

Interpretable unsupervised representation learning for high-precision measurements in particle physics

Xing-Jian Lv^{1,2}, De-Xing Miao^{1,2,*}, Zi-Jun Xu¹, and Jian-Chun Wang¹

¹*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China*

²*University of Chinese Academy of Sciences, Beijing 100049, China*



(Received 22 December 2025; accepted 11 June 2026; published 8 July 2026)

Unsupervised learning has been widely applied to various tasks in particle physics. However, existing models lack precise control over their learned representations, limiting physical interpretability and hindering their use for accurate measurements. We propose the histogram autoencoder (HistoAE), an unsupervised representation learning network featuring a custom histogram-based loss that enforces a physically structured latent space. Applied to silicon microstrip detectors, HistoAE learns an interpretable two-dimensional latent space corresponding to the particle's charge and impact position. After simple postprocessing, it achieves a charge resolution of $0.25e$ and a position resolution of $3\text{ }\mu\text{m}$ on beam-test data, comparable to the conventional approach. These results demonstrate that unsupervised deep learning models can enable physically meaningful and quantitatively precise measurements. Moreover, the generative capacity of HistoAE enables straightforward extensions to fast detector simulations.

DOI: 10.1103/35nc-8xbk

I. INTRODUCTION

Machine learning, and in particular its modern incarnation of deep learning (DL) [1,2], has become an indispensable tool in particle physics, a field that routinely handles vast datasets and nonlinear relationships among observables [3–6]. In recent years, advances in DL have expanded the scope of data-driven progress across the energy, intensity, accelerator, and cosmic frontiers [7,8]. Despite remarkable advancements, most current DL applications in particle physics are supervised, relying either on Monte Carlo (MC) simulations or on labeled experimental data. However, because simulations cannot fully capture the complexity of the real world, a persistent gap between MC simulations and data leads to training bias. Direct training on real data, in turn, demands truth labels for every event. In practice, such labels must be derived from auxiliary detectors, whose raw measurements cannot serve directly as training targets but require elaborate, detector-specific calibration and manual processing to convert into usable labels [9–11]. This labeling pipeline is difficult to generalize and scale. For this reason, the development of unsupervised DL [12,13] is integral for particle physics.

Unsupervised learning has achieved remarkable success in tasks such as clustering and identifying topological phases of matter [14,15], anomaly detection [16,17], and learning representations that capture essential physical degrees of freedom [18,19]. These results demonstrate that unsupervised methods can discover physically meaningful structures from raw data without explicit labels. However, measurements in particle physics demand quantitative precision and interpretability. To date, unsupervised models lack explicit control over their learned representations [20], rendering them unsuitable for direct physical measurement [21].

In this work, we present a fully unsupervised DL approach capable of performing precise measurements. We focus on the charge and position measurements of particles (nuclei) using silicon microstrip detectors (SSDs) [22,23], which are among the most fundamental physical observables in many areas of particle physics, such as cosmic-ray composition and rigidity measurements in spaceborne experiments [24] and final-state particle identification and track reconstruction in heavy-ion experiments [25].

The detector response exhibits a complex and nonlinear dependence on the charge and impact position of the incident particle. Consequently, conventional charge and position reconstruction methods rely on elaborate calibration procedures and prior knowledge of particle species [10,11]. Such priors are typically derived from auxiliary detectors, introducing considerable detection inefficiency and potential bias. In particular, conventional charge reconstruction requires prior species identification from other subdetectors, creating a circular dependency between charge measurement and particle identification. Moreover,

*Contact author: miaodx@ihep.ac.cn

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.

the calibration itself relies on simulation-derived correction functions, so any mismodeling of the detector response propagates directly into the measured charge. An unsupervised method that learns directly from the raw detector signals removes both dependencies, enabling stand-alone reconstruction without auxiliary detectors or simulation-tuned calibrations. Furthermore, no conventional SSD reconstruction algorithm simultaneously determines both charge and position, because the respective calibrations are performed independently with different auxiliary inputs. While supervised multitask neural networks can in principle perform such joint regression [26], they require labeled training data derived from simulation or auxiliary detectors.

Autoencoders (AEs) [27,28] are a widely used framework for unsupervised learning. An AE consists of an encoder that compresses input data into a latent representation and a decoder that reconstructs the original features. By training to minimize reconstruction error, AEs learn compact representations of the data. However, standard AEs lack explicit control over the latent distribution, rendering the learned representations not fully interpretable and thus unsuitable for precise measurements. Previous extensions, such as variational autoencoders (VAEs) [29,30] and Wasserstein autoencoders (WAEs) [31], incorporate regularization terms to impose global constraints on latent distributions, thereby improving overall controllability. Still, there is a lack of fine-grained control over the internal structure of the latent space.

To overcome this limitation, we introduce a novel network architecture, the histogram autoencoder (HistoAE), which employs a custom loss function based on the histogram of the latent space. HistoAE enables explicit control over the latent-space distribution, leading to a fully interpretable latent representation. With this network, we achieve highly precise, unsupervised, and simultaneous reconstruction of particle charge and position.

These results demonstrate that a deep learning model can deliver physically meaningful and quantitatively accurate measurements in a completely unsupervised setting. Beyond its application to SSD, our approach provides a general framework for interpretable, label-free analysis of high-dimensional data.

