

An end-to-end generative diffusion model for heavy-ion collisions

Jing-An Sun^{1,2}, Li Yan^{1,3,*}, Charles Gale², and Sangyong Jeon²

¹Institute of Modern Physics, Fudan University, Shanghai, 200433, China

²Department of Physics, McGill University, Montreal, Quebec H3A 2T8, Canada

³Key Laboratory of Nuclear Physics and Ion-beam Application (MOE), Fudan University, Shanghai 200433, China

Abstract. Heavy-ion collision physics has entered the high precision era, demanding theoretical models capable of generating huge statistics to compare with experimental data. However, traditional hybrid models, which combine hydrodynamics and hadronic transport, are computationally intensive, creating a significant bottleneck. In this work, we introduce DiffHIC, an end-to-end generative diffusion model, to emulate ultra-relativistic heavy-ion collisions. The model takes initial entropy density profiles and transport coefficients as input and directly generates two-dimensional final-state particle spectra. Our results demonstrate that DiffHIC achieves a computational speedup of approximately 10^5 against traditional simulations, while accurately reproducing a wide range of physical observables, including integrated and differential anisotropic flow, multi-particle correlations, and momentum fluctuations. This framework provides a powerful and efficient tool for phenomenological studies in the high-precision era of heavy-ion physics.

1 Introduction

The primary goal of ultra-relativistic heavy-ion collisions conducted at the LHC and RHIC is to create and study the quark-gluon plasma (QGP), a state of matter where quarks and gluons are deconfined [1, 2]. Hydrodynamic models [3–8] have been remarkably successful in describing the collective, fluid-like behavior of the QGP. State-of-the-art simulations typically employ a hybrid approach, coupling initial state models (e.g., TRENTO) with viscous hydrodynamic evolution (e.g., MUSIC), followed by particlization and hadronic transport (e.g., UrQMD).

Despite their success, these traditional models face a severe computational challenge. A single collision event simulation can take approximately 120 minutes on a single CPU core. This makes it computationally prohibitive to achieve the large statistics ($10^9 - 10^{10}$ events) required to study the hyperfine structure of QGP or perform high-precision analyses, such as nuclear structure effects, or the QGP's equation of state.

To overcome this limitation, we leverage machine learning, specifically generative artificial intelligence. We present the first application of a conditional diffusion model [9–11], named DiffHIC (Diffusion Model for Heavy-Ion Collisions), as a fast, end-to-end surrogate

*e-mail: cliyan@fudan.edu.cn

model for heavy-ion collision simulations. This approach aims to learn the complex mapping from initial conditions to final particle spectra, a process governed by non-linear hydrodynamic and Boltzmann transport equations.

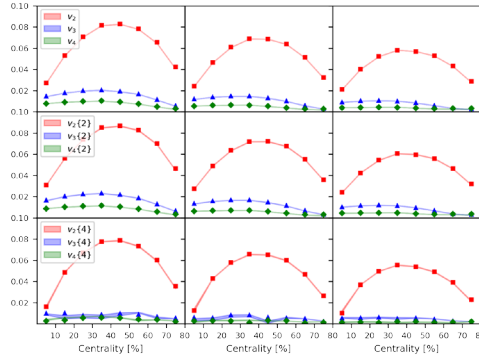


Figure 1. Centrality dependence of integrated anisotropic flow. Filled symbols represent the ground truth from the hybrid model, while colored bands show the results from DiffHIC. The agreement is excellent across different orders (n), cumulants, and shear viscosities.

2 Methodology

2.1 The Generative Diffusion Model

Diffusion models [11] are a class of generative models that learn a data distribution by reversing a gradual noising process. The process is defined by a forward stochastic differential equation (SDE) that transforms data into a simple prior (e.g., Gaussian noise), and a corresponding reverse-time SDE that transforms the prior back into the data distribution. The key to the reverse process is learning the score function, $\nabla_x \log p_t(x)$, which can be approximated by a neural network trained via denoising score matching.

