

---

# $\rho$ -Diffusion: A diffusion-based density estimation framework for computational physics

---

**Maxwell X. Cai**  
Intel Corporation  
maxwell.cai@intel.com

**Kin Long Kelvin Lee**  
Intel Corporation  
kin.long.kelvin.lee@intel.com

## Abstract

In physics, density  $\rho(\cdot)$  is a fundamentally important scalar function to model, since it describes a scalar field or a probability density function that governs a physical process. Modeling  $\rho(\cdot)$  typically scales poorly with parameter space, however, and quickly becomes prohibitively difficult and computationally expensive. One promising avenue to bypass this is to leverage the capabilities of denoising diffusion models often used in high-fidelity image generation to parameterize  $\rho(\cdot)$  from existing scientific data, from which new samples can be trivially sampled from. In this paper, we propose  $\rho$ -DIFFUSION, an implementation of denoising diffusion probabilistic models for multidimensional density estimation in physics, which is currently in active development and, from our results, performs well on physically motivated 2D and 3D density functions. Moreover, we propose a novel hashing technique that allows  $\rho$ -DIFFUSION to be conditioned by arbitrary amounts of physical parameters of interest.

## 1 Introduction

High-quality scientific data—the foundation of all scientific discovery—is difficult and expensive to accumulate. The time and energy cost to design, implement, and execute state-of-the-art experiments (*in silico* and *ex silico*) necessary for many domains remains barely surmountable for all but select problems, further gated behind access to the required resources such as equipment, human expertise, compute, and so on. A significant amount of scientific research could be enabled and/or accelerated if the availability of critical data was improved.

At risk of being overly reductive, a significant amount of scientific research pertains to modeling probability densities,  $\rho(\cdot)$ ; from a probabilistic perspective, by estimating or learning distributions of parameters that govern some physical process, new data and inferences can in principle be obtained through posterior sampling. Traditional density estimators, such as histograms, kernel density estimators (KDEs), and Markov chain Monte Carlo (MCMC) only perform well in low dimensions. But the recent renaissance of generative modeling with deep neural networks has spawned a plethora of approaches ranging from variational inference [1] to adversarial learning [2] to normalizing flows [3, 4] to diffusion models [5, 6]. While most of these methods have demonstrated in various capacities both the potential and ability to be applied to the physical sciences (see e.g., [7, 8, 9, 10, 11]), much of the research focus has remained in select applications particularly in the 2D image domain or toy datasets. Applications of denoising diffusion probabilistic models (DDPM) in physics in particular has seen success in several areas (e.g., [12, 13, 14, 15]).

In this paper, we propose a DDPM-based density modeling framework that provides scientific domain-centered abstractions for conditional data generation, with the primary goal being to lower the barrier for adopting generative practices to scientific workflows. As a proof-of-concept, we demonstrate cross-domain adaptability with data generation in astrophysics and chemistry for 2D/3D data.

## 2 Methods

**Denosing Diffusion Probabilistic Models** Inspired by nonequilibrium thermodynamics, [5] showed that the gradual and systematic noise injection to input data can be considered a Markov process. Given  $\rho^{(t)}(\cdot)$  as the density function of interest at diffusion timestep  $t$ , by using a tractable Gaussian  $\mathcal{N}(\rho^{(t)}; \sqrt{1 - \beta^{(t)}}\rho^{(t)}, \beta_t \mathbf{I})$  as the diffusion kernel, the diffusion trajectory is deterministic, and with a sufficiently large number of steps  $T$ ,  $\rho^{(T)} \rightarrow \mathcal{N}(\mathbf{0}, \mathbf{I})$ . Conversely, the data generation process is elegantly translated into learning a reverse trajectory that starts from  $\mathcal{N}(\rho^{(T)}; \mathbf{0}, \mathbf{I})$  and ends at  $\rho^{(0)}(\cdot)$ . Notably, if the schedule steps size  $\delta\beta := \beta_t - \beta_{t-1}$  is sufficiently small, then the reverse kernel will have the same functional form as the diffusion kernel, i.e., Gaussian. Therefore, the neural network only has to predict the correct time-dependent mean of the reverse Gaussian kernel (the variance of which can be computed directly from  $\beta^{(t)}$ ). Eventually,  $\rho^{(0)}$  can be sampled from pure Gaussian noise through a Langevin-like reverse process [16]. New data that is not part of the original training data can be obtained by interpolating the training data [6]. For details, we refer the readers to the original papers [5, 6].

