
Latent Representation Learning in Heavy-Ion Collisions with MaskPoint Transformer

Jing-Zong Zhang and Shuang Guo

Institute of Modern Physics, Fudan University
Shanghai 200433, China

Shanghai Research Center for Theoretical Nuclear Physics
NSFC and Fudan University, Shanghai 200438, China

Li-Lin Zhu

Department of Physics, Sichuan University
Chengdu 610064, China

Lingxiao Wang

RIKEN Interdisciplinary Theoretical and Mathematical Sciences (iTHEMS),
Wako, Saitama 351-0198, Japan
lingxiao.wang@riken.jp

Guo-Liang Ma

Institute of Modern Physics, Fudan University
Shanghai 200433, China
Shanghai Research Center for Theoretical Nuclear Physics
NSFC and Fudan University, Shanghai 200438, China
glma@fudan.edu.cn

Abstract

A central challenge in high-energy nuclear physics is to extract informative features from the high-dimensional final-state data of heavy-ion collisions (HIC) in order to enable reliable downstream analyses. Traditional approaches often rely on selected observables, which may miss subtle but physically relevant structures in the data. To address this, we introduce a Transformer-based autoencoder trained with a two-stage paradigm: self-supervised pre-training followed by supervised fine-tuning. The pre-trained encoder learns latent representations directly from unlabeled HIC data, providing a compact and information-rich feature space that can be adapted to diverse physics tasks. As a case study, we apply the method to distinguish between large and small collision systems, where it achieves significantly higher classification accuracy than PointNet. Principal component analysis and SHAP interpretation further demonstrate that the autoencoder captures complex non-linear correlations beyond individual observables, yielding features with strong discriminative and explanatory power. These results establish our two-stage framework as a general and robust foundation for feature learning in HIC, opening the door to more powerful analyses of quark-gluon plasma properties and other emergent phenomena. The implementation is publicly available at <https://github.com/Giovanni-Sforza/MaskPoint-AMPT>.

1 Introduction

Relativistic heavy-ion collisions provide a unique environment to study the QCD phase transition of nuclear matter at extreme temperature and density, and to probe the properties of deconfined partonic matter, known as quark–gluon plasma (QGP) [1–5]. Traditional observables such as particle spectra and anisotropic flows have revealed essential QGP features [6–9], yet they are limited in dealing with high-dimensional final-state data. This restricts their effectiveness in capturing the subtle dynamical differences across varying collision system sizes, for example, between large (Pb+Pb) and small (p+Pb) systems [10–17].

Recent advances in machine learning have introduced deep learning techniques into high-energy nuclear physics [18–28]. Since the final-state particle phase space can be naturally represented as unordered point clouds, models such as PointNet and Point Transformer [29–32] have been applied to extract representations directly from their raw data [26]. These approaches enable data-driven exploration of QGP properties beyond traditional observables [33–36].

In this work, we introduce a Transformer-based autoencoder [32] into heavy-ion collision analysis, employing a two-stage paradigm of self-supervised pretraining followed by supervised finetuning. The training data are provided by the AMPT (A Multi-Phase Transport) model [37, 38], a widely used framework that incorporates initial condition, parton cascade, hadronization, and hadronic rescatterings, and has been shown to reproduce many experimental observables. Using Pb+Pb and p+Pb events generated by AMPT, the autoencoder first learns latent representations by minimizing reconstruction error. The pretrained encoder is then finetuned with a lightweight classifier for system identification. Compared with the PointNet baseline [39, 40], our method achieves a significantly higher classification accuracy. Furthermore, principal component analysis, correlation analysis, and SHAP interpretation reveal that the learned features capture both linear and nonlinear connections with traditional observables. While identifying the collision system is not a practical challenge, we employ it as a benchmark task to quantitatively assess the discriminative power of the latent representations learned by our model[41].

Our results show that self-supervised pretraining with a Transformer-based autoencoder can extract physically meaningful structures from high-dimensional collision data. This two-stage framework offers an effective methodology for future studies of QGP properties.

