ColliderML: A High-Luminosity Detector Simulation Dataset for Machine Learning Benchmarks

Paul Gessinger
CERN

Daniel Murnane[*]
Niels Bohr Institute
& Lawrence Berkeley National Lab

Andreas Salzburger
CERN

Stine Kofoed Skov
University of Copenhagen

Andreas Stefl
CERN

Marcus Selchou Raaholt
University of Copenhagen

Anna Zaborowska
CERN

Abstract

We introduce COLLIDERML - a large, open, experiment-agnostic dataset of fully simulated, digitized, and reconstructed proton–proton collisions at HL-LHC conditions ($\sqrt{s} = 14\,\text{TeV}$, mean pile-up $\mu \approx 200$). COLLIDERML provides one million events across ten Standard Model and Beyond Standard Model processes, plus extensive single-particle samples, all produced with modern NLO/LO matrix element calculation and showering, realistic per-event pile-up overlay, a validated OpenDataDetector (ODD) geometry, and standard reconstructions. The release fills a major gap for ML research on detector-level data: it is comprehensive (multiple channels, full detector response), large-scale, and packaged with a lightweight access library. We summarize physics coverage and the generation $\rightarrow$ simulation $\rightarrow$ digitization $\rightarrow$ reconstruction pipeline, describe formats/access, and present initial collider physics benchmarks, including both traditional and ML-based.

1 Introduction & Motivation

Collider-based particle physics has long embraced the cutting edge of machine learning (ML), deploying neural networks in production as early as the 1990s [1]. This synergy has driven both fundamental discoveries, such as the Higgs boson [2, 3], and the exclusion of New Physics classes [4]. However, a widening gap exists between the sophisticated "full" simulations available within LHC collaborations—which utilize resource-intensive tools like Geant4 [5] to access low-level detector readouts—and the data available for public research.

While major collaborations internally access realistic low-level objects, public researchers have traditionally relied on "fast simulations" that produce high-level objects directly. Datasets like JetClass, JetNet, and DarkMachines have spurred significant architectural development [6–9], yet they remain limited by high-level abstractions or, in the case of TrackML [10], limited scope (tracker only) and size (10k events). This limitation is critical as the field moves toward "end-to-end" foundation models—such as Maskformer [11, 12] or Higgsformer [13]—which require granular, low-level inputs to bypass traditional reconstruction hierarchies.

To bridge this gap, we present COLLIDERML Release 1 (the OpenDataDetector Physics Benchmark Dataset 2025). **Release 1** provides one million full-simulation events at HL-LHC pile-up conditions, distributed evenly across 7 Standard Model and 3 Beyond Standard Model channels. The simulation maintains full truth across the realistic OpenDataDetector [14], including an inner tracker and both

---

[*]Corresponding author: `daniel.murnane@nbi.ku.dk`
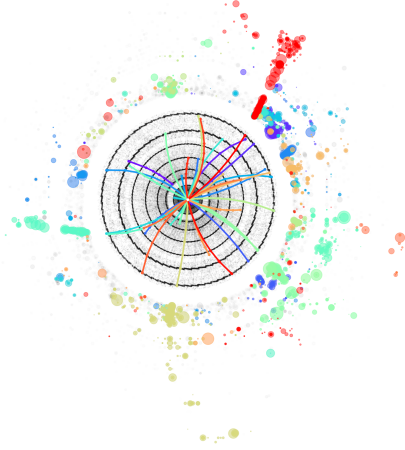
(a) An example COLLIDERMLevent, with inner tracker hits (center), reconstructed tracks, ECal deposits (middle ring) and HCal deposits (outer ring).
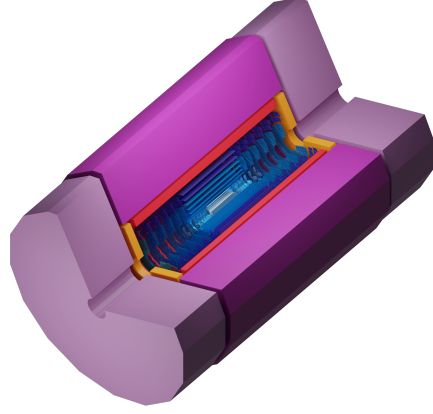
(b) A 3D view of the OpenDataDetector, with (light blue) pixel track, (dark blue) strip tracker, (red and orange) electromagnetic calorimeters, and (purple) hadronic calorimeters.