**Architecture** A neural network that learns  $\mu_\theta(\rho^{(t)}, t)$  must be able to accept  $n$ -dimensional  $\rho^{(t)}(\cdot)$  at an arbitrary timestep  $t$ . We follow the practice in e.g., [6, 17] to use a U-Net [18] for the parameterization, which is enhanced with residual [19] building blocks, multi-head attention [20], and a Transformer-style sinusoidal position embedding to allow parameters to be shared across time. Additionally, thanks to the close form of  $q(\rho^{(t)}|\rho^{(0)}) = \mathcal{N}(\rho^{(t)}; \sqrt{\bar{\alpha}^{(t)}}\rho^{(0)}, (1 - \bar{\alpha}^{(t)})\mathbf{I})$  where  $\bar{\alpha}^{(t)} := \prod_{s=1}^t (1 - \beta^{(s)})$ , it is possible to obtain a noised sample  $\rho^{(t)}$  directly from  $\rho^{(0)}$  [6]. Therefore, instead of looping over  $t$ , training can be accelerated in a (distributed) data-parallel fashion.

**Implementation** We are aware of the existence of many open-source DDPM implementations. While most existing implementations are designed for 2D images, a few (e.g., DIFFUSERS from HuggingFace and GUIDEDDIFFUSION from OpenAI) have extended the API to 1D and 3D data. However, most of these implementations are optimized for tasks like image generation or text-to-image translation, and the underlying software architectures do not necessarily provide the flexibility to accommodate the domain-specific requirements and properties of physical data. As such, we eventually opted for a customized, physics-centric implementation from the stretch, which allows us to integrate physics-motivated models and inductive biases into the framework in the subsequent development. We generalized the DDPM backbone of  $\rho$ -DIFFUSION to  $n$ -dimensional, with a current limitation of  $n = \{1, 2, 3\}$ , which corresponds to sequential data, 2D PDFs/images, and 3D PDF/volumetric data, respectively. The model is implemented with PYTORCH, Intel<sup>®</sup> Extension for PyTorch (IPEX) and Intel<sup>®</sup> oneAPI Deep Neural Network Library (oneDNN). The model training and inference were performed using Intel<sup>®</sup> Data Center GPU Max Series.

**Conditioned data generation with physical parameters** In physics, generative models are particularly useful if the data generation process can be conditioned directly by the physical parameters of interest. In the DDPM case, it boils down to learning  $p(\rho^{(t)}|\rho^{(t-1)}, c)$ , where  $c$  is the condition. In the traditional class-conditioned DDPM image generation approach, the class labels are typically categorical, which are then passed to an embedding layer to calculate the embedding vector. Such an approach is inadequate to handle the scenario where  $\rho(\cdot)$  is a continuous and multivariate function. Instead, we allow the input variables of  $\rho(\cdot)$  to be expressed in a PYTHON dict structure and subsequently use the SHA512 algorithm to hash the parameter dict. The hexadecimal digest is then converted into a float vector, which serves as the embedding of params. The uniqueness of the embedding is guaranteed by the SHA512 algorithm, which is deterministic and has a hash collision rate of  $1/2^{512}$ , and the length of the embedding is fixed, independent of the number of parameters given. This makes it possible for  $\rho$ -DIFFUSION to be conditioned by an arbitrary amount of physical parameters of interest.

This hash-based embedding approach is computationally efficient, and by virtue of being non-parametric, does not require domain adaptation and can be used for arbitrary formats. Its main weakness however is that it is not possible to interpolate the model directly in the physical parameter space: while the avalanche effect is a desirable property for cryptography, minuscule changes in the original parameters will result in a completely different embedding. In the case of  $\rho$ -DIFFUSION,