Our model, DiffHIC, is a conditional diffusion model. It is trained to generate the final charged particle momentum spectra \mathbf{S} conditioned on the initial entropy density profile \mathbf{I} and the shear viscosity to entropy density ratio, η/s . The core of DiffHIC is a U-Net neural network architecture, which is widely adept in image-related tasks. The initial profile, conditioning parameters (η/s), and the diffusion timestep t are fed into the network to predict the noise added at that step. Once trained, the model can generate new particle spectra by solving the reverse ODE starting from random noise, a process that is both deterministic and computationally efficient.

2.2 Training and Validation Datasets

To train DiffHIC, we generated a large dataset of Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV using a well-established hybrid model. The initial entropy density profiles were generated using the TRENTO model. The subsequent (2+1)D viscous hydrodynamic evolution was simulated with MUSIC, followed by particlization via the Cooper-Frye formula and hadronic rescattering with UrQMD. We generated 12,000 minimum bias events for each of three distinct values of shear viscosity: $\eta/s = 0.0, 0.1, \text{ and } 0.2$. For validation, an additional 10,000 events per η/s value were generated across all centralities.

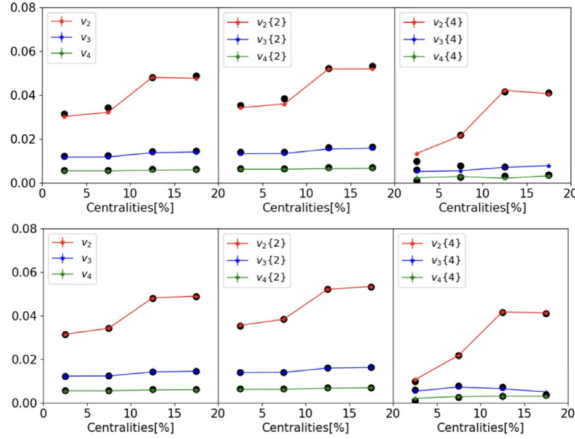


Figure 2. Symbols represent results from a hydrodynamic model, while lines show the output of a Generative AI model. The top row shows the pre-trained model, and the bottom row shows the improved results after fine-tuning with 500 new events.

3 Results and Performance

3.1 Computational Efficiency

The most significant advantage of DiffHIC is its computational efficiency. Generating a single event with DiffHIC takes approximately 0.1 seconds on a single NVIDIA GeForce GTX 4090 GPU. This represents a speedup factor of roughly 10^5 compared to the ~ 120 minutes required by the traditional hybrid model on a single CPU. This acceleration transforms previously infeasible large-scale statistical studies into routine computations.

3.2 Physics Fidelity

To validate the physical accuracy of DiffHIC, we compared its generated observables with the ground truth from the traditional hybrid model.

- **Anisotropic Flow:** Anisotropic flow coefficients, v_n , are crucial for characterizing the collective dynamics. As shown in Fig. 1, DiffHIC accurately reproduces the centrality dependence of integrated flow (v_2, v_3, v_4) calculated from both the event-plane method and multi-particle cumulants ($v_n\{2\}, v_n\{4\}$). The colored bands (DiffHIC) show excellent agreement with the filled symbols (ground truth) across all centralities and η/s values.
- **Comprehensive Comparison:** We conducted a comprehensive comparison across a wide range of observables, including pT-dependent flow, mean transverse momentum fluctuations, and mixed harmonic correlations. As presented in Ref. [12], the ratio of the observables generated by DiffHIC to the ground truth values are consistently close to unity, indicating the model’s high fidelity.
- **Flexibility and Fine-Tuning:** A key feature of our approach is its adaptability. We demonstrated that a pre-trained DiffHIC model can be quickly fine-tuned with a small dataset (e.g., 500 events) to incorporate new physics, such as the effects of nuclear deformation in Uranium-Uranium collisions, as shown in Figure 2. This highlights the model’s potential as a flexible tool for exploring a wide parameter space without retraining from scratch.

4 Conclusion and Outlook

We have developed an end-to-end generative diffusion model, DiffHIC, for simulating heavy-ion collisions. This model achieves a remarkable 10^5 -fold speedup over traditional hybrid simulations while maintaining high physical fidelity across a lot of observables. DiffHIC effectively learns the intricate physics of QGP evolution and hadronization, providing a powerful and efficient framework for phenomenological studies.