This article is organized as follows. Section II introduces the basics of SSDs and the training and testing dataset for this work. Section III presents the HistoAE framework in detail. Section IV reports the main results, and Sec. V provides a summary.

II. DATA SAMPLE

A. Silicon microstrip detector for particle measurement

SSDs are extensively used for precise position and charge measurements of ionizing particles across several fields in collider experiments [32–34] and nuclear physics [35], as well as in many spaceborne experiments [36–38] including the Alpha Magnetic Spectrometer (AMS-02) [39,40].

The SSD used in this work was designed for the layer-0 tracking detector upgrade of the AMS-02 experiment, aiming to triple the statistical sample and enhance the nuclei identification capability [41]. The layer-0 detector is scheduled to be launched to the International Space Station in 2027.

The SSD consists of a large number of parallel, equally spaced readout strips fabricated on a silicon substrate. When a charged particle traverses the detector, it deposits energy along its path through ionization, generating electron-hole pairs in the silicon bulk. Under an applied electric field, charge carriers drift toward the nearest readout strips, causing the signals of several adjacent channels to exceed the threshold. These adjacent channels collectively form what is referred to as a “cluster.” According to the Bethe-Bloch formula [42], the total deposited energy is proportional to the square of the particle’s charge, Z^2 . However, the signal amplitudes of individual channels within a cluster exhibit a much more complex behavior.

First, due to the nonlinearity and saturation effects of the readout electronics, the signal amplitude recorded by each channel is not strictly proportional to the amount of charge collected by the corresponding readout strip. Moreover, the charge-collection efficiency of the SSD depends on the impact position of the particle: when the particle hits closer to a readout strip, the efficiency is higher. Consequently, even for the same particle species, clusters produced at different impact positions can exhibit different total amplitudes and shapes. In addition, the presence of δ rays leads to a non-Gaussian signal distribution for individual particles [43]. Figure 1 shows that the signal amplitude of the maximum channel within a cluster, hereafter referred to as the seed

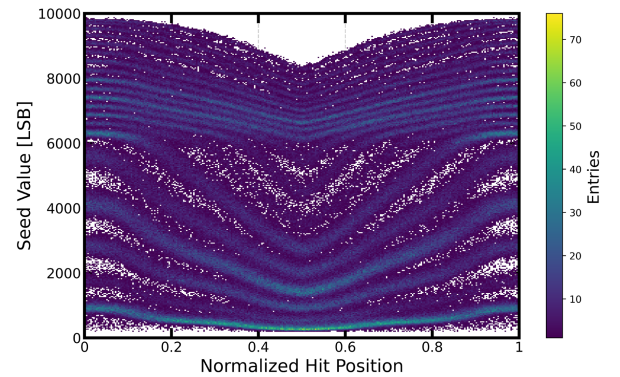


FIG. 1. Distribution of the maximum channel signal within a cluster (referred to as the seed value) as a function of the normalized impact position for mixed-nuclei beam-test events. The x-axis represents the impact position normalized to the pitch between two readout strips, where a value of 0.5 denotes the midpoint between adjacent strips. The y-axis gives the maximum channel signal in units of the digitized electronics readout (LSB, least significant bit). The horizontal bands from bottom to top correspond to different nuclear species, with the lowest visible band corresponding to $Z = 3$ after hydrogen and helium events are removed.

value, varies with the particle's impact position, normalized to the interval between two readout strips. The signal is expressed in LSB, the least significant bit of the digitized electronics readout. Each band corresponds to a specific species; in the figure shown here, the lowest visible band starts at $Z = 3$ because hydrogen and helium events are removed. The shapes of these bands differ significantly and exhibit complex nonlinear behavior. This complexity makes it challenging to reconstruct the particle's charge and impact position from the cluster signals.

B. High-energy nuclei beam test

In this work, all training and testing data are real data obtained from beam-test measurements, which were carried out at CERN Super Proton Synchrotron. The beam was generated by bombarding a lead primary beam of 150 GeV/n on a beryllium target. The resulting secondary fragmented beam consisted of a mixture of various nuclei and was selected by the magnetic optics with a mass-to-charge ratio of $A/Z = 2$, while the beam momentum was controlled by the magnet system with a precision of 0.1% [44]. A total of 5×10^6 data samples were collected for training and testing.

We have a total of nine SSD layers with identical designs. Among them, eight layers serve as a beam telescope, providing reference measurements of the particle charge and impact position, while the central layer is the detector under test (DUT). The “truth” position used in this study refers to the predicted impact point derived from the telescope reconstructed track. The truth charge is defined as the average charge value obtained from independent measurements of the same particle track in different telescope layers. The results presented for the network training and testing are the DUT's completely independent measurements of the same particle tracks. We emphasize that the truth charge and position from the telescope are used exclusively for post-training evaluation and are never provided to the network during training. The HistoAE model is trained in a purely unsupervised manner using only the raw DUT cluster signals as both input and reconstruction target.

III. MODEL ARCHITECTURE

A. Network design

HistoAE adopts a deep neural network architecture composed of fully connected layers for both the encoder and decoder. The model is designed to learn compact, physically interpretable latent representations of detector responses, corresponding to the impact position and charge of incident particles. The architecture of HistoAE is shown in Fig. 2.