ID	Label	Process	Notes
1	ttbar	$pp \to t\bar{t}$	Top-quark pair; SM benchmark / BSM background.
2	zee	$pp \to e^+ e^-$	Drell–Yan; electroweak probe.
3	zmumu	$pp \to \mu^+ \mu^-$	Drell–Yan; electroweak probe.
4	diphoton	$pp \to \gamma\gamma$	Loop-induced; bump-hunt candidate.
5	ggf	$gg \to H$	Main Higgs production.
6	dihiggs	$gg \to HH$	Higgs self-coupling probe.
7	multijet	$pp \to 2j, 3j, 4j$	Pure QCD jets; dominant background in hadron colliders.
8	susy	$pp \to \tilde{g}\tilde{g}$	Gluino pair; dark-matter motivated.
9	zprime	$q\bar{q} \to Z'$	Heavy neutral resonance (BSM).
10	hiddenvalley	$pp \to Z' \to v\text{-hadrons}$	Hidden sector; long-lived signatures.

Table 1: Production channels available in COLLIDERML. Each channel contains 100,000 events

electromagnetic and hadronic calorimeters. We provide digitized energy deposits emulating detector response and reconstructed track objects with fitted parameters. Future releases will expand this to include topoclustering, particle flow (PF) constituents via Pandora [15], and clustered jets. The data is released with documentation, a lightweight Python access library, and reproduction recipes.

2 Dataset Description

2.1 Physics Coverage

In **Release 1**, we provide 1 million pp, $\sqrt{s} = 14$TeV collision events, divided equally across 10 physics processes (and a variety of particle gun samples). These channels and some comments on why they were chosen are provided in table 1.

Also included are "particle gun" channels, where single particles are fired at set energy levels in a uniformly sampled initial direction. In these `singleparticle` channels, available are 200,000 each of single μ^-, single e^-, single π^+, single K^+ and single γ.

Hard-scatter matrix element calculation is state-of-the-art, with next-to-leading order (NLO) calculation of the main process and up to two jets, in Madgraph aMC@NLO v3.58 [16]. Loop-induced processes (GGF and di-Higgs), and BSM channels are modelled at LO. Showering is provided with

Pythia v8.313 [17], with FxFx matching for NLO processes, and CKKW matching for LO processes. Known mismodelling in VBF is avoided in the choice of physics processes [18].

Soft QCD scatter ("pile-up") is generated uniquely at LO for every hard scatter event. Re-sampling of pile-up, as is typical in data generation, has been observed to pollute train/test splits in ML algorithms [19], and we propose a wider community investigation of this. Final events are the combination of one hard-scatter and N pile-up sub-events, where N is sampled from a Poisson distribution with μ=200. Sub-event ID is preserved, so that events can be down-sampled to the preferred luminosity.

2.2 Geometry

A core contribution of this work is the validation of the OpenDataDetector (ODD) geometry [20, 21], a project driven by the ACTS [22] collaboration, as an experiment-agnostic way to provide realistic physics simulations in high-luminosity and future collider environments. The ODD combines design choices from existing and future detectors: a wide-coverage silicon inner tracker (IT) inspired by the ATLAS Phase 2 ITk, an electromagnetic calorimeter (ECal) resembling several FCCee detector proposals, and a hadronic calorimeter (HCal) similar to the CMS HGCal subdetector and future detectors CLD, AHCAL and SiD [23–29].

The tracker system is realistically composed of three subsystems: several layers of high resolution pixel modules, two layers of medium-length "strixel" modules, and two layers of strip modules with 2D resolution. Time is accurately modelled and can be included in later digitisation stages. Together, the tracker detector layers provide at least 12 layers of sensitive material across a pseudorapidity of $\eta \in [-3, 3]$, and 8 layers across $\eta \in [-3.5, 3.5]$. A solenoid surrounds the IT, with an inner strength of 3T (a hybrid of ATLAS and CMS conditions) and an outer return field of 0.5T. The outer boundary of the sensitive tracking material is taken to be $r \le 1080$ mm and $z \in [-3030 \text{ mm}, 3030 \text{ mm}]$. This is used to define the truth handling of particles incident on the calorimeter (described in section 2.3).