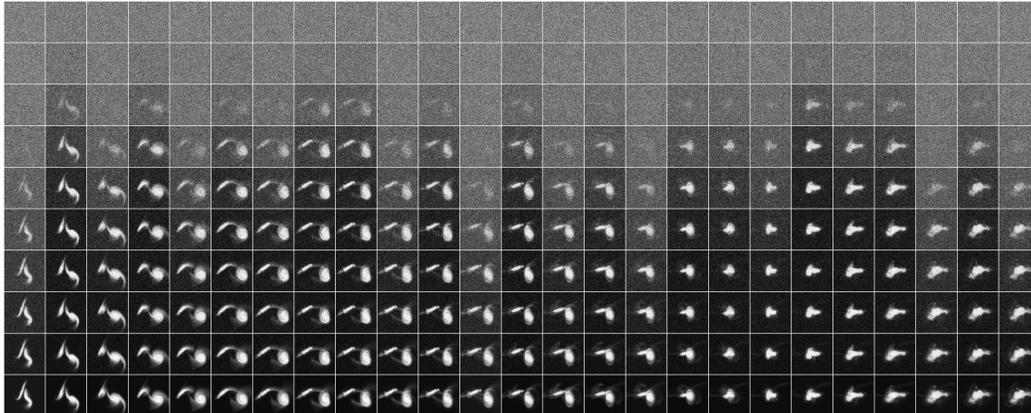


Figure 1: Generated galaxy merger image samples (128x128 px). From top to down: the denoising process at  $t = [900, 800, \dots, 100, 0]$ , respectively; From left to right: the merging process from initial approaching to tidal disruption to collision, where each column corresponds to a snapshot  $5 \times 10^6$  years apart from the adjacent columns in physical timescale.

ideally the embedding function is not only differentiable, but also maps perturbations from physical parameters to the embedding space in a physically meaningful capacity. A possible improvement in future work—inspired by CLIP[21]—is to use train a language model jointly with the DDPM, along with the physical parameters as the guidance to condition the  $\rho(\cdot)$  sampling.

### 3 Experiments

In the following experiments, we demonstrate that  $\rho$ -DIFFUSION works on both traditional image data and real physical data (with proper normalization, if necessary).

So far, we have shown success in applying our workflow to 2D and 3D data: the combination of UNet employing either 2D and 3D convolutions with Langevin sampling can produce conditional densities with high fidelity and accuracy. However, we see the approach break down for the 1D case of pure rotational spectra, which correspond to observable transitions between quantized rotational energy levels of molecules [22]. This demonstrates a number of complexities not captured: notably, regions of high “density” or signal are significantly more sparse in high-resolution molecular spectra than the other data covered here and likely require architectures that recover long-range correlations.

**2D Data: Images of galaxy merger simulations** Galaxy mergers are fundamentally important processes driven through the history of galactic evolution [23, 24]. In this process, the two colliding galaxies are tidally disrupted, creating a time-evolving morphology. It is computationally expensive to simulate galaxy mergers. In this experiment, we use a Barnes-Hut tree [25] code BONSAI [26] to simulate the dynamical process with a grid of parameters and subsequently generate 2D images from the visualization simulation snapshots. The 2D images are then used to train  $\rho$ -DIFFUSION, in which the physical parameters used to run BONSAI are used to condition the sampling. Figure 1 shows that by varying the time parameter while keeping the others fixed,  $\rho$ -DIFFUSION can generate the merger process directly, thus bypassing the need to carry out expensive simulations.

**3D data: Volumes of spherical harmonics** Spherical harmonics are a mathematical basis with diverse applications, ranging from molecule spectroscopy [22] and quantum chemistry [27] to astronomy [28]. In the context of density estimation, spherical harmonics are an attractive source of data to generate and test arbitrarily complex and structured two and three-dimensional densities; in spherical coordinates, each spherical harmonic  $Y$  is defined or “conditioned” by two parameters, the order  $m$  and degree  $l$ :

$$Y_l^m(\theta, \phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} e^{im\theta} P_l^m(\cos(\phi)) \quad (1)$$

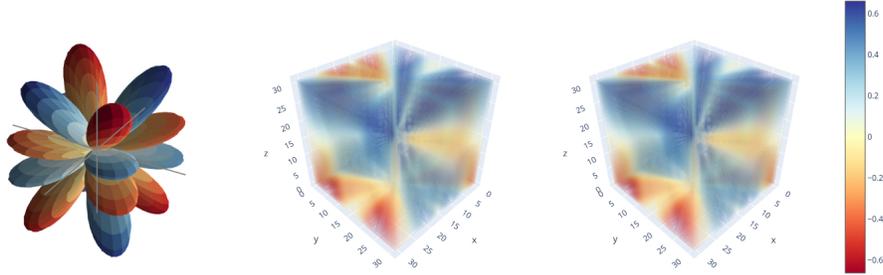


Figure 2: Left: A surface plot of  $Y_5^2(\theta, \phi)$ ; Center: the voxelized ground truth volume; Right: the generated volume. The real part is color-coded, and all three subplots share the same colorbar. The generated volume is visually indistinguishable from the ground truth, and has a Wasserstein distance of  $\sim 0.035$  from the ground truth.