B. Input representation

For each event, we use the readout signals of the five channels with the highest amplitudes within the cluster associated with the incident particle. The raw signals span

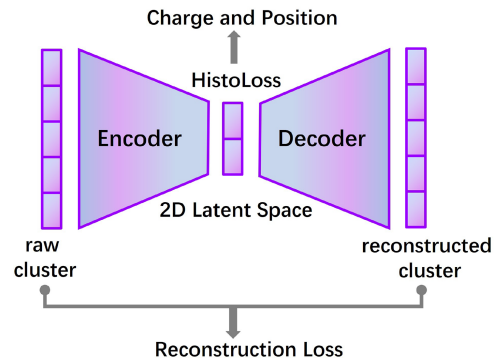


FIG. 2. The network architecture of HistoAE. It consists of an encoder and a decoder connected through a two-dimensional latent space. A reconstruction loss ensures that the reconstructed clusters closely match the original ones, while the HistoLoss constrains the two latent dimensions to represent charge and position, respectively.

several orders of magnitude—from $\mathcal{O}(1)$ to $\mathcal{O}(10^4)$ —owing to variations in impact position and particle charge. Direct application of standard normalization techniques, such as min-max or mean-variance scaling, fails to alleviate this dynamic range problem, since they preserve the multiplicative scale differences between large and small signals. Logarithmic or power transformations can compress the range but distort the absolute differences between signal amplitudes from adjacent nuclei, thereby hindering the learning of relative charge patterns.

To address this, we introduce a digitwise vector encoding scheme, referred to as vecoding. Each scalar input value is decomposed into its individual decimal digits and mapped to a fixed-length vector representation. For instance, a value of 9764.4 is transformed into the vector $[0, 9, 7, 6, 4, 4]$, corresponding to a six-dimensional encoding with one decimal place retained. This approach ensures that every element entering the network lies within the range $[0, 9]$, thereby maintaining numerical stability while preserving the intrinsic structure of the original signals. Crucially, this transformation does not alter the relative or absolute information content of the input data. Related digit-level encoding strategies have appeared in machine learning models for mathematical reasoning [45] and in continuous numerical tokenization for scientific data [46]; our vecoding is motivated by the specific dynamic range requirements of physical detector signals.

C. Encoder and decoder design

The input vectors from the five channels are flattened into a single 30-dimensional vector (5×6) before being fed into the encoder. The encoder consists of a stack of fully connected layers with intermediate dimensions of 256, 128, 64, and 32, each followed by batch normalization and sigmoid linear unit activation functions. The final encoder layer projects the high-dimensional input into a