2 Model Architecture and Methodology

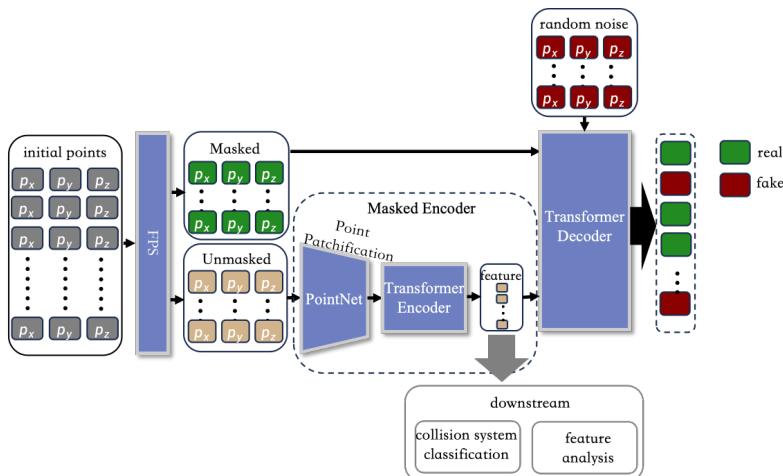


Figure 1: Architecture of the masked autoencoder. The input point cloud is first partitioned into patches by PointNet and then encoded by a Transformer. During pre-training, the encoder output is passed to a Transformer decoder for an discrimination task between real and fake particles. For fine-tuning and analysis, only the pretrained encoder is retained as the feature extractor.

This study utilizes p+Pb and Pb+Pb collision events from the AMPT model. Each event is represented as a point cloud composed of the three-momenta (p_x, p_y, p_z) of 128 final-state particles within the kinematic window of pseudorapidity $|\eta| < 2.4$ and transverse momentum $p_T > 0.4$ GeV/c [42, 40]. To learn the intrinsic physical structure of the data, rather than degenerating into a trivial coordinate-memorization task, we designed a masked autoencoder based on a discrimination scheme for pre-training (see Figs. 1) [32]. Specifically, we randomly mask 25% of each event's point cloud using Farthest Point Sampling (FPS), and the unmasked portion, $\mathbf{U} \in \mathbb{R}^{96 \times 3}$, is fed into an encoder. This encoder first extracts local features using a PointNet network, then captures global information through a 6-layer Transformer architecture, generating a 96-dimensional feature vector \mathbf{f} . Subsequently, a single-layer Transformer decoder employs a cross-attention mechanism to integrate the feature vector \mathbf{f} with either the masked "real point cloud" or a randomly sampled "fake point cloud," which is then judged by an MLP binary classifier head using cross-entropy loss function. This discrimination task compels the encoder to learn high-quality physical features, as detailed in the appendixA.1.1. After pre-training, we freeze the encoder parameters for downstream tasks.

For the task of collision system identification, we remove the decoder and feed the features \mathbf{f} extracted by the encoder into a simple MLP classifier. Both the pre-training and fine-tuning stages were run for up to 300 epochs, with the best-performing model saved based on validation performance.

To provide a complete overview of our experimental configuration and validation, detailed settings and results are summarized in the appendixA. Specifically, Tables1 and 2 list the hyperparameters for the pre-training and fine-tuning stages, while Table3 presents ablation results validating our key design choices, such as the PointNet preprocessor and a 25% masking ratio. Table4 further compares the model sizes, showing that our superior performance arises from the two-stage training paradigm rather than from an increased parameter count. To analyze the physical meaning of the features, we use Principal Component Analysis for dimensionality reduction, then calculate the linear correlations between the principal components and physical observables. Finally, we combine a Random Forest model with SHAP (SHapley Additive exPlanations) analysis to investigate non-linear associations between them [43].

3 Results and discussion

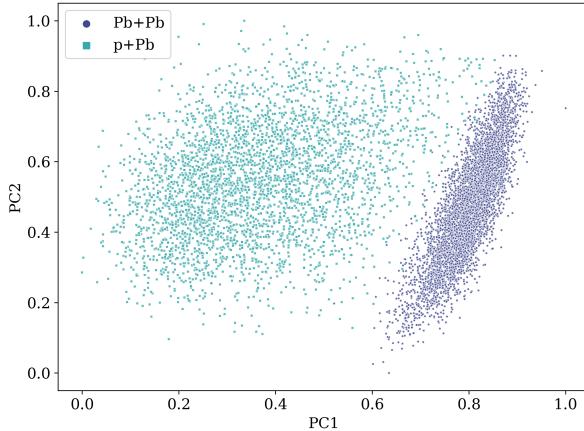


Figure 2: PCA projection of the latent features learned by the autoencoder during pre-training with 3D momentum inputs, shown in the PC1–PC2 plane. The two colors denote two different systems. The clear clustering indicates that the model, even without labels, captures and distinguishes intrinsic physical differences between the two systems.

