}$) and the standard deviation (σ_γ) in each of the generated datasets to calculate oscillation length d, as calculated by Eq. (1), which can be useful for understanding reconstruction efficiencies. See the text for details.

m_Q (GeV)	Λ (eV)	$\bar{\gamma}$	σ_γ	d [cm]	Efficiency (%)	*Well-behaved* fraction (%)
100	100	4.4	3.4	670	91.0	88.3
	500	3.7	2.7	21.6	82.8	77.0
	1000	3.1	2.05	4.2	77.8	79.7
	2000	3.0	2.0	1.0	60.4	83.4
	3000	3.0	1.9	0.44	35.6	83.6
	4000	2.9	1.8	0.24	34.5	84.5
	5000	2.9	1.7	0.15	24.2	85.0
500	100	1.9	0.7	896	92.0	82.3
	500	1.8	0.6	31	79.0	51.2
	1000	1.7	0.6	7.3	64.3	53.1
	2000	1.7	0.5	1.7	60.9	59.9
	3000	1.7	0.5	0.8	59.6	62.3
	4000	1.6	0.5	0.4	42.6	63.0
	5000	1.6	0.5	0.2	39.2	63.8
1000	100	1.5	0.3	950	94.6	80.5
	500	1.4	0.3	32	63.6	40.2
	1000	1.4	0.3	7.6	62.7	42.3
	2000	1.4	0.3	1.8	69.2	48.6
	3000	1.3	0.2	0.8	54.1	51.7
	4000	1.3	0.2	0.4	59.7	52.9
5000	100	1.04	0.03	420	84.8	40.2
	500	1.03	0.02	11	69.8	32.6
	1000	1.03	0.02	2.7	69.5	35.2
	2000	1.03	0.02	0.7	49.2	39.6
	3000	1.03	0.02	0.3	36.4	40.8
	4000	1.03	0.02	0.2	34.6	41.2

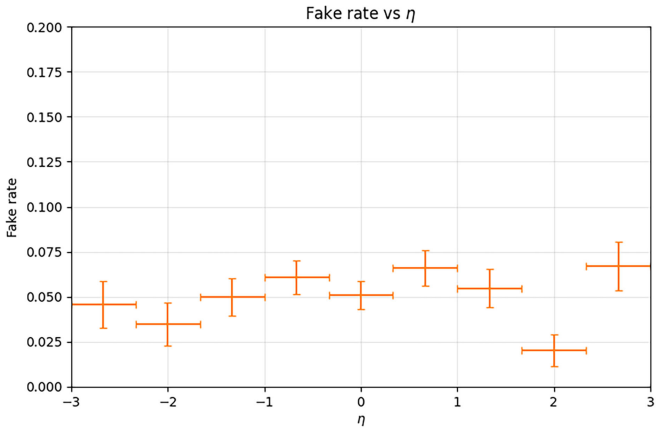


FIG. 6. Fake rate for $m_Q = 500$ GeV, $\Lambda = 500$ eV, as a function of track pseudorapidity η.

noisy track candidates, as demonstrated in [50,63]. In these works, a walkthrough of the graph effectively separates and cleans noisy tracks such as those seen in our candidate set. Fake production across pseudorapidity is presented in Fig. 6, where stable performance is observed across the detector coverage. The duplicate rate, defined as the rate of reconstructable particles being matched by more than one track candidate, is by construction zero, as the double-majority matching overconstrains against this possibility. Background tracks are also included in training discussed in subsequent studies in the following sections, as well as the results in Table I.

E. Dependence on geometry

To explore the dependence of the number of tracking layers, the study was repeated in the 25-layer setting, yielding an efficiency of 79.0%, somewhat higher than the efficiency in the eight-layer case. This suggests that the nature of the problem does not sensitively depend on the number of layers. Figure 7 presents an example of reconstructed and truth tracks with 25 layers for $m_Q = 500$ GeV, $\Lambda = 500$ eV. The 25-layer geometry is used in studies in the following sections.

F. Dependence on quirk parameters

Results above demonstrate effective tracking for a single value of quirk parameters, but the task of quirk track reconstruction is much broader, and the nature of the tracking challenge varies strongly with the parameter values. Table I shows the typical oscillation length and measured reconstruction efficiency for quirks over a range of quirk parameters m_Q, Λ.

For very long oscillation lengths, $\mathcal{O}(\mathrm{m})$, no oscillations occur in the tracking volume and the problem is essentially indistinguishable from that of the SM tracking, and efficiencies are very high, ≈90%. For the target regime of intermediate-length oscillation lengths, $\mathcal{O}(\mathrm{cm})$, performance is in the range of 60–80%, which is a dramatic improvement over the ≈10% efficiency for quirk tracks when using a SM-trained tracker (compare Tables I and II). As the oscillation length becomes very small, reported efficiency drops by a factor of ≈2, and this is due to the two quirk tracks overlapping nearly exactly, often leading to only one reconstructed track, though in principle two exist, see Fig. 8.