with associated Legendre polynomial  $P_l^m$ , and azimuthal and polar coordinates  $\theta$  and  $\phi$ . Larger values of  $l$  and  $m$  progressively introduce symmetric nodal structure, which may prove increasingly difficult for models to reproduce. To construct these basis functions as 3D objects, we rely on their implementation in NUMPY [29] and SCIPY [30] and compute them on linearly spaced cartesian grids, with the value at each point corresponding to the real part of  $Y_l^m(x, y, z \rightarrow \theta, \phi)$ . We compute the spherical harmonics for different combinations of  $(m, l)$  and voxelize them to volumes of  $32^3$ . Figure 2 shows a non-cherry-picked generated sample of  $Y_5^2(\theta, \phi)$ .

## 4 Discussions and Conclusions

DDPM models are known for their capability of high-quality image density synthesis. In this project, we extend this capability to physics-based density estimation tasks. Theoretically, the DDPM framework places no limits on the types of density functions. In practice, however, we observe several difficulties when training DDPM to learn  $\rho(\cdot)$  on physical data due to their unique properties:

- Large temporal ranges: Sequential data, especially time series of a dynamical system, may exhibit long-range dependencies. Despite the enhancements on the U-Net with attention blocks and residual networks, the model may not be informed about the global structure of the sequence due to the small convolutional kernel sizes.
- Large spatial ranges: in many fields in physics, especially in astrophysics, the range of the data is common to cross multiple orders of magnitudes. In the DDPM framework, the use of  $\mathcal{N}(\rho^{(t)}; \sqrt{1 - \beta_t}\mathbf{x}^{(t-1)}, \beta_t\mathbf{I})$  as a diffusion kernel has essentially required the data to be scaled to a small interval, ideally  $[-1, 1]$ . However, squeezing raw physical data into such a range will sacrifice the model’s resolution on small details.
- Lack of inductive biases: Although diffusion models are inspired by nonequilibrium thermodynamics, the DDPM framework does not necessarily place physical constraints on an architectural level. This may cause slower convergence and limitations on the model’s capability to understand the physical laws behind the data.
- Expensive sampling process: Unlike other generative models, sampling DDPM is an iterative process that typically requires  $\sim 10^3$  steps of model inference. The nonlinear behavior in physical data may require even more steps due to the fluctuating gradients. We expect this issue to be alleviated with improved sampling algorithms (e.g., DDIM sampling[31]).

Despite the challenges, we demonstrate that under certain conditions, DDPM-based models are capable of capturing the fine details of some density functions in physics. As a proof-of-concept, we propose  $\rho$ -DIFFUSION as a framework to handle physical data of different dimensionalities with a unified application programming interface (API), and provide a workflow for training and inference. The framework is currently in active development<sup>1</sup> and contributions are welcome.