This study demonstrates the successful application of a masked autoencoder within a self-supervised pre-training and supervised fine-tuning paradigm to extract physically meaningful representations from the final-state three-dimensional momentum data of heavy-ion collisions. In the pre-training phase, the model, without any label information, learned to spontaneously distinguish between p+Pb

and Pb+Pb collision systems in its latent space (Fig. 2), highlighting the potential of self-supervised learning to uncover intrinsic data structures.

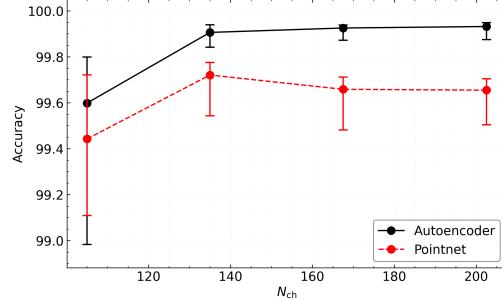


Figure 3: Classification accuracy between large and small systems across N_{ch} bins for the fine-tuned autoencoder (black) and PointNet (red). The autoencoder consistently outperforms PointNet, validating the effectiveness of the “self-supervised pre-training + supervised fine-tuning” strategy.

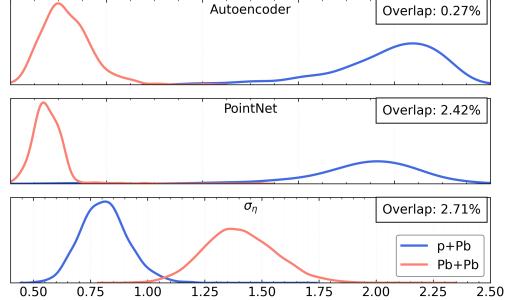


Figure 4: Distributions of PC1 from the autoencoder (top) and PointNet (middle), compared with σ_η (bottom). The overlap between Pb+Pb and p+Pb is 0.27% for the autoencoder, lower than PointNet (2.42%) and σ_η (2.71%). While PointNet approaches the theoretical limit of σ_η , the autoencoder surpasses it, demonstrating stronger discriminative power from self-supervised pre-training.

In the subsequent fine-tuning phase, our model’s superiority is quantitatively demonstrated by its classification accuracy, which significantly surpasses the PointNet baseline across all tested multiplicity ranges (Fig. 3). Error bars represent the range of the accuracies over 10 independent trials, where the test set is split into 10 subsets for each trial. To investigate the origin of this performance advantage, we analyzed the principal components (PCs) of the features learned by both models. This analysis reveals a striking difference in the discriminative power of their leading PCs (Fig. 4). The distribution overlap for our autoencoder’s PC1 is merely 0.27%, substantially lower than the 2.42% overlap for PointNet’s PC1. Crucially, PointNet’s overlap closely mirrors the 2.71% overlap of the raw physical observable, the standard deviation of pseudorapidity distribution (σ_η), suggesting that its learned representation is closely tied to this single variable.