G. Generalization across quirk parameters

The studies above demonstrate the capacity for the pipeline to reconstruct tracks for particular quirk

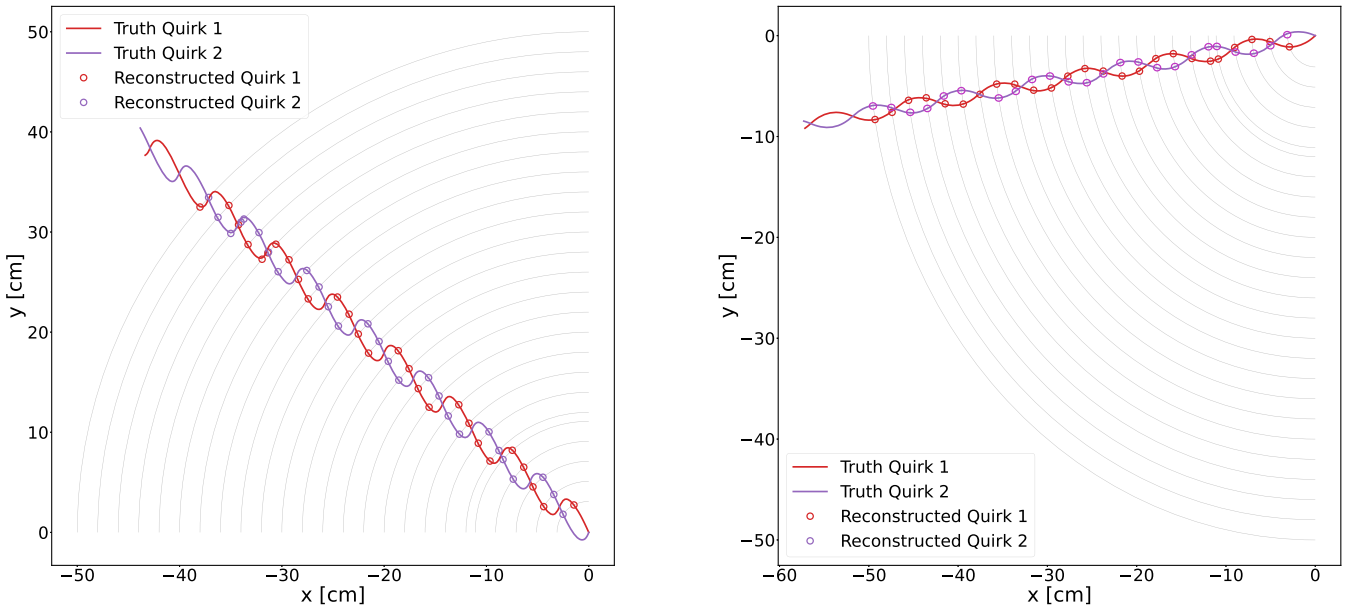


FIG. 7. Example truth and reconstructed quirk tracks for $m_Q = 500$ GeV, $\Lambda = 500$ eV in the 25-layer geometry.