<sup>1</sup>Source code: <https://github.com/intel/rho-diffusion>

## References

- [1] Diederik P Kingma and Max Welling. Auto-Encoding Variational Bayes. *arXiv e-prints*, page arXiv:1312.6114, December 2013.
- [2] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative Adversarial Networks. *arXiv e-prints*, page arXiv:1406.2661, June 2014.
- [3] Danilo Jimenez Rezende and Shakir Mohamed. Variational Inference with Normalizing Flows. *arXiv e-prints*, page arXiv:1505.05770, May 2015.
- [4] Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud. Neural Ordinary Differential Equations. *arXiv e-prints*, page arXiv:1806.07366, June 2018.
- [5] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In Francis Bach and David Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 2256–2265, Lille, France, 07–09 Jul 2015. PMLR.
- [6] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising Diffusion Probabilistic Models. *arXiv e-prints*, page arXiv:2006.11239, June 2020.
- [7] Mathieu Germain, Karol Gregor, Iain Murray, and Hugo Larochelle. MADE: Masked Autoencoder for Distribution Estimation. *arXiv e-prints*, page arXiv:1502.03509, February 2015.
- [8] Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. Density estimation using Real NVP. *arXiv e-prints*, page arXiv:1605.08803, May 2016.
- [9] George Papamakarios, Theo Pavlakou, and Iain Murray. Masked Autoregressive Flow for Density Estimation. *arXiv e-prints*, page arXiv:1705.07057, May 2017.
- [10] Qiao Liu, Jiaye Xu, Rui Jiang, and Wing Hung Wong. Density estimation using deep generative neural networks. *Proceedings of the National Academy of Sciences*, 118(15):e2101344118, 2021.
- [11] Will Grathwohl, Ricky T. Q. Chen, Jesse Bettencourt, Ilya Sutskever, and David Duvenaud. FFJORD: Free-form Continuous Dynamics for Scalable Reversible Generative Models. *arXiv e-prints*, page arXiv:1810.01367, October 2018.
- [12] Yihang Wang, Lukas Herron, and Pratyush Tiwary. From data to noise to data for mixing physics across temperatures with generative artificial intelligence. *Proceedings of the National Academy of Science*, 119(32):e2203656119, August 2022.
- [13] Michael J. Smith, James E. Geach, Ryan A. Jackson, Nikhil Arora, Connor Stone, and Stéphane Courteau. Realistic galaxy image simulation via score-based generative models. *MNRAS*, 511(2):1808–1818, April 2022.
- [14] Dule Shu, Zijie Li, and Amir Barati Farimani. A physics-informed diffusion model for high-fidelity flow field reconstruction. *Journal of Computational Physics*, 478:111972, April 2023.
- [15] Rui Li, Gabriel della Maggiora, Vardan Andriasyan, Anthony Petkidis, Artsemi Yushkevich, Mikhail Kudryashev, and Artur Yakimovich. Microscopy image reconstruction with physics-informed denoising diffusion probabilistic model. *arXiv e-prints*, page arXiv:2306.02929, June 2023.
- [16] Max Welling and Yee Whye Teh. Bayesian learning via stochastic gradient langevin dynamics. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ICML’11, page 681–688, Madison, WI, USA, 2011. Omnipress.
- [17] Prafulla Dhariwal and Alex Nichol. Diffusion Models Beat GANs on Image Synthesis. *arXiv e-prints*, page arXiv:2105.05233, May 2021.
- [18] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv e-prints*, page arXiv:1505.04597, May 2015.
- [19] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. *arXiv e-prints*, page arXiv:1512.03385, December 2015.
- [20] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention Is All You Need. *arXiv e-prints*, page arXiv:1706.03762, June 2017.

- [21] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning Transferable Visual Models From Natural Language Supervision. *arXiv e-prints*, page arXiv:2103.00020, February 2021.
- [22] Charles H. Townes and Arthur L. Schawlow. *Microwave Spectroscopy*. Dover Publications, New York, 1975.
- [23] Alar Toomre. Mergers and Some Consequences. In Beatrice M. Tinsley and D. Campbell Larson, Richard B. Gehret, editors, *Evolution of Galaxies and Stellar Populations*, page 401, January 1977.
- [24] Joshua E. Barnes and Lars Hernquist. Dynamics of interacting galaxies. *Annual Review of Astronomy and Astrophysics*, 30:705–742, January 1992.
- [25] Josh Barnes and Piet Hut. A hierarchical  $O(N \log N)$  force-calculation algorithm. *Nature*, 324(6096):446–449, December 1986.
- [26] Jeroen Bédorf, Evghenii Gaburov, and Simon Portegies Zwart. A sparse octree gravitational N-body code that runs entirely on the GPU processor. *Journal of Computational Physics*, 231(7):2825–2839, April 2012.
- [27] H. Bernhard Schlegel and Michael J. Frisch. Transformation between Cartesian and pure spherical harmonic Gaussians. *International Journal of Quantum Chemistry*, 54(2):83–87, 1995.
- [28] L. Brett. Methods of Spherical Harmonic Analysis. *Quarterly Journal of the Royal Astronomical Society*, 29:129, June 1988.
- [29] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, September 2020.
- [30] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, 17:261–272, 2020.
- [31] Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising Diffusion Implicit Models. *arXiv e-prints*, page arXiv:2010.02502, October 2020.