The electromagnetic calorimeter ("ECal") comprises 48 layers of 0.5 mm thick silicon sensors, each divided into cells of size 5.1 mm × 5.1 mm and interlaid with 1.9 mm tungsten absorption layers. The readout, spacers, and addition material accounts for additional thickness of 2.65 mm per layer. The hadronic calorimeter ("HCal") is composed of 36 layers of 3 mm thick polystyrene scintillator, divided into cells of 30 mm × 30 mm size. The absorber is 30 mm thick steel and the remaining 16 mm corresponds to the readout. A muon detection system is underway for inclusion within the OpenDataDetector geometry.

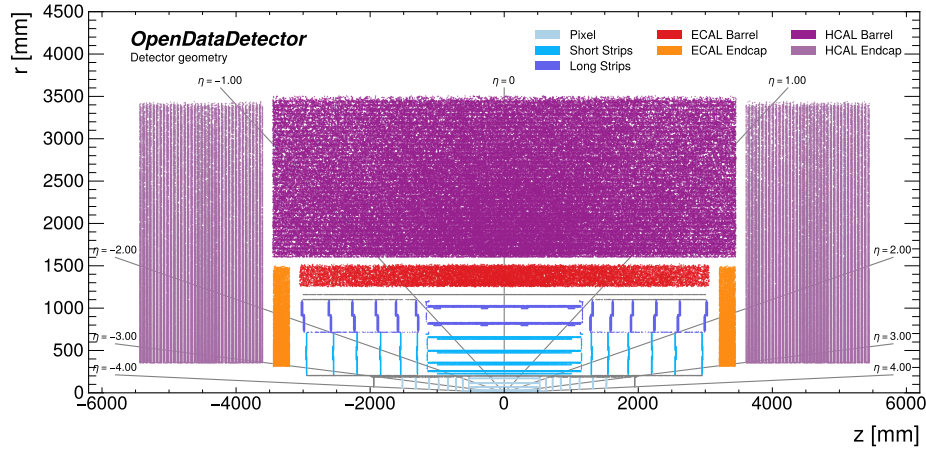## 2.3 Simulation & Digitisation

Figure 2: A longitudinal view of the OpenDataDetector sensitive material, with (light blue) pixel track, (dark blue) strip tracker, (red and orange) electromagnetic calorimeters, and (purple) hadronic calorimeters.

3

Simulations are produced with the pile-up-merged events and ODD geometry, through the DD4hep library [30, 31]. A custom version of the library is used with two major additions. First, a new truth handler is included, to provide users with much more granular information of particle decay chains in the calorimeters. Any particle in the tracker that leaves an energy deposit *or* has energy above 100MeV is preserved. Any particle in the calorimeter that leaves and energy deposit *and* has energy above 100MeV is preserved. A full description is provided in appendix B. Second, experimental multithreading is used, which increases simulation throughput by four-fold, with the same computational resources [32]. Each primary vertex is smeared according to a normal distribution $\mathcal{N} \sim ((\mu_x, \mu_y, \mu_z, \mu_t), (\sigma_x, \sigma_y, \sigma_z, \sigma_t))$, with $\mu_x = \mu_y = \mu_z = \mu_t = 0$, $\sigma_x = \sigma_y = 12.5\mu$m, $\sigma_z = 55.5$mm and $\sigma_t = 185$ps.

All measurements in **Release 1** are digitised with an approximate detector response. A Common Tracking Software (ACTS) [33] is used to digitise the hits registered at truth level by Geant4. Geometric digitisation is implemented by module-wise segmentation of the sensitive material into cells, with low-energy thresholding. Calorimeter digitisation follows the simplified pathway in Key4hep [34] for CLD simulation [35]. That is, every contribution in the calorimeter has a time correction applied based on the time-of-flight from the origin. With this approximate time of production, a window of [-1, 10]ns is applied to all contributions. Those within the window are included in the event, and energies summed per-cell. Total cell energies have an energy threshold of 50keV (250keV) in the electromagnetic (hadronic) calorimeter. The timing and energy thresholds reduce the number of calorimeter hits by a factor of around 4x.

2.4 Reconstruction & Baselines

A key feature of ColliderML **Release 1** is the inclusion of realistically reconstructed track objects. Tracking is performed by the ACTS library, which is the standard solution for many detector collaborations. Tracks are simultaneously found and fit using a Combinatorial Kalman Filter (CKF) applied to a population of triplet seeds from the inner pixel layers [36]. Primary tracks above 1GeV are targeted by the CKF, and ambiguities are resolved to produce the final set of included tracks. The quality of these tracks (efficiencies and resolutions) is similar to those seen in ATLAS and CMS (for example, as seen in fig. 3b).