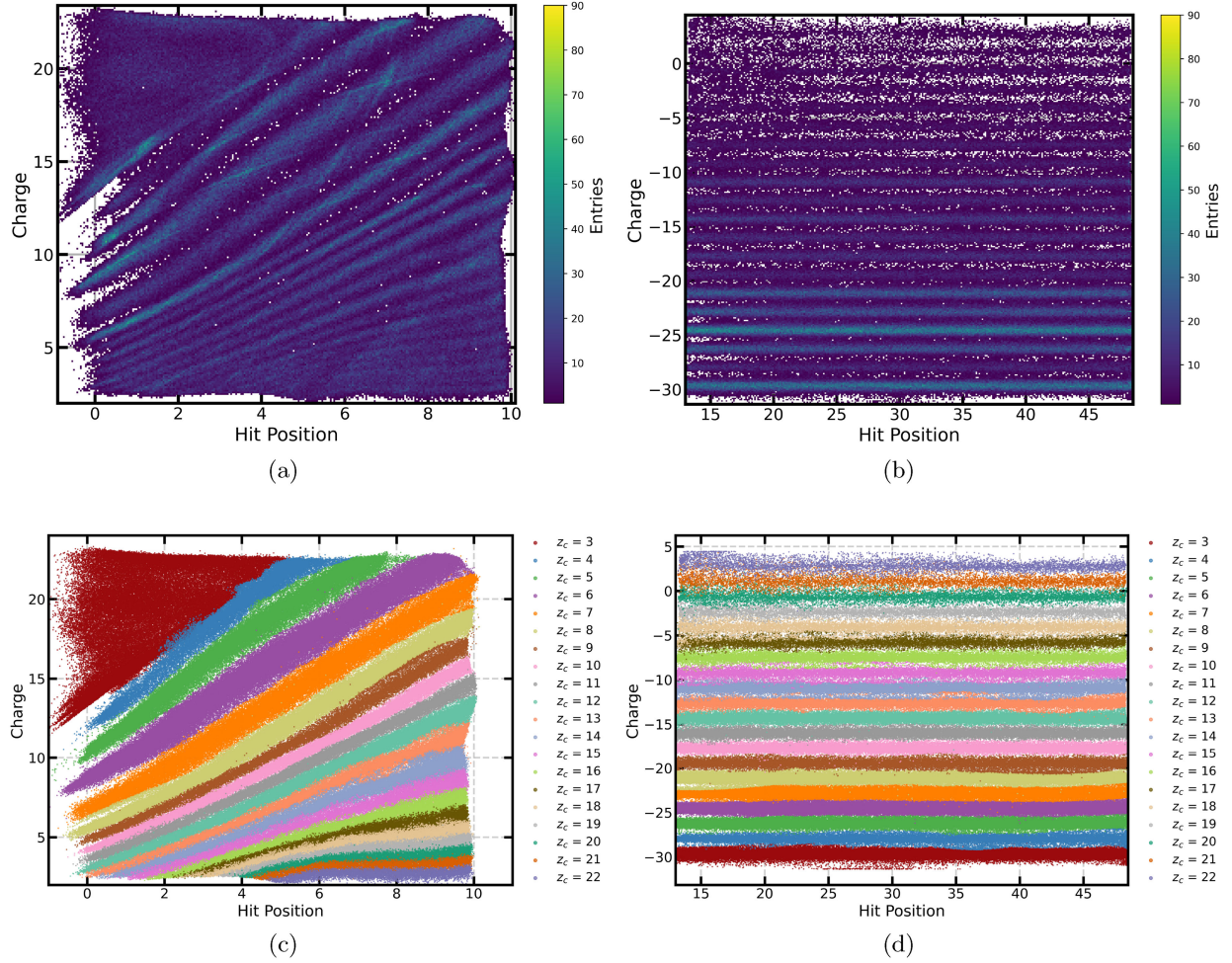


FIG. 3. Latent-space visualizations. (a),(b) Two-dimensional histograms of the latent variables obtained with the WAE and HistoAE, respectively. (c),(d) The corresponding latent spaces color coded by the true nuclear charge, revealing overlapping diagonal bands for the WAE and well-separated horizontal parallel structures for the HistoAE.

two-dimensional latent space, representing the particle's impact position and charge. The decoder mirrors the encoder with hidden widths 32, 64, 128, and 256 in reverse order to reconstruct the input readout patterns from the latent representation. The use of batch normalization in all hidden layers stabilizes training and improves convergence in the presence of large dynamic ranges [47].

1. HistoLoss

Physically, a cluster is fully determined by the impact position and the particle charge. In principle, any linear or nonlinear combination of these two variables would allow the decoder to reconstruct the observed cluster. However, only a properly disentangled latent representation can be interpreted in physically meaningful terms. Therefore, a central objective of our approach is to achieve accurate control of the latent-space distribution.

In the broader machine learning community, several frameworks have been proposed to impose global constraints

on the latent distribution, such as VAE and WAE. However, when trained on our data, the resulting latent distributions typically exhibit bandlike structures—each band corresponding to a distinct particle species—but these bands are often curved and irregular, lacking clear physical interpretation along either latent dimension as shown in Figs. 3(a) and 3(c). This arises because such models regularize only the overall support of the latent space [20], without providing the fine-grained control needed for quantitative measurements in our problem, where signal distributions from different particle species intrinsically overlap due to finite resolution.

To explicitly regulate the geometry of the latent space, we introduce a loss function based directly on a two-dimensional histogram, referred to as HistoLoss. A target latent distribution is first defined, and the histogram of the learned latent variables is computed. The L1 loss between the two histograms is then minimized to align the learned distribution with the target.

In our case, the incident particles carry integer charges. After accounting for detector resolution, the charge

dimension is expected to follow a Gaussian mixture model distribution. The incident position, normalized to lie between two adjacent channels, can be regarded as uniformly distributed. Since the charge and position are independent, the target latent distribution can therefore be expressed as a function of the charge coordinate q and the position coordinate x ,

$$H_{\text{target}}(q, x) = \sum_{k=1}^K w_k \mathcal{N}(q | \mu_k, \sigma_k^2) \frac{1}{x_{\text{max}} - x_{\text{min}}},$$

where μ_k , σ_k^2 , and w_k are the mean, variance, and normalized weight of each charge component, respectively. $H_{\text{target}}(q, x)$ represents the target probability density at charge coordinate q and position coordinate x . We stress that H_{target} is constructed entirely from generic physical priors: incident particles carry integer charges, which after detector-resolution broadening produce a Gaussian mixture, and the normalized interstrip impact position is approximately uniform. Here, K denotes the number of charge components retained in the selected beam sample, and w_k represents their relative abundances in the mixed beam. The parameters μ_k correspond to the integer charge values. The σ_k parameters provide an estimate of the intrinsic detector resolution, which is reflected in the relative widths of the bands in Fig. 1. These parameter values are used only for initialization. After 5000 epochs, they are released and allowed to be learned by the model, enabling them to adapt to the characteristics of the data themselves.

Denote z_q and z_x as the latent variables corresponding to charge and position, respectively. For the latent-space distribution, let $\{(z_{q,i}, z_{x,i})\}_{i=1}^N$ represent the latent variables for N events in a minibatch. We first normalize $z_{q,i}$ and $z_{x,i}$ to the range $[0, 1]$ and compute their two-dimensional histogram. To enable backpropagation, the histogram is smoothed using a Gaussian kernel, yielding the empirical latent density $H_{\text{emp}}(q, x)$,

$$H_{\text{emp}}(q, x) = \frac{1}{N} \sum_{i=1}^N \mathcal{K}_{\sigma}(q - z_{q,i}) \mathcal{K}_{\sigma}(x - z_{x,i}),$$

where $\mathcal{K}_{\sigma}$ is a Gaussian kernel with bandwidth σ .

The histogram loss is computed as the elementwise L_1 distance between the empirical and target densities,

$$\mathcal{L}_{\text{histo}} = \|H_{\text{emp}}(q, x) - H_{\text{target}}(q, x)\|_1,$$

evaluated on the binned latent grid.