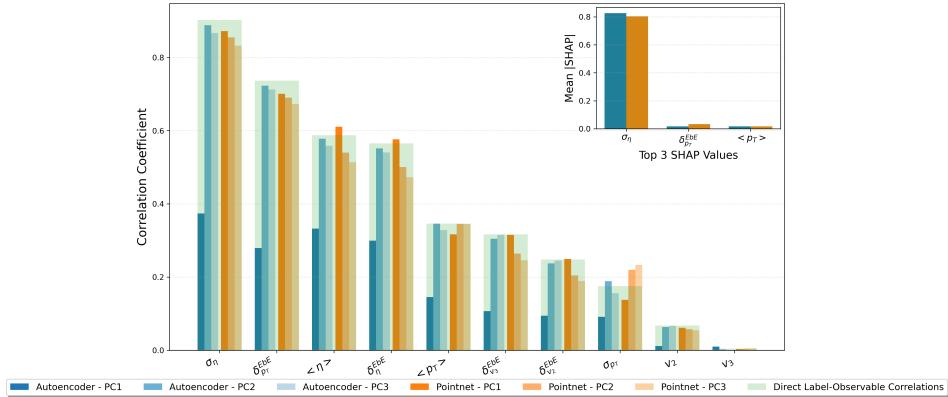


Figure 5: Correlation coefficients between PCs and observables, with SHAP contributions to PC1. Autoencoder PC1 shows near-zero linear correlations but a high SHAP weight from σ_η , indicating that it encodes key information in a non-linear manner, which explains its superior performance.

This observation motivates a deeper analysis using SHAP values and correlation coefficients to probe the structure of the learned features. Fig. 5 shows that PointNet’s PC1 is strongly linearly correlated with σ_η , indicating that it mainly reproduces a linear representation of the most discriminative observable. By contrast, autoencoder’s PC1 exhibits almost zero linear correlation with σ_η , yet SHAP analysis identifies σ_η as its most influential contributor. This combination of low linear correlation

and high SHAP importance demonstrates that the autoencoder, through self-supervised pre-training, learns a more complex non-linear representation. Rather than duplicating a single observable, the pretrained encoder extracts richer and more abstract features, which explains its markedly lower distribution overlap and the resulting superior classification performance. Therefore, this work not only provides a high-performance model for identifying heavy ion collision systems, but clearly demonstrates that selfsupervised pretraining can guide a model to move beyond simple linear fitting of key physical observables and instead learn deeper, non-linear patterns within the data, offering a powerful new perspective for AI-driven discovery in fundamental high-energy nuclear physics.

4 Conclusion, Limitations and Outlook

This work demonstrates the effectiveness of a self-supervised pre-training and fine-tuning paradigm with a Transformer-based autoencoder for heavy-ion collision analysis. Our model learns deep and physically meaningful representations from unlabeled data, outperforming PointNet by capturing complex non-linear features rather than merely reproducing obvious traditional observables.

Despite these advantages, several limitations remain. Our study relies solely on AMPT-generated events and only considers p+Pb and Pb+Pb systems. The input is limited to particle three-momenta (p_x, p_y, p_z), without incorporating four-momentum or additional features such as charge or spin. Addressing these aspects is crucial for enhancing robustness and broad applicability.

Looking forward, promising directions include exploring advanced generative and Transformer variants [44–49], embedding physics-inspired priors such as conservation laws and symmetries [50], and strengthening the connection between latent representations and established physical observables [51, 52]. Beyond the generic masking strategy adopted here, pre-training can further be designed with physics-informed objectives to construct data-driven observables [53–58], potentially uncovering subtle signals such as the Chiral Magnetic Effect [59–61] or nuclear deformation [62, 63]. Such strategies could transform machine learning into an active partner in the AI discovery of new physics in high-energy heavy-ion collisions.

5 Acknowledgments

This work is partially supported by the National Natural Science Foundation of China under Grants No.12147101 and No. 12325507, the National Key Research and Development Program of China under Grant No. 2022YFA1604900, the Guangdong Major Project of Basic and Applied Basic Research under Grant No. 2020B0301030008 (J.Z, S.G. and G.M.), the RIKEN TRIP initiative (RIKEN Quantum), JSPS KAKENHI Grant No. 25H01560, and JST-BOOST Grant No.JPMJBY24H9 (L.W.). We also thank the DEEP-IN working group at RIKEN-iTHEMS for support in the preparation.

References

- [1] J. Adams, M. Aggarwal, Z. Ahammed, J. Amonett, B. Anderson, D. Arkhipkin et al., *Experimental and theoretical challenges in the search for the quark–gluon plasma: The star collaboration’s critical assessment of the evidence from rhic collisions*, *Nuclear Physics A* **757** (2005) 102.
- [2] K. Adcox, S.S. Adler, S. Afanasiev, C. Aidala, N. Ajitanand, Y. Akiba et al., *Formation of dense partonic matter in relativistic nucleus–nucleus collisions at rhic: experimental evaluation by the phenix collaboration*, *Nuclear Physics A* **757** (2005) 184.
- [3] ALICE collaboration, *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002.
- [4] A. Bzdak, S. Esumi, V. Koch, J. Liao, M. Stephanov and N. Xu, *Mapping the Phases of Quantum Chromodynamics with Beam Energy Scan*, *Phys. Rept.* **853** (2020) 1 [1906.00936].
- [5] X. Luo and N. Xu, *Search for the QCD Critical Point with Fluctuations of Conserved Quantities in Relativistic Heavy-Ion Collisions at RHIC : An Overview*, *Nucl. Sci. Tech.* **28** (2017) 112 [1701.02105].