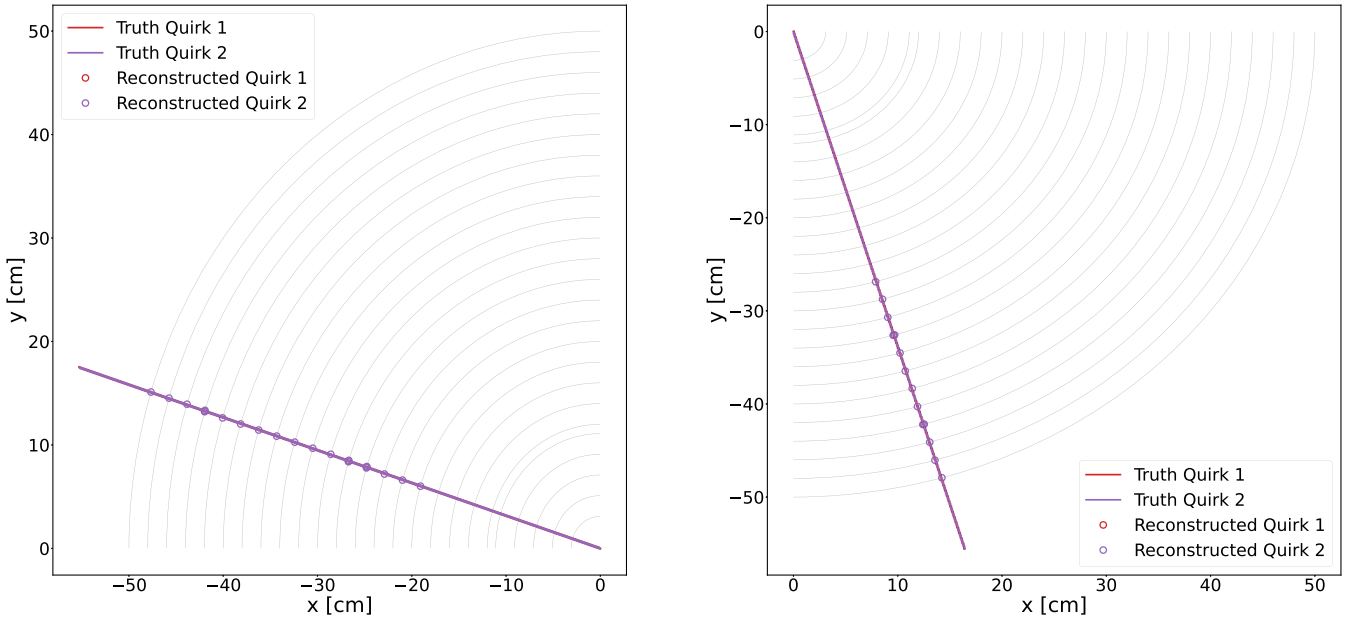


FIG. 8. Examples of truth and reconstructed tracks for well-behaved quirks with $m_Q = 500$ GeV, $\Lambda = 5000$ eV, in which there is no visible oscillation and the two quirks overlap.

parameters when those tracks are present in the training sample. It would be much more desirable, however, to develop a *single* pipeline capable of reconstructing *any* quirk track, or more ambitiously, any smooth, continuous particle path. Achieving such a broad capacity is left to future work, but the rough ability of the current pipeline to generalize can already be assessed by training on a sample with several quirk parameter sets, rather than a single choice, and testing on a sample of tracks from a parameter set which does not appear in the training sample.

The efficiency to reconstruct quirks generated with $m_Q = 100$ GeV, $\Lambda = 1000$ eV is 52.6% when training with a mixture that does not include that set, compared to 77.8% when training only on tracks from the same parameter set as the testing sample. The mixed set includes m_Q (GeV), Λ (eV) $= (100, 100), (500, 100), (1000, 100)$ $(100, 500)(5000, 1000)$. This demonstrates some significant but imperfect ability to generalize. Another working point with $m_Q = 1000$ GeV, $\Lambda = 1000$ eV achieves 49.3% efficiency for the generalized training, compared to 62.7% for the specific training, a similar performance ratio. The nature of the tracking challenge varies with quirk parameter values, but each working point clearly has enough in common for the tracker to learn general features of quirk tracks, not just the features of quirks for one particular quirk parameter choice. Further work is required to understand which ML choices contribute to the highest generalizability.

H. Hit location uncertainty

To explore the sensitivity of the pipeline to imperfect detector resolution, we smear the hit locations in training and testing with a Gaussian of width 1 mm, which is larger than typical LHC tracking resolution [21,69]. A spot test for $m_Q = 500$ GeV, $\Lambda = 500$ eV quirks shows no significant degradation in tracking performance, with performance consistent within statistical uncertainties.

VI. BACKGROUNDS AND SENSITIVITY

The pipeline has demonstrated the capacity to find quirks with a reasonable efficiency, which is necessary to observe quirks in experiments at the LHC. However, equally important to control is the background (or "fake") rate. Quirks are an exciting target because their signature is striking enough to suggest the possibility of very low background, potentially even background-free, discoveries. Unfortunately, SM tracks have a very high rate, so even a tiny probability to mimic the quirk signature could significantly impact discovery potential.

To examine the rate of background from SM tracks, we collect samples of fake tracks reconstructed by the pipeline which are not matched to quirks. We perform two fits to each track, one to the quirk hypothesis and a second to the helix hypothesis, and use the ratio of χ^2 values to discriminate between quirks and the background. In both the quirk and helix hypothesis, the χ^2 is the sum of the squared distances between a hit and the distance of closest approach by the track, with an implicit $\sigma^2 = 1$ cm^2 for simplicity.

The helix fitter finds the value of the five helical parameters that minimize the χ^2_{helix}, starting from an initial estimate [70] and refined by the Nelder-Mead optimization algorithm. This approach does not match the sophistication