In the under-development **Release 2** a calorimeter clustering algorithm is performed, analogous to the ATLAS topoclustering approach [37], whereby high-energy cells in each layer seed clusters and neighbors are consumed according to a simple set of heuristics. Orthogonally, the Pandora Particle Flow algorithm is applied to produce particle flow objects (PFOs). This algorithm uses sophisticated matching of tracks and calorimeter deposits by extrapolating beyond the inner tracker. Topoclusters and PFOs may be used for complimentary studies to understand their value for downstream tasks. One such task is jet reconstruction, which is to be provided in **Release 2**. Anti-k_t clustering in FastJet [38] is applied to PFOs and simple heuristics are applied to provide a baseline performance on flavour classification, isolation, and energy regression of each jet. While COLLIDERMLhas "only" 1 million events, each event may contain many jets, and these can be regionalised[2] to produce a much larger full-simulation jet-based dataset.

3 ML Use-Cases

The ColliderML Benchmark dataset is well-suited for ML algorithmic studies. Release 1 provides all detector-level objects, as well as track objects. These can be used for comparisons between traditional and ML-based track reconstruction. fig. 4a shows the track finding efficiency for GNN-based tracking [19, 40, 41]. The inclusion of timing data is also explored, and can be seen to boost performance, motivating the development of precise time detection in future detectors [42, 43]. We also show the ability of a transformer trained on track, jet and event-level ("analysis") objects for top-antitop vs. gluon-gluon fusion Higgs event tagging at zero pile-up, fig. 4b. Each object type performs similarly alone, but combined the "multiscale" tagger increases BR by a factor of 1.5 over analysis-level. The finding mirrors that seen in recent work by ATLAS and CMS that the inclusion of lower-level objects improves downstream tagging tasks [44, 45].

---

[2]We note that the regionalisation of jets is non-trivial, and may need to be handled carefully for certain topologies. For now, this is left to users to implement for their particular use-case.
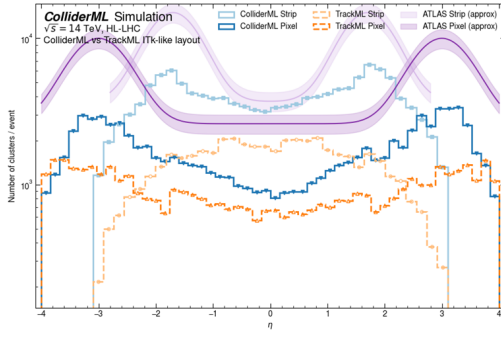
(a) Number of inner tracker deposits across ColliderML (blue), TrackML (orange) and ATLAS (purple - approximate)
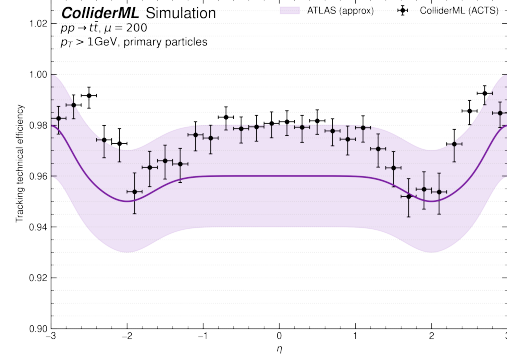

(b) ACTS tracking efficiency vs. η (black) and approximate ATLAS performance (purple, adapted from [39]).