- [6] F. Karsch, A. Bazavov, H.-T. Ding, P. Hegde, O. Kaczmarek, E. Laermann et al., *Conserved charge fluctuations from lattice qcd and the beam energy scan*, *Nuclear Physics A* **956** (2016) 352.
- [7] H.-j. Xu, Z. Li and H. Song, *High-order flow harmonics of identified hadrons in 2.76 a tev pb+pb collisions*, *Physical Review C* **93** (2016) 064905.
- [8] X. Zhu, Y. Zhou, H. Xu and H. Song, *Correlations of flow harmonics in 2.76 a tev pb-pb collisions*, *Physical Review C* **95** (2017) 044902.
- [9] N.-B. Chang, S. Cao and G.-Y. Qin, *Probing medium-induced jet splitting and energy loss in heavy-ion collisions*, *Physics Letters B* **781** (2018) 423.
- [10] K. Dusling, W. Li and B. Schenke, *Novel collective phenomena in high-energy proton-proton and proton-nucleus collisions*, *Int. J. Mod. Phys. E* **25** (2016) 1630002 [1509.07939].
- [11] C. Loizides, *Experimental overview on small collision systems at the LHC*, *Nucl. Phys. A* **956** (2016) 200 [1602.09138].
- [12] J.L. Nagle and W.A. Zajc, *Small System Collectivity in Relativistic Hadronic and Nuclear Collisions*, *Ann. Rev. Nucl. Part. Sci.* **68** (2018) 211 [1801.03477].
- [13] H.-S. Wang and G.-L. Ma, *Testing the collectivity in large and small colliding systems with test particles*, *Phys. Rev. C* **106** (2022) 064907 [2208.06854].
- [14] X.-L. Zhao, Z.-W. Lin, L. Zheng and G.-L. Ma, *A transport model study of multiparticle cumulants in p+p collisions at 13 TeV*, *Phys. Lett. B* **839** (2023) 137799 [2112.01232].
- [15] J.-L. Pei, G.-L. Ma and A. Bzdak, *Effect of transverse momentum conservation and flow on symmetric cumulants $sc_{2,3}\{4\}$ and $sc_{2,3,4}\{6\}$* , 2503.12846.
- [16] G.-L. Ma and A. Bzdak, *Flow in small systems from parton scatterings*, *Nucl. Phys. A* **956** (2016) 745.
- [17] A. Bzdak and G.-L. Ma, *Elliptic and triangular flow in p+Pb and peripheral Pb+Pb collisions from parton scatterings*, *Phys. Rev. Lett.* **113** (2014) 252301 [1406.2804].
- [18] ALICE collaboration, *Using machine learning for particle identification in ALICE*, *JINST* **17** (2022) C07016 [2204.06900].
- [19] M. ZHOU, Y. LUO and H. SONG, *Applications of machine learning in relativistic heavy ion physics*, *Sci. China Phys. Mech. Astron.* **52** (2022) 252002.
- [20] L. WANG, L. PANG and K. ZHOU, *Applications of deep learning in high energy nuclear physics*, *Sci. China Phys. Mech. Astron.* **52** (2022) 252003.
- [21] W. HE, J. HE, R. Wang and Y. MA, *Machine learning applications in nuclear physics*, *Sci. China Phys. Mech. Astron.* **52** (2022) 252004.
- [22] Y.-L. DU, D. PABLOS and K. Tywoniuk, *Applications of deep learning in jet quenching*, *Sci. China Phys. Mech. Astron.* **52** (2022) 252017.
- [23] F. LI, L. PANG and X. WANG, *Application of machine learning to the study of QCD transition in heavy ion collisions*, *Nucl. Sci. Tech.* **46** (2023) 040014.
- [24] W. He, Q. Li, Y. Ma, Z. Niu, J. Pei and Y. Zhang, *Machine learning in nuclear physics at low and intermediate energies*, *Sci. China Phys. Mech. Astron.* **66** (2023) 282001 [2301.06396].
- [25] Y.-G. Ma, L.-G. Pang, R. Wang and K. Zhou, *Phase Transition Study Meets Machine Learning*, *Chin. Phys. Lett.* **40** (2023) 122101 [2311.07274].
- [26] K. Zhou, L. Wang, L.-G. Pang and S. Shi, *Exploring QCD matter in extreme conditions with Machine Learning*, *Prog. Part. Nucl. Phys.* **135** (2024) 104084 [2303.15136].
- [27] A. Boehnlein et al., *Colloquium: Machine learning in nuclear physics*, *Rev. Mod. Phys.* **94** (2022) 031003 [2112.02309].