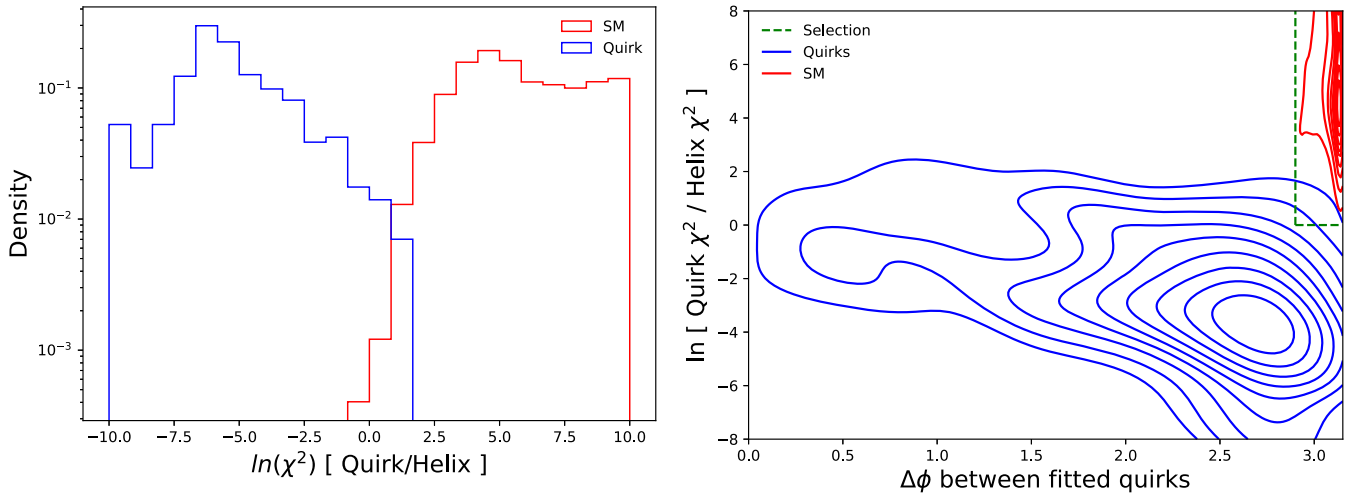


FIG. 9. Distributions of ratio χ^2 under the quirk and helix hypothesis, for $m_Q = 500$ GeV, $\Lambda = 500$ eV quirk and SM tracks. The lower panel shows the ratio of the two χ^2 metrics versus the fitted quirk opening angle for the same m_Q, Λ, and the selection used to remove SM tracks is shown in the dashed red lines.

of track fitting in a realistic setting, which benefits from additional heuristics, such as outlier removal and ambiguity resolution. For our purposes, this provides a conservative estimate on our ability to distinguish between true quirks and background. The distribution of χ^2_{helix} is shown in the upper left panel of Fig. 9 for true quirks at our benchmark point $(m_Q = 500, \Lambda = 500)$ and for background particles. There is a clear separation between the two distributions, as many quirks cannot be well fit to a helix. A significant population of background tracks also receive a χ^2_{helix} larger than what might be expected, due to incorrect hits and local minima in the fitting.

In the quirk hypothesis, we use the propagation algorithm as described in Sec. III, which requires that we fit simultaneously for the four-vectors of the quirk pair given values of (m_Q, Λ), though only one track is required and included in the χ^2 evaluation. The distribution of the ratio of χ^2_{quirk} to χ^2_{helix} values are shown in the left panel of Fig. 9 for true quirks and for background SM tracks. Both background tracks can be accommodated by the quirk hypothesis, due to the freedom provided by the second quirk's unconstrained parameters. To mimic a high-momentum SM track, the fitter finds solutions in which the two quirks are back to back, generating a perfectly straight line. To suppress these, we remove tracks whose fitted quirks are back to back $(\Delta\phi > 2.8)$ or whose ratio of $\ln(\chi^2_{\text{quirk}}/\chi^2_{\text{helix}}) < -1$, see the lower panel of Fig. 9. For $m_Q = 500$ GeV, $\Lambda = 500$ eV, this requirement is 94%

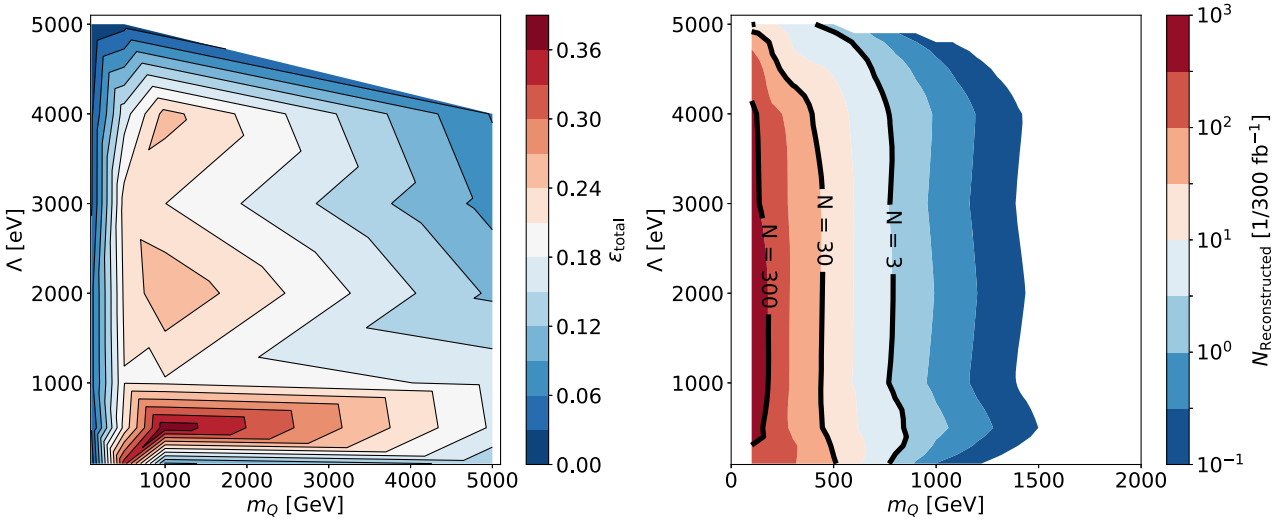


FIG. 10. Left: efficiency of quirk finding versus mass m_Q and string tension Λ. Right: expected reconstructed quirk event rate per 300 fb^{-1} in the same parameter plane.