A key feature of COLLIDERML that is not typically included in such datasets is the retention of the truth-level primary vertex of all particles. We propose this as a mechanism to study scaling behaviour: Events can easily be downsampled to the hard-scatter event (primary vertex 1) and a pile-up of n (primary vertices $2, ..., n + 1$). We suggest that this scaling from "easy" to "hard" environments should be the de facto approach for ML studies (for example replacing artificial cuts on low energy particles and energy deposits). One can downsample with the ColliderML library, which simply removes those contributions from tracker and calorimeters from dropped pile-up. A re-digisation procedure is necessary, which for now follows the simple energy thresholding effect described above.
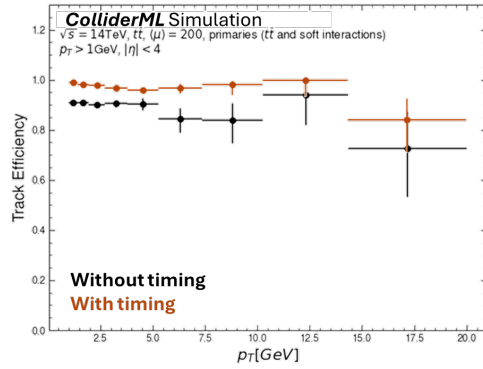
(a) Track reconstruction performance of the GNN4ITk pipeline, with and without smeared timing feature included.
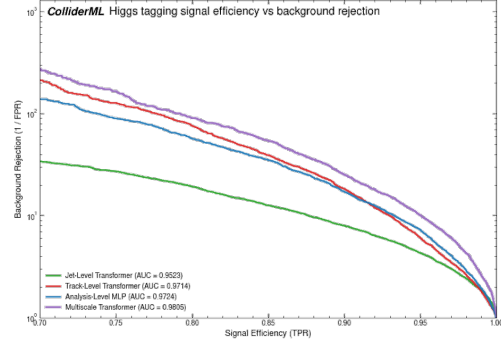
(b) The background rejection vs. signal efficiency of a ttbar vs. GGF Higgs classifying transformer, applied to each scale of data, or all three ("multiscale").

4 Data Format & Access

Release 1 provides two formats for end-users. First, full-truth data in intermediate and final stages is structured in EDM4hep ROOT files [46], which are augmented with extra features where necessary using the Podio interface [47]. This requires approximately 300TB to preserve. In order to make this collection maximally useful for machine learning experiments, we process this into a set of four parquet files: `particles` (table 2), `tracker_hits` (table 3), `calo_hits` (table 4) and `tracks` (table 5). This dataset is approximately 30TB, and made available in full on HuggingFace, at `https://huggingface.co/OpenDataDetector`. The format of the parquet files closely mirrors that of the EDM4hep, with simplified linking, and many unnecessary branches removed, described in appendix A. A lightweight PyPi library is provided to inspect the various physics objects available in Release 1, and produce some simple data augmentations and manipulations.

5

A Parquet formats

Field	Type	Description
event_id	int32	Unique event identifier
particle_id	list<uint64>	Unique particle ID within event
pdg_id	list<int64>	PDG particle code
mass	list<float32>	Particle rest mass (GeV/c²)
energy	list<float32>	Particle total energy (GeV)
charge	list<float32>	Electric charge (in units of e)
px, py, pz	list<float32>	Momentum components (GeV/c)
vx, vy, vz	list<float32>	Vertex position (mm)
time	list<float32>	Production time (ns)
num_tracker_hits	list<uint16>	Number of hits in tracker
num_calo_hits	list<uint16>	Number of hits in calorimeter
vertex_primary	list<uint16>	Primary vertex ID (1=hard scatter, [2,...,N]=pile-up)
perigee_d0	list<float32>	Radial impact parameter of the particle's fitted perigee
perigee_z0	list<float32>	Longitudinal impact parameter of the particle's fitted perigee
parent_id	list<int64>	ID of parent particle

Table 2: Particle data structure. Note that parent_id is given as -1 in case the particle has no parent.

Field	Type	Description
event_id	uint32	Unique event identifier
x, y, z	list<float32>	Measured hit position (mm)
true_x, true_y, true_z	list<float32>	True hit position before digitization (mm)
time	list<float32>	Hit time (ns)
particle_id	list<uint64>	Truth particle that created this hit
volume_id	list<uint8>	Detector volume identifier
layer_id	list<uint16>	Detector layer number
surface_id	list<uint32>	Sensor surface identifier
detector	list<uint8>	Detector subsystem code

Table 3: Tracker hits data structure

Field	Type	Description
event_id	uint32	Unique event identifier
detector	list<uint8>	Calorimeter subsystem name
cell_id	list<string>	Calorimeter cell identifier
total_energy	list<float32>	Total energy deposited in cell (GeV)
x, y, z	list<float32>	Cell center position (mm)
contrib_particle_ids	list<list<uint64»	IDs of particles contributing to this cell
contrib_energies	list<list<float32»	Energy contribution from each particle (GeV)
contrib_times	list<list<float32»	Time of each contribution (ns)

Table 4: Calorimeter hits data. Note the nested lists for contributions.