- [28] G. Aarts, K. Fukushima, T. Hatsuda, A. Ipp, S. Shi, L. Wang et al., *Physics-driven learning for inverse problems in quantum chromodynamics*, *Nature Rev. Phys.* **7** (2025) 154 [2501.05580].
- [29] R. Zhang, Z. Guo, P. Gao, R. Fang, B. Zhao, D. Wang et al., *Point-m2ae: multi-scale masked autoencoders for hierarchical point cloud pre-training*, *Advances in neural information processing systems* **35** (2022) 27061.
- [30] R. Zhang, L. Wang, Y. Qiao, P. Gao and H. Li, *Learning 3d representations from 2d pre-trained models via image-to-point masked autoencoders*, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 21769–21780, 2023.
- [31] X. Yu, L. Tang, Y. Rao, T. Huang, J. Zhou and J. Lu, *Point-bert: Pre-training 3d point cloud transformers with masked point modeling*, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 19313–19322, 2022.
- [32] H. Liu, M. Cai and Y.J. Lee, *Masked discrimination for self-supervised learning on point clouds*, in *European Conference on Computer Vision*, pp. 657–675, Springer, 2022.
- [33] T. Golling, L. Heinrich, M. Kagan, S. Klein, M. Leigh, M. Osadchy et al., *Masked particle modeling on sets: towards self-supervised high energy physics foundation models*, *Machine Learning: Science and Technology* **5** (2024) 035074.
- [34] J. Birk, A. Hallin and G. Kasieczka, *Omnijet- α : the first cross-task foundation model for particle physics*, *Machine Learning: Science and Technology* **5** (2024) 035031.
- [35] P. Harris, M. Kagan, J. Krupa, B. Maier and N. Woodward, *Re-simulation-based self-supervised learning for pre-training foundation models*, *arXiv preprint arXiv:2403.07066* (2024) .
- [36] V. Mikuni and B. Nachman, *Method to simultaneously facilitate all jet physics tasks*, *Physical Review D* **111** (2025) 054015.
- [37] Z.-W. Lin, C.M. Ko, B.-A. Li, B. Zhang and S. Pal, *Multiphase transport model for relativistic heavy ion collisions*, *Physical Review C—Nuclear Physics* **72** (2005) 064901.
- [38] Z.-W. Lin and L. Zheng, *Further developments of a multi-phase transport model for relativistic nuclear collisions*, *Nuclear Science and Techniques* **32** (2021) 113.
- [39] C.R. Qi, H. Su, K. Mo and L.J. Guibas, *Pointnet: Deep learning on point sets for 3d classification and segmentation*, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 652–660, 2017.
- [40] S. Guo, H.-S. Wang, K. Zhou and G.-L. Ma, *Machine learning study to identify collective flow in small and large colliding systems*, *Physical Review C* **110** (2024) 024910.
- [41] Y.S. Lai, J. Mulligan, M. Płoskoń and F. Ringer, *The information content of jet quenching and machine learning assisted observable design*, *Journal of High Energy Physics* **2022** (2022) 1.
- [42] CMS collaboration, *Multiplicity and Transverse Momentum Dependence of Two- and Four-Particle Correlations in pPb and PbPb Collisions*, *Phys. Lett. B* **724** (2013) 213 [1305.0609].
- [43] S.M. Lundberg and S.-I. Lee, *A unified approach to interpreting model predictions*, in *Advances in Neural Information Processing Systems*, I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan et al., eds., vol. 30, Curran Associates, Inc., 2017.
- [44] K. Dohi, *Variational autoencoders for jet simulation*, *arXiv preprint arXiv:2009.04842* (2020) .
- [45] B. Ostdiek, *Deep set auto encoders for anomaly detection in particle physics*, *SciPost Physics* **12** (2022) 045.
- [46] V. Mikuni, B. Nachman and M. Pettee, *Fast point cloud generation with diffusion models in high energy physics*, *Physical Review D* **108** (2023) 036025.
- [47] A. Butter, N. Huetsch, S. Palacios Schweitzer, T. Plehn, P. Sorrenson and J. Spinner, *Jet diffusion versus jetgpt—modern networks for the lhc*, *SciPost Physics Core* **8** (2025) 026.