efficient for quirks and removes all SM tracks in our samples. For other m_Q, Λ comparable efficiencies of ≈80% are found, except for the heaviest case $m_Q = 5000$ GeV, due to their larger opening angles (however, 5 TeV quirks have a small cross section and are not produced at any appreciable rate). To ensure a negligible background, these requirements can be tuned in data, but these studies suggest that such suppression is possible. Further suppression is available by fitting two quirk candidate tracks simultaneously, if needed. Additional discrimination power may be available from the momentum of the recoiling jet, which is correlated with the quirk pair momentum.

To estimate the potential of discovering the quirk signal among the SM background, we estimate the number of reconstructed quirk events, $N_{\text{Reconstructed}}$, as follows. Following Ref. [18], we select quirk events that include a $p_T > 200$ GeV jet to pass the high-level trigger, and calculate the number of reconstructed quirk events as

$$N_{\text{Reconstructed}} = \mathcal{L} \times \sigma(Q\bar{Q}j, p_T^{\text{jet}} > 200\ \text{GeV}) \times \epsilon_{\text{total}}, \quad (3)$$

where ϵ_{total} is the total efficiency for reconstruction, selection, and also includes the fraction of well-behaved quirks at this parameter point. The efficiency ϵ_{total} varies from $\lesssim$1% to ≈30% across the plane, as shown in the left panel of Fig. 10. The right panel of Fig. 10 shows the expected quirk event rate $N_{\text{Reconstructed}}$ in 300 fb^{-1} of pp collisions at the LHC, showing $N > 3$ across $\Lambda \lesssim 5000$ eV for $m_Q \lesssim 700$ GeV.

VII. DISCUSSION

Reconstruction of quirk tracks in the region where the oscillation length (d) is larger than detector resolution but smaller than the detector is expected to be very challenging. The application of ML-based tracking demonstrates a newfound capability to reconstruct such tracks with reasonable efficiency, in excess of 50% across a range of parameter values where $d \approx 1$–10 cm. For comparison, previous work which searches for coplanar hits to avoid fitting quirk trajectories [18] has a similar reconstruction efficiency for small oscillation lengths, $d < 1$ cm, but drops very quickly for cm-scale oscillations, reaching ≈10% for $m_Q, \Lambda = 800$ GeV, 1000 eV. Training ML trackers explicitly on quirks can significantly extend the sensitivity range, complementing the coplanar approach. The technique presented in Sec. V is found to be sensitive to $m_Q \lesssim 700$ GeV, for 100 eV $< \Lambda <$ 5000 eV.

One approach for the implementation of an ML-based quirk-finding algorithm is to analyze a set of events which are selected for other reasons, such as large p_T^{miss}. But a more ambitious and much more challenging direction would be to generalize the core track finding in the reconstruction pipeline so that it is capable of identifying helical and nonhelical tracks, whether those are quirks or other unexpected particles.

Reference [18] estimated the trigger efficiency for quirks directly to be low with existing tools, relying on primarily muon triggers. However, an ML-based tracker for quirks opens a new possibility of identifying quirks directly in the trigger, using knowledge distillation [71] and hardware accelerators such as GPUs to fit the GNN within the trigger budget. The computational cost of this pipeline at high luminosity remains to be studied. However, we observe, with no dedicated tuning of the Exa.TrkX pipeline, the metric learning stage to produce approximately 180 edges per node. This is a factor of 3 more dense than those produced for standard model tracks [50]. This gives an approximate scaling factor for the extra cost of a quirk tracking pass. This is comparable with, for example, the extra computational cost of a large-radius tracking pass used in the ATLAS experiment [72].

VIII. CONCLUSIONS

The analysis of data from particle physics experiments necessarily sacrifices some general sensitivity for the sake of practical computability, leaving us blind to many potentially striking signatures of new particles. An example of this is quirks, confining particles where the quirk mass, m_Q, is greater than the confinement scale, Λ, opposite to that of quarks. This hierarchy leads to an oscillating pair of nonhelical tracks from the particles once they are pair produced at the LHC and propagate through the detector. Current tracking algorithms fail to reconstruct the quirk tracks when their oscillation is on the scale of the detector, $\mathcal{O}$(cm). Larger oscillations would appear as two helical tracks, and shorter ones would appear as a nearly straight track through the detector.