Field	Type	Description
event_id	uint32	Unique event identifier

Field	Type	Description
track_id	list<uint16>	Unique track identifier within event
majority_particle_id	list<uint64>	Truth particle with most hits on this track
d0	list<float32>	Transverse impact parameter (mm)
z0	list<float32>	Longitudinal impact parameter (mm)
phi	list<float32>	Azimuthal angle (radians)
theta	list<float32>	Polar angle (radians)
qop	list<float32>	Charge divided by momentum (e/GeV)
hit_ids	list<list<uint32»	List of tracker hit IDs assigned to this track

Table 5: Track data structure. Track parameters are those fitted to a helix by ACTS

B   Truth Handling

To enable sophisticated studies of particle decay reconstruction or simulation, we enforce very conservative truth handling. The DD4hep algorithm proceeds as: In a particle decay/interaction, should we retain the daughter information? If yes, then the daughter is registered as a particle with all its details entering the particle record and all its energy deposits attributed to it. If no, then the daughter is marked to be discarded, and its energy deposits are attributed to its parent. This attribution occurs after the simulation, and thus we may have the case where a to-be-discarded daughter itself produces a to-be-retained grand-daughter. As expected then, the daughter's hits are attributed to its parent, while the grand-daughter's hits are attributed to itself, and so on. Given this framework, the choice of DD4hep truth handling amounts to: What are the conditions in which a particle is discarded or retained?

The default truth handling is shown in fig. 5a. While this leads to compact particle truth records, and is likely suitable for the majority of standard analysis use-cases, it leads to a variety of behaviour that is non-ideal. For example, backscatter from calorimeter into the tracker is not correctly handled. Additionally, tracking can become confusing in this environment since all low energy particles are assigned to their parent, and thus their hits will appear to create tree-like tracks in the tracker. Finally, all calorimeter decay information is lost.
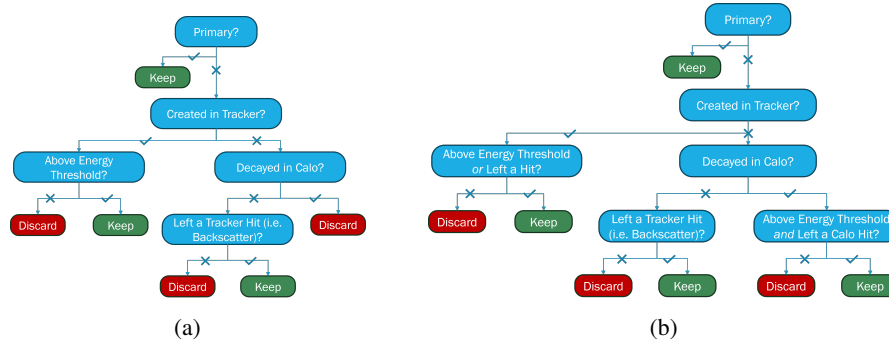

Figure 5: Truth handling logical flow. a) The default DD4hep truth handling. b) The custom COLLIDERMLhandling, allowing more detailed tracking of particle information through the detector

References

[1] L. Lönnblad, C. Peterson, and T. Rögnvaldsson. Using neural networks to identify jets. *Nucl. Phys. B*, 349:675–702, 1991. doi: 10.1016/0550-3213(91)90190-V.

[2] ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Phys. Lett. B*, 716:1–29, 2012. doi: 10.1016/j.physletb. 2012.08.020.

[3] CMS Collaboration. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Phys. Lett. B*, 716:30–61, 2012. doi: 10.1016/j.physletb.2012.08.021.

[4] ATLAS Collaboration. Summary of searches for squarks and gluinos as realised in run 2 of the lhc with the atlas detector. Technical report, CERN, 2021. URL `https://cds.cern.ch/record/2772288`.

[5] S. Agostinelli et al. Geant4—a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250–303, 2003. doi: 10.1016/S0168-9002(03)01368-8.

[6] Huilin Qu, Congqiao Li, and Sitian Qian. Particle transformer for jet tagging. *arXiv preprint*, 2022. Presents the 100M-jet *JetClass* dataset.

[7] Raghav Kansal et al. Particle cloud generation with message passing generative adversarial networks. *NeurIPS*, 2021. Introduces the JetNet dataset.