Future work will focus on extending the model's capabilities. This includes incorporating more physical parameters (e.g., bulk viscosity, equation of state), moving towards 3D hydrodynamic simulations, and generating full particle clouds for greater kinematic detail. The success of DiffHIC paves the way for a new generation of fast, AI-driven simulators to meet the challenges of the high-precision era in heavy-ion physics.

References

- [1] E. Shuryak, Strongly coupled quark-gluon plasma in heavy ion collisions, *Rev. Mod. Phys.* **89**, 035001 (2017), 1412.8393. [10.1103/RevModPhys.89.035001](https://doi.org/10.1103/RevModPhys.89.035001)
- [2] U. Heinz, R. Snellings, Collective flow and viscosity in relativistic heavy-ion collisions, *Ann. Rev. Nucl. Part. Sci.* **63**, 123 (2013), 1301.2826. [10.1146/annurev-nucl-102212-170540](https://doi.org/10.1146/annurev-nucl-102212-170540)
- [3] L. Yan, A flow paradigm in heavy-ion collisions, *Chin. Phys. C* **42**, 042001 (2018), 1712.04580. [10.1088/1674-1137/42/4/042001](https://doi.org/10.1088/1674-1137/42/4/042001)
- [4] C. Shen, L. Yan, Recent development of hydrodynamic modeling in heavy-ion collisions, *Nucl. Sci. Tech.* **31**, 122 (2020), 2010.12377. [10.1007/s41365-020-00829-z](https://doi.org/10.1007/s41365-020-00829-z)
- [5] B. Schenke, S. Jeon, C. Gale, Elliptic and triangular flow in event-by-event (3+1)D viscous hydrodynamics, *Phys. Rev. Lett.* **106**, 042301 (2011), 1009.3244. [10.1103/PhysRevLett.106.042301](https://doi.org/10.1103/PhysRevLett.106.042301)
- [6] B. Schenke, S. Jeon, C. Gale, (3+1)D hydrodynamic simulation of relativistic heavy-ion collisions, *Phys. Rev. C* **82**, 014903 (2010), 1004.1408. [10.1103/PhysRevC.82.014903](https://doi.org/10.1103/PhysRevC.82.014903)
- [7] J.F. Paquet, C. Shen, G.S. Denicol, M. Luzum, B. Schenke, S. Jeon, C. Gale, Production of photons in relativistic heavy-ion collisions, *Phys. Rev. C* **93**, 044906 (2016), 1509.06738. [10.1103/PhysRevC.93.044906](https://doi.org/10.1103/PhysRevC.93.044906)
- [8] C. Gale, J.F. Paquet, B. Schenke, C. Shen, Multimessenger heavy-ion collision physics, *Phys. Rev. C* **105**, 014909 (2022), 2106.11216. [10.1103/PhysRevC.105.014909](https://doi.org/10.1103/PhysRevC.105.014909)
- [9] J. Ho, A. Jain, P. Abbeel, Denoising diffusion probabilistic models, in *Proceedings of the 34th International Conference on Neural Information Processing Systems* (Curran Associates Inc., Red Hook, NY, USA, 2020), NIPS '20, ISBN 9781713829546, <https://proceedings.neurips.cc/paper/2020/file/4c5bcfec8584af0d967f1ab10179ca4b-Paper.pdf>
- [10] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, S. Ganguli, Deep unsupervised learning using nonequilibrium thermodynamics, in *International conference on machine learning* (PMLR, 2015), pp. 2256–2265, <https://proceedings.mlr.press/v37/sohl-dickstein15.pdf>
- [11] Y. Song, J. Sohl-Dickstein, D.P. Kingma, A. Kumar, S. Ermon, B. Poole, Score-Based Generative Modeling through Stochastic Differential Equations, in *International Conference on Learning Representations* (2021), <https://openreview.net/forum?id=PXTIG12RRHS>
- [12] J.A. Sun, L. Yan, C. Gale, S. Jeon, An end-to-end generative diffusion model for heavy-ion collisions (2024), 2410.13069, <https://arxiv.org/abs/2410.13069>