Recent advances in ML-based reconstruction allow for new inroads into the previously inaccessible territory. In this work we have shown that graph neural network tracking can be trained to reconstruct quirk tracks with macroscopic oscillation lengths in the problematic centimeter range. We have generated quirk pseudodata across the m_Q, Λ parameter space in a simplified model of the ATLAS detector and trained the Exa.TrkX machine learning algorithm to reconstruct quirks. When trained on only SM tracks, the algorithm naturally fails to reconstruct tracks, but when instead trained on quirks the algorithm reaches comparatively large reconstruction efficiencies, ∼40%–90%. We have also tested the robustness of the algorithm, by including background tracks from the associated jet in the signal process, and found that these additional backgrounds do not significantly degrade our reconstruction performance or the sensitivity to discovering quirks. It remains to be understood the performance of traditional algorithms such as the CKF in nonhelical track searches. Certainly those configured for helical track finding have very low track-finding efficiency, as shown

in Sec. VA with a GNN proxy for a helical CKF. Tuning such an algorithm to nonhelical tracks has, to our knowledge, never been performed. It is unclear how onerous building such an algorithm would be, particularly in defining the seed construction stage. We note for future work that the Exa.TrkX pipeline has been successfully used to produce track seeds, which can then be used for a track extension algorithm, and this approach may render a CKF quirk finder more tractable [73].

Effects from pile-up remain to be studied rigorously in future work, but existing implementations of GNN reconstruction in high pile-up [high luminosity (HL)-LHC-like] conditions are performant with comparable sensitivity to traditional techniques [25,50]. Although quirk trajectories are clearly more topologically complex than a SM hadron, we have no reason to suspect that their reconstruction is particularly sensitive to pile-up and noise effects, and remain optimistic that future studies in this direction will maintain the high efficiency demonstrated in this work. The agnostic nature of the ML pipeline to the specific features of these tracks suggests a potentially broader capability to reconstruct nonhelical tracks in a model-independent way.

The application of such techniques to the existing vast LHC dataset might identify quirk candidates, which could be nearly background-free, as the probability for a SM track to exhibit such oscillations or of random hits creating such a path are near zero.

ACKNOWLEDGMENTS

The authors are grateful to Simon Knapen and Michele Papucci for useful discussions and for sharing the quirk trajectory code from [18]. We are also grateful to Markus Elsing and Xiangyang Ju for their comments and expertise. D. W. is funded by the DOE Office of Science. Y. F. and Q. S. are supported partially by the Institute of High Energy Physics, Chinese Academy of Sciences, under the innovative project on sciences and technologies with No. E3545BU210. D. M. is supported by the Danish Data Science Academy, which is funded by the Novo Nordisk Foundation (No. NNF21SA0069429). The work of M. F. was supported in part by U.S. National Science Foundation Grants No. PHY-2111427 and No. PHY-2210283, as well as by NSF Graduate Research Fellowship Award No. DGE-1839285.

DATA AVAILABILITY

The data that support the findings of this article are openly available [37].

APPENDIX: SM EFFICIENCY

TABLE II. Efficiency to reconstruct quirk tracks for various values of quirk mass (m_Q) and confinement scale (Λ), when training on the SM trackers and inference on the quirk trackers based on a 25-layer and *well-behaved* setting.

	m_Q (GeV)			
Λ (eV)	100 (%)	500 (%)	1000 (%)	5000 (%)
100	81.3	89.9	93.2	80.5
500	19.0	7.0	5.4	2.2
1000	18.8	12.8	4.8	13.0
2000	15.4	11.8	31.2	8.0
3000	18.6	32.2	24.6	15.2
4000	17.0	21.0	29.6	18.8
5000	29.8	24.6	...	...

TABLE III. Score cut for different points.

m_Q (GeV)	Λ (eV)	$\bar{\gamma}$	σ_γ	d [cm]	Efficiency (%)	Score cut
100	100	4.4	3.4	670	91.0	0.9
	500	3.7	2.7	21.6	82.8	0.9
	1000	3.1	2.05	4.2	77.8	0.9
	2000	3.0	2.0	1.0	60.4	0.6
	3000	3.0	1.9	0.44	35.6	0.5
	4000	2.9	1.8	0.24	34.5	0.5
	5000	2.9	1.7	0.15	24.2	0.5
500	100	1.9	0.7	896	92.0	0.9
	500	1.8	0.6	31	79.0	0.9
	1000	1.7	0.6	7.3	64.3	0.9
	2000	1.7	0.5	1.7	60.9	0.6
	3000	1.7	0.5	0.8	59.6	0.6
	4000	1.6	0.5	0.4	42.6	0.5
	5000	1.6	0.5	0.2	39.2	0.5
1000	100	1.5	0.3	950	94.6	0.9
	500	1.4	0.3	32	63.6	0.9
	1000	1.4	0.3	7.6	62.7	0.6
	2000	1.4	0.3	1.8	69.2	0.7
	3000	1.3	0.2	0.8	54.1	0.6
	4000	1.3	0.2	0.4	59.7	0.6
5000	100	1.04	0.03	420	84.8	0.9
	500	1.03	0.02	11	69.8	0.9
	1000	1.03	0.02	2.7	69.5	0.6
	2000	1.03	0.02	0.7	49.2	0.5
	3000	1.03	0.02	0.3	36.4	0.5
	4000	1.03	0.02	0.2	34.6	0.3