[8] T. Aarrestad et al. The dark machines anomaly score challenge: Benchmark data and model independent event classification for the large hadron collider. *SciPost Phys.*, 12:043, 2022. doi: 10.21468/SciPostPhys.12.1.043.

[9] G. Kasieczka et al. The lhc olympics 2020: A community challenge for anomaly detection in high energy physics. *Rept. Prog. Phys.*, 84(12):124201, 2021. doi: 10.1088/1361-6633/ac36b9.

[10] S. Amrouche et al. The tracking machine learning challenge: Accuracy phase. *arXiv preprint*, 2019.

[11] J. Gong et al. Maskformers for particle physics: End-to-end secondary vertex reconstruction. *Phys. Rev. D*, 108:076025, 2023.

[12] Samuel Van Stroud, Philippa Duckett, Max Hart, Nikita Pond, Sébastien Rettie, Gabriel Facini, and Tim Scanlon. Transformers for charged particle track reconstruction in high energy physics, 2024. URL `https://arxiv.org/abs/2411.07149`.

[13] F. Zhou et al. Higgsformer: Transformer-based end-to-end search for higgs pair production from raw detector hits. *arXiv*, 2025. Preprint, 2025-08-26.

[14] P. Gessinger-Befurt et al. Opendatadetector tracking system: Design and prototyping. In *IEEE NSS/MIC 2023*, 2023.

[15] J. S. Marshall and M. A. Thomson. Pandora particle flow algorithm. *Eur. Phys. J. C*, 75(439), 2015. doi: 10.1140/epjc/s10052-015-3659-3.

[16] J. Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014. doi: 10.1007/JHEP07(2014)079.

[17] T. Sjöstrand et al. An introduction to pythia 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015. doi: 10.1016/j.cpc.2015.01.024.

[18] S. Höche et al. Vector-boson fusion at next-to-leading order qcd with parton showers. *SciPost Phys.*, 12:091, 2022.

[19] ATLAS Collaboration. Atlas itk track reconstruction with a gnn-based pipeline. ATL-ITK-PROC-2022-006, 2022. URL `https://cds.cern.ch/record/2816063`.

[20] OpenDataDetector Team. Opendatadetector (odd) geometry and docs. `https://opendatadetector.gitlab.io/`. Accessed 2025-08-30.

[21] P. Gessinger-Befurt et al. The open data detector tracking system. Instrumentation Days (IN2P3) 2023, slides, 2023. URL `https://indico.in2p3.fr/event/28938/contributions/123916/`.

[22] Xiaocong Ai, Corentin Allaire, Noemi Calace, Angéla Czirkos, Markus Elsing, Irina Ene, Ralf Farkas, Louis-Guillaume Gagnon, Rocky Garg, Paul Gessinger, Hadrien Grasland, Heather M. Gray, Christian Gumpert, Julia Hrdinka, Benjamin Huth, Moritz Kiehn, Fabian Klimpel, Bernadette Kolbinger, Attila Krasznahorkay, Robert Langenberg, Charles Leggett, Georgiana Mania, Edward Moyse, Joana Niermann, Joseph D. Osborn, David Rousseau, Andreas

Salzburger, Bastian Schlag, Lauren Tompkins, Tomohiro Yamazaki, Beomki Yeo, and Jin Zhang. A common tracking software project. *Computing and Software for Big Science*, 6(1):8, 2022. doi: 10.1007/s41781-021-00078-8. URL https://doi.org/10.1007/s41781-021-00078-8.

[23] ATLAS Collaboration. Technical design report for the atlas inner tracker pixel detector. Technical Report CERN-LHCC-2017-021, ATLAS-TDR-030, CERN, 2017. URL https://cds.cern.ch/record/2285585.

[24] ATLAS Collaboration. Technical design report for the atlas inner tracker strip detector. Technical Report CERN-LHCC-2017-005, ATLAS-TDR-025, CERN, 2017. URL https://inspirehep.net/literature/1614102.

[25] CLICdp/FCC-ee detector working groups. Cld — a detector concept for fcc-ee. https://edms.cern.ch/document/2390058/1, 2019. CERN-LCD-2019-001.

[26] CMS Collaboration. The phase-2 upgrade of the cms endcap calorimeter. Technical Report CERN-LHCC-2017-023, CMS-TDR-019, CERN, 2017. URL https://cds.cern.ch/record/2293646.