The total training objective combines this histogram loss with the standard reconstruction loss,

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{recon}} + \lambda_{\text{histo}} \mathcal{L}_{\text{histo}},$$

where λ_{histo} controls the strength of the latent-space regularization.

2. Training details

All model implementation and training were performed using the PyTorch framework [48]. The network was trained end to end using the AdamW optimizer [49] with a weight decay of 10^{-4} and a base learning rate of 10^{-3} . All weights were initialized following the Kaiming scheme [50] to ensure stable convergence. A linear warm-up [51] of the learning rate was applied during the first 10 epochs, followed by a cosine annealing schedule [52] that gradually decreased the rate to 10^{-5} . Mixed-precision training with automatic gradient scaling [53] was employed to enhance numerical stability and computational efficiency on GPUs.

The dataset was randomly split into 80% for training and 20% for testing. Training was conducted for up to 20 000 epochs. To improve convergence and ensure physical consistency of the latent representation, we adopted a two-stage training strategy. In the first stage, the network was pretrained using data with charge number $3 \leq Z \leq 13$ and a batch size of 10 000. In the second stage, the model was post-trained with an enlarged batch size of 40 000 and data covering the range $3 \leq Z \leq 22$. The use of a large batch size is particularly important because the number of high- Z events in the beam-test dataset is relatively small. Smaller batches would contain too few samples of these rare species to yield stable gradients or meaningful statistics for the empirical latent histogram $H_{\text{emp}}(q, x)$ used in the HistoLoss. We start from $Z = 3$ for the same reason, since hydrogen and helium dominate the statistics and can be easily selected with a simple value cut.

IV. RESULTS

- (1) *Precise control of the latent-space distribution.* Compared with a plain AE or its variational and Wasserstein variants, the latent representation learned by HistoAE aligns exceptionally well with the physically motivated target distribution, namely a Gaussian mixture along the charge axis and a direct product with a uniform distribution along the position axis. As shown in Figs. 3(a) and 3(c), when applying previous solutions, such as the Wasserstein distance-based loss, the latent variables indeed occupy an approximately square region in latent space, indicating that the support of the distribution is bounded as expected. However, the internal geometry within this region remains uncontrolled: the distribution exhibits irregular, curved bandlike patterns with no clear physical meaning. While incorporating the HistoLoss, both the support and the detailed shape of the latent-space distribution align closely with the target, producing

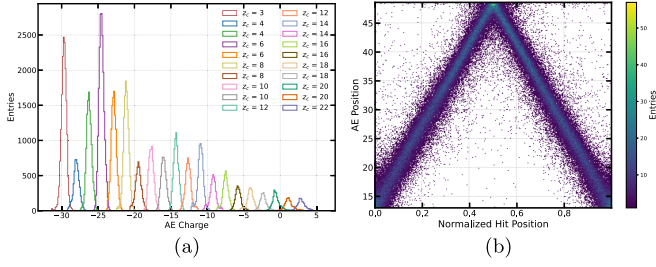


FIG. 4. Physical interpretation of the HistoAE latent space. (a) Projection onto the charge dimension, where each point is color coded by the independently measured charge from the telescope, shows that each AE-derived charge peak corresponds to the correct nuclei. (b) Comparison between the position dimension and the telescope-predicted impact position, normalized to the interval between two strips, reveals two distinct linear segments.

well-separated, linear structures as shown in Figs. 3(b) and 3(d).

- (2) *Fully interpretable latent space.* Because the latent distribution learned by HistoAE matches the desired target manifold, the two latent dimensions acquire clear and independent physical interpretations. The dimension with a Gaussian mixture structure corresponds to the particle charge, while the dimension with a uniform distribution represents the interstrip position on the detector. Figure 4 shows the physical interpretation of the latent space along the two dimensions. For the charge dimension, we color code with the true charge of the incident particle. Each peak in Fig. 4(a) corresponds to one charge species, ordered sequentially from low to high Z . For the position dimension, we compare with the impact position predicted by the telescope in Fig. 4(b), revealing a piecewise linear relation. This demonstrates that the latent variables jointly encode the complete physical information of the strip-signal pattern.
- (3) *Unsupervised precision measurement of physical quantities.* With this interpretable latent distribution, we perform unsupervised measurement of physical quantities directly from the learned latent variables. Because the latent-space control in HistoLoss operates on normalized coordinates, a postprocessing step is applied to map these to physical quantities. For the charge dimension, we identify the charge peaks and assign them to the corresponding integer charges ($Z = 3$ – 22) and interpolate in between, as shown in Figs. 5(a) and 5(b). The resulting charge distribution is shown in Fig. 5(c), where the individual peak is fitted with a Gaussian function. The σ values of the Gaussian represent the charge resolution for the corresponding nuclei. The charge resolution of nuclei between $Z = 3$ (Li) and $Z = 22$ (Ti) is shown in Fig. 5(d), which is better than $0.3e$ for all nuclei. For position measurement, a