[1] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **716**, 1 (2012).
[2] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **716**, 30 (2012).
[3] T. Gehrmann and B. Malaescu, Annu. Rev. Nucl. Part. Sci. **72**, 233 (2022).
[4] P. Ferreira da Silva, Annu. Rev. Nucl. Part. Sci. **73**, 255 (2023).
[5] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **115**, 072001 (2015).
[6] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **113**, 151601 (2014).
[7] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. Lett. **105**, 252303 (2010).
[8] V. Khachatryan *et al.* (CMS Collaboration), J. High Energy Phys. 09 (2010) 091.
[9] A. Hayrapetyan *et al.* (CMS Collaboration), J. High Energy Phys. 07 (2024) 142.
[10] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. 06 (2023) 200.
[11] A. Hayrapetyan *et al.* (CMS Collaboration), Rep. Prog. Phys. **88**, 037801 (2025).
[12] L. Lee, C. Ohm, A. Soffer, and T.-T. Yu, Prog. Part. Nucl. Phys. **106**, 210 (2019); **122**, 103912(E) (2022).
[13] J. Alimena *et al.*, J. Phys. G **47**, 090501 (2020).
[14] A. Hayrapetyan *et al.* (CMS Collaboration), Phys. Rep. **1115**, 448 (2025).
[15] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. 11 (2023) 112.
[16] A. Hayrapetyan *et al.* (CMS Collaboration), Phys. Rev. Lett. **134**, 131802 (2025).
[17] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **96**, 201801 (2006).
[18] S. Knapen, H. K. Lou, M. Papucci, and J. Setford, Phys. Rev. D **96**, 115015 (2017).
[19] A. Strandlie and R. Fruhwirth, Rev. Mod. Phys. **82**, 1419 (2010).
[20] M. Aaboud *et al.* (ATLAS Collaboration), Eur. Phys. J. C **77**, 673 (2017).
[21] S. Chatrchyan *et al.* (CMS Collaboration), J. Instrum. **9**, P10009 (2014).
[22] R. O. Duda and P. E. Hart, Commun. ACM **15**, 11 (1972).
[23] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, IEEE Signal Process. Mag. **34**, 18 (2017).
[24] X. Ju *et al.* (Exa.TrkX Collaboration), in *33rd Annual Conference on Neural Information Processing Systems* (Neural Information Processing Systems Foundation, CA, USA, 2020).
[25] X. Ju *et al.* (Exa.TrkX Collaboration), Eur. Phys. J. C **81**, 876 (2021).
[26] J. Kang and M. A. Luty, J. High Energy Phys. 11 (2009) 065.
[27] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **105**, 211803 (2010).
[28] M. Farina and M. Low, Phys. Rev. Lett. **119**, 111801 (2017).
[29] J. A. Evans and M. A. Luty, J. High Energy Phys. 06 (2019) 090.
[30] J. L. Feng, J. Li, X. Liao, J. Ni, and J. Pei, J. High Energy Phys. 06 (2024) 197.
[31] J. Li, J. Pei, L. Ran, and W. Zhang, J. High Energy Phys. 12 (2021) 109.
[32] J. Li, X. Liao, J. Ni, and J. Pei, Phys. Rev. D **109**, 095005 (2024).
[33] R. Mammen Abraham, J. Adhikary, J. L. Feng, M. Fieg, F. Kling, J. Li, J. Pei, T. R. Rabemananjara, J. Rojo, and S. Trojanowski, J. High Energy Phys. 01 (2025) 094.
[34] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, J. High Energy Phys. 07 (2014) 079.
[35] C. Bierlich *et al.*, SciPost Phys. Codebases **2022**, 8 (2022).
[36] M. Luscher and P. Weisz, J. High Energy Phys. 07 (2002) 049.
[37] M. Fieg, Quirk and background hits (2025), 10.5281/zenodo.17941765.
[38] J. Duarte and J.-R. Vlimant, in *Artificial Intelligence for High Energy Physics* (World Scientific, Singapore, 2022), pp. 387–436.
[39] G. DeZoort, P. W. Battaglia, C. Biscarat, and J.-R. Vlimant, Nat. Rev. Phys. **5**, 281 (2023).
[40] S. Thais, P. Calafiura, G. Chachamis, G. DeZoort, J. Duarte, S. Ganguly, M. Kagan, D. Murnane, M. S. Neubauer, and K. Terao, arXiv:2203.12852.
[41] J. Kieseler, Eur. Phys. J. C **80**, 1 (2020).
[42] K. Lieret and G. DeZoort, EPJ Web Conf. **295**, 09004 (2024).
[43] D. Murnane, EPJ Web Conf. **295**, 09016 (2024).
[44] A. Huang, Y. Melkani, P. Calafiura, A. Lazar, D. T. Murnane, M.-T. Pham, and X. Ju, arXiv:2402.10239.
[45] L. Heinrich, B. Huth, A. Salzburger, and T. Wettig, arXiv:2401.16016.
[46] L. H. Vaage, in *Connecting the Dots Workshop (CTD)* (Zenodo, Geneva, Switzerland, 2022).
[47] T. Kortus, R. Keidel, N. R. Gauger *et al.* (pCT Collaboration) *et al.*, *IEEE Trans. Pattern Anal. Mach. Intell.* 45, 15820 (2023).
[48] S. Farrell, P. Calafiura, M. Mudigonda, Prabhat, D. Anderson, J.-R. Vlimant, S. Zheng, J. Bendavid, M. Spiropulu, G. Cerati *et al.*, arXiv:1810.06111.
[49] X. Ju, S. Farrell, P. Calafiura, D. Murnane, L. Gray, T. Klijnsma, K. Pedro, G. Cerati, J. Kowalkowski, G. Perdue *et al.*, arXiv:2003.11603.
[50] S. Caillou, P. Calafiura, C. Rougier, J. Stark, D. T. Murnane, A. Vallier, X. Ju, and S. A. Farrell, CERN Technical Report No. ATL-COM-ITK-2022-057, 2022.
[51] Y. Verma and S. Jena, arXiv:2012.08515.
[52] J. Hewes, A. Aurisano, G. Cerati, J. Kowalkowski, C. Lee, W.-k. Liao, A. Day, A. Agrawal, M. Spiropulu, J.-R. Vlimant *et al.*, EPJ Web Conf. **251**, 03054 (2021).
[53] F. Drielsma, K. Terao, L. Dominé, and D. H. Koh, arXiv:2102.01033.
[54] A. Akram and X. Ju, arXiv:2208.12178.
[55] X. Jia, X. Qin, T. Li, X. Huang, X. Zhang, N. Yin, Y. Zhang, and Y. Yuan, EPJ Web Conf. **295**, 09006 (2024).