- [48] H. Qu, C. Li and S. Qian, *Particle transformer for jet tagging*, in *International Conference on Machine Learning*, pp. 18281–18292, PMLR, 2022.
- [49] S. Qiu, S. Han, X. Ju, B. Nachman and H. Wang, *Holistic approach to predicting top quark kinematic properties with the covariant particle transformer*, *Physical Review D* **107** (2023) 114029.
- [50] A. Bogatskiy, S. Ganguly, T. Kipf, R. Kondor, D.W. Miller, D. Murnane et al., *Symmetry group equivariant architectures for physics*, *arXiv preprint arXiv:2203.06153* (2022) .
- [51] R. Gambhir, A. Osathapan and J. Thaler, *Streamlining latent spaces in machine learning using moment pooling*, *Physical Review D* **110** (2024) 074020.
- [52] A. Hallin, G. Kasieczka, S. Kraml, A. Lessa, L. Moureaux, T. von Schwartz et al., *Universal new physics latent space*, *Physical Review D* **111** (2025) 016006.
- [53] S. Gong, Q. Meng, J. Zhang, H. Qu, C. Li, S. Qian et al., *An efficient lorentz equivariant graph neural network for jet tagging*, *Journal of High Energy Physics* **2022** (2022) 1.
- [54] A. Bogatskiy, T. Hoffman, D.W. Miller and J.T. Offermann, *Pelican: Permutation equivariant and lorentz invariant or covariant aggregator network for particle physics*, *arXiv preprint arXiv:2211.00454* (2022) .
- [55] J. Spinner, V. Bresó, P. De Haan, T. Plehn, J. Thaler and J. Brehmer, *Lorentz-equivariant geometric algebra transformers for high-energy physics*, *Advances in neural information processing systems* **37** (2024) 22178.
- [56] J. Brehmer, V. Bresó, P. de Haan, T. Plehn, H. Qu, J. Spinner et al., *A lorentz-equivariant transformer for all of the lhc*, *arXiv preprint arXiv:2411.00446* (2024) .
- [57] J. Spinner, L. Favaro, P. Lippmann, S. Pitz, G. Gerhartz, T. Plehn et al., *Lorentz local canonicalization: How to make any network lorentz-equivariant*, *arXiv preprint arXiv:2505.20280* (2025) .
- [58] L. Favaro, G. Gerhartz, F.A. Hamprecht, P. Lippmann, S. Pitz, T. Plehn et al., *Lorentz-equivariance without limitations*, *arXiv preprint arXiv:2508.14898* (2025) .
- [59] K. Fukushima, D.E. Kharzeev and H.J. Warringa, *The Chiral Magnetic Effect*, *Phys. Rev. D* **78** (2008) 074033 [0808.3382].
- [60] Y.-S. Zhao, L. Wang, K. Zhou and X.-G. Huang, *Detecting the chiral magnetic effect via deep learning*, *Phys. Rev. C* **106** (2022) L051901 [2105.13761].
- [61] S. Guo, L. Wang, K. Zhou and G.-L. Ma, *Neural Unfolding of the Chiral Magnetic Effect in Heavy-Ion Collisions*, *Chin. Phys. Lett.* **42** (2025) 110101 [2507.05808].
- [62] J. Jia, G. Giacalone and C. Zhang, *Separating the Impact of Nuclear Skin and Nuclear Deformation in High-Energy Isobar Collisions*, *Phys. Rev. Lett.* **131** (2023) 022301 [2206.10449].
- [63] X.-L. Zhao and G.-L. Ma, *Search for the chiral magnetic effect in collisions between two isobars with deformed and neutron-rich nuclear structures*, *Phys. Rev. C* **106** (2022) 034909 [2203.15214].
- [64] K. He, X. Chen, S. Xie, Y. Li, P. Dollar and R. Girshick, *Masked autoencoders are scalable vision learners*, in *Proceedings - 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022*, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 15979–15988, IEEE Computer Society, 2022, DOI.