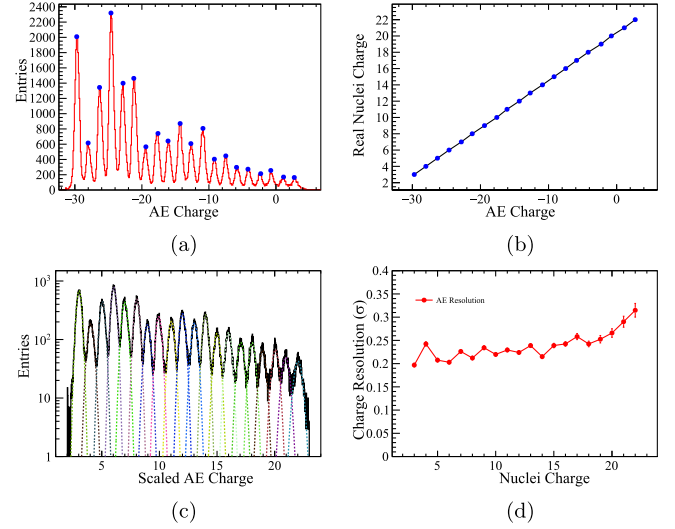


FIG. 5. Charge measurement using HistoAE. (a) Local peak finding to determine the AE charge value corresponding to each integer charge nucleus. (b) Linear correlation between the AE charge and the true integer charge. (c) Rescaled AE charge to the physical charge space through interpolation based on the linear relation, and each nucleus peak is fitted with a Gaussian function. (d) The Gaussian width represents the charge resolution for each nucleus obtained with HistoAE.

normalization procedure is also required, in which the AE position is scaled to the range of 0–1. It can also be seen from Fig. 4(b) that the AE position exhibits a piecewise linear relationship with the true impact position. This arises because particles hitting symmetrically between two strips produce mirror-symmetric clusters, while no left-right information is provided to the network. However, this ambiguity can be easily resolved by comparing the two largest channel signals within a cluster, denoted as (S_L) and (S_R), corresponding to the left and right signals, respectively. Define $\eta = \frac{S_R}{S_L + S_R}$. When $\eta < 0.5$, the impact point lies in the left half between the two strips, and vice versa. Therefore, we can express

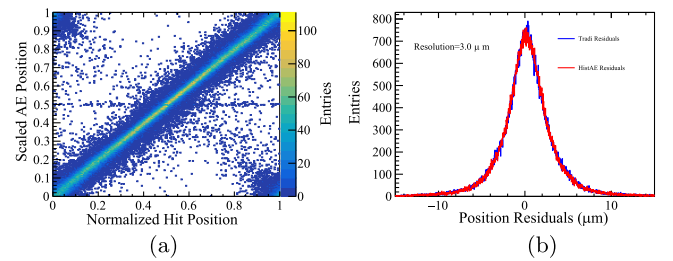


FIG. 6. Hit position measurement using HistoAE. (a) Two-dimensional distribution of the scaled AE position versus the true impact position predicted by the telescope. (b) Residual distribution of the AE reconstructed positions (red) and those obtained with the conventional method (blue). Both distributions are fitted with a double-Gaussian function, yielding a width of $3 \mu\text{m}$.

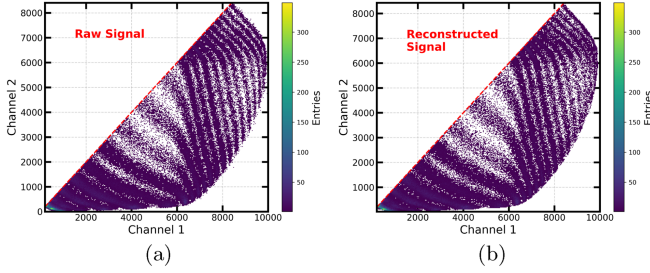


FIG. 7. Comparison between the original and HistoAE reconstructed clusters, illustrated by the two-dimensional distribution of the two largest channel signals: (a) original cluster, (b) reconstructed cluster.

$$\text{scaled AE position} = \begin{cases} \text{AE position}, & \eta < 0.5, \\ 1 - \text{AE position}, & \eta \geq 0.5. \end{cases}$$

As shown in Fig. 6(a), the distribution of the scaled impact position versus the true position is presented. The residual distribution is fitted with a double-Gaussian function, from which the position resolution achieved by HistoAE is determined to be $3 \mu\text{m}$, as shown in Fig. 6(b). It can also be seen that the position resolution obtained by HistoAE is consistent with that derived from the conventional method, which has already reached the intrinsic performance limit of the detector [54]. This indicates that HistoAE introduces no additional measurement uncertainty.

This demonstrates, for the first time, to the best of our knowledge, that highly precise measurements are achieved with a completely unsupervised learning method. Notably, HistoAE is also the first unsupervised algorithm capable of performing simultaneous charge and position reconstruction with SSDs, without requiring any labeled training data.

- (4) *Fast simulation of SSD response.* Accurate and fast simulation of particle physics processes is crucial for the high-energy physics community [55,56]. Regarding the use of HistoAE for fast simulation, this possibility is enabled by the fact that the learned latent-space distribution intrinsically captures the detector-response fluctuations, while the decoder maps these latent representations back to realistic detector clusters. The reconstruction quality of the decoder is illustrated directly in Fig. 7, where the reconstructed cluster distribution reproduces the main structures of the original data. For example, for incident particles with charge $Z = 10$, one may smear the charge coordinate according to the corresponding learned latent-space distribution and then pass the resulting latent representations through the decoder to generate detector clusters. In this way, the generated clusters naturally reproduce the characteristic band

structure (just like the pattern shown in Fig. 1) associated with $Z = 10$.