[27] M. Bacchetta et al. Cld — a detector concept for the fcc-ee. *CERN Report*, 2019. arXiv:1911.12230.

[28] C. Adloff et al. Construction and commissioning of the CALICE analog hadron calorimeter prototype. *JINST*, 5:P05004, 2010. doi: 10.1088/1748-0221/5/05/P05004.

[29] H. Aihara, P. Burrows, M. Oreglia, et al. Sid letter of intent. *arXiv*, 2009.

[30] M. Frank, F. Gaede, M. Petric, and A. Sailer. DD4hep: toolkit for detector description. In *J. Phys. Conf. Ser.*, volume 513, page 022010, 2014. doi: 10.1088/1742-6596/513/2/022010.

[31] F. Gaede et al. The dd4hep detector description toolkit. *EPJ Web Conf.*, 245:02004, 2020. doi: 10.1051/epjconf/202024502004.

[32] wdconinc. DD4hep. https://github.com/wdconinc/DD4hep, 2021. GitHub Repository, version v01-17-00, accessed 2025.

[33] ACTS Collaboration. A common tracking software (acts). *EPJ Web Conf.*, 245:02028, 2020. doi: 10.1051/epjconf/202024502028.

[34] M. Brondolin, D. R. Ferreira, E. Sicking, A. Sailer, et al. The Key4HEP software stack: Recent progress. In *EPJ Web Conf.*, volume 295, page 05010, 2024. doi: 10.1051/epjconf/202429505010.

[35] Key4HEP Collaboration. k4detperformance: Cld/key4hep reconstruction and digitisation examples. https://github.com/key4hep/k4DetPerformance, 2023. Accessed 2025-08-30.

[36] J. Gao et al. Track reconstruction with the acts combinatorial kalman filter and seeding. In *STCF 2023*, 2023.

[37] ATLAS Collaboration. Topological cell clustering in the atlas calorimeters and its performance in lhc run 1. *Eur. Phys. J. C*, 77:490, 2017. doi: 10.1140/epjc/s10052-017-5004-5.

[38] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. Fastjet user manual: (for version 3.0.2). *The European Physical Journal C*, 72(3), March 2012. ISSN 1434-6052. doi: 10.1140/epjc/s10052-012-1896-2. URL http://dx.doi.org/10.1140/epjc/s10052-012-1896-2.

[39] G. et al Aad. Expected tracking performance of the atlas inner tracker at the high-luminosity lhc. *Journal of Instrumentation*, 20(02):P02018, February 2025. ISSN 1748-0221. doi: 10.1088/1748-0221/20/02/p02018. URL http://dx.doi.org/10.1088/1748-0221/20/02/P02018.

[40] X. Ju et al. Graph neural networks for particle tracking and reconstruction. In *ICLR 2021 Workshop on AI for Science*, 2021.

[41] J. Duarte, H. J. Chang, S. Hsu, et al. End-to-end particle tracking and reconstruction with gnns at the hl-lhc. *arXiv preprint*, 2022.

[42] ATLAS Collaboration. Technical design report: A high-granularity timing detector (hgtd) for the atlas phase-ii upgrade. Technical Report CERN-LHCC-2018-023, ATLAS-TDR-031, CERN, 2018. URL `https://cds.cern.ch/record/2623663`.

[43] ATLAS Collaboration. Technical design report for the atlas high-granularity timing detector (hgtd). Technical Report CERN-LHCC-2020-007; ATLAS-TDR-031, CERN, 2020. URL `https://cds.cern.ch/record/2719855`.

[44] ATLAS Collaboration. Transforming jet flavour tagging at atlas, 2025. URL `https://arxiv.org/abs/2505.19689`.

[45] A unified approach for jet tagging in Run 3 at $\sqrt{s}$=13.6 TeV in CMS. 2024. URL `https://cds.cern.ch/record/2904702`.

[46] F. Gaede, P. Mato, G. Ganis, A. Sailer, et al. Edm4hep: A common event data model for hep. In *EPJ Web Conf.*, volume 251, page 03026, 2021. doi: 10.1051/epjconf/202125103026.

[47] A. Carceller, T. Madigan, F. Gaede, A. Sailer, et al. Towards podio v1.0. In *EPJ Web Conf.*, volume 295, page 06018, 2024. doi: 10.1051/epjconf/202429506018.