[56] C. Biscarat, S. Caillou, C. Rougier, J. Stark, and J. Zahreddine, EPJ Web Conf. **251**, 03047 (2021).
[57] M. Zdybal, M. Kucharczyk, and M. Wolter, Comput. Sci. **25**, 25 (2024).
[58] M. Andrews, B. Burkle, S. Chaudhari, D. Di Croce, S. Gleyzer, U. Heintz, M. Narain, M. Paulini, and E. Usai, EPJ Web Conf. **251**, 03057 (2021).
[59] G. DeZoort, S. Thais, J. Duarte, V. Razavimaleki, M. Atkinson, I. Ojalvo, M. Neubauer, and P. Elmer, Comput. Softw. Big Sci. **5**, 1 (2021).
[60] A. Correia, F. I. Giasemis, N. Garroum, V. V. Gligorov, and B. Granado, J. Instrum. **19**, P12022 (2024).
[61] A. Elabd, V. Razavimaleki, S.-Y. Huang, J. Duarte, M. Atkinson, G. DeZoort, P. Elmer, S. Hauck, J.-X. Hu, S.-C. Hsu *et al.*, Front. Big Data **5**, 828666 (2022).
[62] J. E. Hopcroft and R. E. Tarjan, Commun. ACM **16**, 372 (1973).
[63] D. T. Murnane, R. Liu, L. H. J. Condren, A. Vallier, D. Whiteson, A. Lazar, and X. Ju (ATLAS Collaboration), High performance graph segmentation for ATLAS GNN track reconstruction (2024), https://cds.cern.ch/record/2915350.
[64] M. Kiehn *et al.*, EPJ Web Conf. **214**, 06037 (2019).
[65] S. Amrouche, L. Basara, P. Calafiura, V. Estrade, S. Farrell, D. R. Ferreira, L. Finnie, N. Finnie, C. Germain, V. V. Gligorov *et al.*, *The Tracking Machine Learning Challenge: Accuracy Phase* (Springer International Publishing, New York, 2019), pp. 231–264.
[66] S. Amrouche, L. Basara, P. Calafiura, D. Emeliyanov, V. Estrade, S. Farrell, C. Germain, V. V. Gligorov, T. Golling, S. Gorbunov *et al.*, Comput. Softw. Big Sci. **7**, 1 (2023).
[67] H. Torres (ATLAS Collaboration) (2023), https://cds.cern.ch/record/2876457.
[68] G. Aad *et al.*, J. Instrum. **20**, P02018 (2025).
[69] B. Mindur (ATLAS Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **845**, 257 (2017).
[70] M. Hansroul, H. Jeremie, and D. Savard, Nucl. Instrum. Methods Phys. Res., Sect. A **270**, 498 (1988).
[71] G. Hinton, O. Vinyals, and J. Dean, arXiv:1503.02531.
[72] J. C. Burzynski (ATLAS Collaboration), Tech. Rep., CERN, Geneva, 2022, https://cds.cern.ch/record/2815527.
[73] B. Huth, Dissertation, Universität Regensburg, 2024, https://epub.uni-regensburg.de/55604/.