Therefore, the decoder of HistoAE can serve as a physically motivated generator of SSD responses. At the same time, for applications to downstream physics analyses, the quantitative accuracy of such a fast-simulation approach still needs to be evaluated in future studies.

V. SUMMARY

Unsupervised learning is widely used, yet its application to high-precision physical measurements has remained challenging. We present HistoAE, an unsupervised deep learning model capable of performing precise, interpretable measurements. When applied to silicon strip detectors, HistoAE learns a two-dimensional latent space whose axes correspond directly to particle charge and impact position. It achieves charge and position resolutions of approximately $0.25e$ and $3 \mu\text{m}$, on par with conventional approaches. Moreover, its decoder can be repurposed for fast detector simulation. At present, the method is suitable for low-dimensional latent spaces. For tasks involving higher-dimensional latent representations, larger statistical samples are required to achieve effective control of the latent space. It will be addressed in future work.

Looking ahead, we also plan to further develop HistoAE for the future in-orbit operation of the AMS layer-0 detector. In this context, HistoAE may offer two potential advantages: first, by reconstructing charge and position simultaneously in a unified representation learning framework, it may help reduce the mutual error propagation present in conventional sequential corrections; second, because it does not rely on event-by-event labels from other subdetectors during training, it may help preserve more usable physics events in the large-acceptance L0 configuration. These possible advantages will need to be investigated carefully with real in-orbit data in future studies.

ACKNOWLEDGMENTS

We would like to express our sincere gratitude to Luya Lou (AMSS), Shudong Wang (IHEP), Yiming Wang (IHEP), Zhiyu Xiang (CSU), Jiaoyang Xu (BNU), Yepeng Yan (IHEP), and Tongtian Zhu (ZJU) for the inspiring discussions that greatly facilitated the design of HistoAE. This study was supported by the National Key Research and Development Program of China (Grant No. 2022YFA1604800) and the National Natural Science Foundation of China (Grant No. 12342503).