Table 1: Pretraining stage training hyperparameter settings

hyperparameter	value
epochs	300
optimizer	AdamW
learning rate	0.001
weight decay	0.05
LR schedule	cosine decay
warmup epochs	3
batch size	256
mask ratio	0.25
mask type	random

Table 2: Finetuning phase training hyperparameter settings

hyperparameter	value
epochs	300
optimizer	AdamW
learning rate	0.0005
weight decay	0.05
LR schedule	cosine decay
warmup epochs	10
batch size	128

A Technical Appendices and Supplementary Material

Hardware and Training Time

All experiments were conducted on a laptop equipped with an NVIDIA RTX 4070 GPU (8 GB VRAM), an Intel i7-12800HX CPU, and 32 GB of system memory. Training for 300 epochs required approximately 45,600 seconds.

Code Availability

The implementation of our model, including pre-training and fine-tuning procedures, is publicly available at: <https://github.com/Giovanni-Sforza/MaskPoint-AMPT>.

A.1 Justification of Pre-Training Design Choices

Here, we elaborate on the key design choices made during the self-supervised pre-training phase.

A.1.1 Discrimination vs. Reconstruction Task

In the domain of image processing, masked autoencoders often follow a reconstruction paradigm: they apply patch-wise masking to an input image, retain the positional encodings of the masked patches, and then use a decoder to reconstruct the RGB values of the masked pixels from the visible regions [64]. The underlying assumption is that the geometric position of a pixel is known framework information, while its color is the core information to be modeled.

However, this paradigm is ill-suited for relativistic heavy-ion collisions. In this context, the momentum-space coordinates of the final-state particles constitute the core physical information, as their distribution is directly linked to the thermodynamic properties of the collision. A direct application of the image reconstruction paradigm would allow the decoder to simply copy the positional encodings to reconstruct the momentum coordinates, causing the model to degenerate into a trivial memorization task without learning deep physical features.

To overcome this challenge, we adopted a discrimination-based pre-training scheme [32]. As illustrated in Figure 1, this approach tasks the model with distinguishing between the "real" masked

Table 3: Ablation studies: (1) Effect of PointNet preprocessing; (2) Effect of different masking ratios.

Experiment	Setting	Accuracy
PointNet Preprocessing	With PointNet	0.9775
	Without PointNet	0.7303
Masking Ratio	25%	0.9775
	50%	0.7853
	75%	0.7232

Table 4: Number of parameters for different models. Note that while the pretraining stage includes a larger model due to the decoder, the finetuned model has a parameter count comparable to the PointNet baseline. The observed performance gains thus primarily arise from the two-stage pretraining and finetuning strategy rather than model scale.

Model	Stage	Total Parameters	Trainable Parameters
encoder + decoder	Pretrain	2.32M	1.16M
encoder + downstream head	Finetune	1.07M	1.07M
PointNet	Baseline	0.74M	0.74M

particles and randomly sampled "fake" particles. This forces the encoder to learn the underlying distribution and correlations within the particle point cloud to successfully identify what constitutes a physically plausible particle within the context of the event, thereby learning a much richer and more meaningful feature representation.

A.1.2 Sampling Strategy: Farthest Point Sampling

We chose Farthest Point Sampling (FPS) over a uniform random sampling strategy to select the 25% of points to be masked. FPS iteratively selects points that are maximally distant from the set of points already chosen. This ensures that the unmasked points provide a more uniform coverage of the entire point cloud's geometric structure.

A.1.3 Dataset and Training Procedure

Both the self-supervised pre-training and the supervised fine-tuning stages utilized the entire available AMPT-generated dataset. The models were trained for a maximum of 300 epochs. We employed a checkpointing strategy, saving the model weights after each epoch if the performance on a validation set improved. The final models used for analysis correspond to the epoch with the best validation performance.