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

- [1] Y. LeCun, Y. Bengio, and G. Hinton, *Nature (London)* **521**, 436 (2015).
- [2] J. Schmidhuber, *Neural Netw.* **61**, 85 (2015).
- [3] L. Lonnblad, C. Peterson, and T. Rognvaldsson, *Phys. Rev. Lett.* **65**, 1321 (1990).
- [4] B. H. Denby, *Comput. Phys. Commun.* **49**, 429 (1988).
- [5] K. Cranmer, U. Seljak, and K. Terao, in *Review of Particle Physics*, edited by R. L. Workman *et al.* (Lawrence Berkeley National Laboratory, 2023), Chap. 41, pp. 711–744, <https://pdg.lbl.gov/2023/reviews/rpp2023-rev-machine-learning.pdf>.
- [6] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, *Rev. Mod. Phys.* **91**, 045002 (2019).
- [7] A. Radovic, M. Williams, D. Rousseau, M. Kagan, D. Bonacorsi, A. Himmel, A. Aurisano, K. Terao, and T. Wongjirad, *Nature (London)* **560**, 41 (2018).
- [8] D. Guest, K. Cranmer, and D. Whiteson, *Annu. Rev. Nucl. Part. Sci.* **68**, 161 (2018).
- [9] L. Yu, J. Wang, R. Qiao, K. Gong, W. Peng, J. Wei, B. Lu, D. Guo, Y. Liu, X. Liu *et al.*, *Astron. Comput.* **53**, 100986 (2025).
- [10] G. Ambrosi, V. Choutko, C. Delgado, A. Oliva, Q. Yan, and Y. Li, *Nucl. Instrum. Methods Phys. Res., Sect. A* **869**, 29 (2017).
- [11] Y. Jia, Q. Yan, V. Choutko, H. Liu, and A. Oliva, *Nucl. Instrum. Methods Phys. Res., Sect. A* **972**, 164169 (2020).
- [12] Y. Bengio, A. Courville, and P. Vincent, *IEEE Trans. Pattern Anal. Mach. Intell.* **35**, 1798 (2013).
- [13] L. Jing and Y. Tian, *IEEE Trans. Pattern Anal. Mach. Intell.* **43**, 4037 (2021).
- [14] J. F. Rodriguez-Nieva and M. S. Scheurer, *Nat. Phys.* **15**, 790 (2019).
- [15] M. S. Scheurer and R.-J. Slager, *Phys. Rev. Lett.* **124**, 226401 (2020).
- [16] K. Kottmann, P. Huembeli, M. Lewenstein, and A. Acín, *Phys. Rev. Lett.* **125**, 170603 (2020).
- [17] C. L. Cheng, G. Singh, and B. Nachman, *Phys. Rev. Lett.* **135**, 021801 (2025).
- [18] R. Iten, T. Metger, H. Wilming, L. del Rio, and R. Renner, *Phys. Rev. Lett.* **124**, 010508 (2020).
- [19] B. Hou, J. Wu, and D. Y. Qiu, *Nat. Commun.* **15**, 9481 (2024).
- [20] S. Kolouri, P. E. Pope, C. E. Martin, and G. K. Rohde, in *International Conference on Learning Representations (ICLR)* (ICLR, New Orleans, LA, USA, 2019).
- [21] K. Fraser, S. Homiller, R. K. Mishra, B. Ostdiek, and M. D. Schwartz, *J. High Energy Phys.* **03** (2021) 066.
- [22] S. Seidel, *Phys. Rep.* **828**, 1 (2019).
- [23] J.-J. Wei, J.-H. Guo, and Y.-M. Hu, *Nucl. Sci. Tech.* **31**, 97 (2020).
- [24] M. Aguilar *et al.* (AMS Collaboration), *Phys. Rep.* **894**, 1 (2021).
- [25] X. Zhou and J. Yang (HIAF project Team Collaboration), *AAPPS Bull.* **32**, 35 (2022).
- [26] J. Pata, J. Duarte, J.-R. Vlimant, M. Pierini, and M. Spiropulu, *Eur. Phys. J. C* **81**, 381 (2021).
- [27] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Learning internal representations by error propagation, in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations* (MIT Press, Cambridge, MA, USA, 1986), p. 318–362.
- [28] P. Baldi and K. Hornik, *Neural Netw.* **2**, 53 (1989).
- [29] D. P. Kingma and M. Welling, in *Proceedings of the 2nd International Conference on Learning Representations (ICLR)* (ICLR, Banff, AB, Canada, 2014).
- [30] D. J. Rezende, S. Mohamed, and D. Wierstra, in *Proceedings of the 31st International Conference on Machine Learning (ICML)* (PMLR, Beijing, China, 2014), pp. 1278–1286.
- [31] I. Tolstikhin, O. Bousquet, S. Gelly, and B. Schölkopf, in *Proceedings of the 6th International Conference on Learning Representations (ICLR)* (ICLR, Vancouver, BC, Canada, 2018).
- [32] A. A. Alves *et al.* (LHCb Collaboration), *J. Instrum.* **3**, S08005 (2008).
- [33] S. Chatrchyan *et al.* (CMS Collaboration), *J. Instrum.* **3**, S08004 (2008).
- [34] G. Aad *et al.* (ATLAS Collaboration), *J. Instrum.* **3**, S08003 (2008).
- [35] R. He *et al.*, *Nucl. Sci. Tech.* **34**, 205 (2023).
- [36] S. Straulino *et al.* (PAMELA Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **530**, 168 (2004).
- [37] Y.-F. Dong, F. Zhang, R. Qiao, W.-X. Peng, R.-R. Fan, K. Gong, D. Wu, and H.-Y. Wang, *Chin. Phys. C* **39**, 116202 (2015).
- [38] J. T. Vievering, L. Glesener, P. S. Athiray, J. C. Buitrago-Casas, S. Musset, D. Ryan, S.-n. Ishikawa, J. Duncan, S. Christe, and S. Krucker, *Astrophys. J.* **913**, 15 (2021).
- [39] K. Lubelsmeyer *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **654**, 639 (2011).
- [40] M. Duranti (AMS 02 Tracker Collaboration), *Proc. Sci., Vertex2012* (2013) 052.
- [41] D. Miao *et al.*, *arXiv:2505.23050*.
- [42] H. A. Bethe, *Ann. Phys. (Berlin)* **397**, 325 (1930).
- [43] M. Boronat, C. Marinas, A. Frey, I. García, B. Schwenker, M. Vos, and F. Wilk, *IEEE Trans. Nucl. Sci.* **62**, 381 (2015).
- [44] X. Altuna, C. Arimatea, R. Bailey, P. Baudrenghien, G. Crockford, G. De Rijk, C. Despas, A. Faugier *et al.*, A momentum calibration of the SPS proton beam, Technical Report, CERN-SL-92-32-OP, CERN, 1992.
- [45] F. Charton, *arXiv:2308.15594*.
- [46] S. Golkar, M. Pettee, M. Eickenberg, A. Bietti, M. Cranmer, G. Krawezik, F. Lanusse, M. McCabe, R. Ohana, L. Parker *et al.*, *arXiv:2310.02989*.
- [47] S. Ioffe and C. Szegedy, in *Proceedings of the 32nd International Conference on Machine Learning (ICML)* (PMLR, Lille, France, 2015), Vol. 37, pp. 448–456.
- [48] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, in *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)* (Curran Associates, Inc., Red Hook, NY, USA, 2019), pp. 8024–8035.
- [49] I. Loshchilov and F. Hutter, in *International Conference on Learning Representations (ICLR)* (ICLR, New Orleans, LA, USA, 2019).
- [50] K. He, X. Zhang, S. Ren, and J. Sun, in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*

- (IEEE Computer Society, Los Alamitos, CA, USA, 2015), pp. 1026–1034.
- [51] P. Goyal, P. Dollár, R. Girshick, P. Noordhuis, L. Wesolowski, A. Kyrola, A. Tulloch, Y. Jia, and K. He, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops* (2017), [arXiv:1706.02677](#).
- [52] I. Loshchilov and F. Hutter, in *International Conference on Learning Representations (ICLR)* (ICLR, Toulon, France, 2017).
- [53] torch.amp.GradScaler.
- [54] R. Turchetta, *Nucl. Instrum. Methods Phys. Res., Sect. A* **335**, 44 (1993).
- [55] J. Apostolakis *et al.*, *Front. Phys.* **10**, 913510 (2022).
- [56] B. Hashemi and C. Krause, *Rev. Phys.* **12**, 100092 (2024).