
Physics-Informed Neural Networks for Modeling Galactic Gravitational Potentials

Charlotte Myers¹
c_myers@mit.edu

Nathaniel Starkman^{1,2,3}
starkman@mit.edu

Lina Necib^{1,4}
lnebib@mit.edu

¹**MIT Department of Physics and MIT Kavli Institute for Astrophysics and Space Research,**
Massachusetts Institute of Technology, Cambridge, MA 02139, USA

²Brinson Prize Fellow

³Visiting Scholar, Case Western Reserve University

⁴The NSF AI Institute for Artificial Intelligence and Fundamental Interactions,
Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract

We introduce a physics-informed neural framework for modeling static and time-dependent galactic gravitational potentials. The method combines data-driven learning with embedded physical constraints to capture complex, small-scale features while preserving global physical consistency. We quantify uncertainty through a Bayesian framework, and model time evolution using a neural ODE approach. Applied to mock systems of varying complexity, the model achieves reconstruction errors at the sub-percent level (0.14% mean acceleration error) and improves dynamical consistency compared to analytic baselines.

1 Introduction

Modeling the gravitational potential of a galaxy is a central problem in astrophysics, providing the link between the observed motions of stars and gas to the underlying distribution of mass [1]. Since most of this mass is in the form of dark matter – interacting only via its gravitational influence – potential models are among the few probes of its distribution and properties, while also serving as a foundation for modeling a galaxy’s small-scale structure and dynamical evolution [2, 3].

Existing gravity modeling methods span a trade-off between interpretability, flexibility, and efficiency. Analytic models employ interpretable, physically motivated forms (e.g. NFW halos), but struggle to model substructures and deviations from the idealized form [4]. Basis function expansions (BFEs) trade for increased flexibility by representing the potential as a sum over modes. However, for complex systems they can be inefficient due to requiring a large number of modes and misleading in their interpretation due to unphysical terms (e.g. negative mass components) and representation of noise by high-order terms [5].

Physics-informed neural networks (PINNs) address many of these limitations by generalizing the concept of BFEs, learning flexible and adaptive basis functions directly from data while embedding physical constraints into the training objective. By constraining the hypothesis space to physically consistent functions, PINNs reduce the required parameter count and improve generalization compared to purely empirical networks [6, 7]. Recent advances, particularly the PINN-GM-III architecture, have achieved state-of-the-art performance in reconstructing gravitational fields of terrestrial objects with high accuracy and robust extrapolation [8].

In this work we make three main contributions: (1) we extend the PINN-GM-III framework to model galactic potentials, where larger spatial scales, diverse structural components, and non-axisymmetric

perturbations pose new challenges; (2) we introduce a time-dependent neural ordinary differential equation (NODE) formulation that captures dynamical evolution; (3) we incorporate Bayesian inference to quantify uncertainty in both static and evolving fields.

2 Model design

Our model combines analytic structure with learned residuals through six ingredients: (i) a loss that enforces the physical relation $\mathbf{a} = -\nabla\phi$; (ii) a coordinate transform that compactifies the spatial domain; (iii) radial scaling and (iv) analytic fusing, which factor out the dominant large-scale trends and reserve capacity for higher-order perturbations; (v) a Bayesian treatment to capture epistemic uncertainty and to infer analytic scale parameters jointly with the residual field; and (vi) a neural ODE formulation to constrain temporal evolution to a continuous, causal trajectory.

2.1 Architecture

The model accepts N paired vectors of positions and accelerations, $(\mathbf{x}_i, \mathbf{a}_i)$, sampled from the target gravitational field. The network, parameterized by $\boldsymbol{\theta}$, outputs the predicted potential $\phi(\mathbf{x}_i)$, which is differentiated with respect to position to produce an acceleration vector. The physics-informed loss \mathcal{L} enforces the constraint $\mathbf{a} = -\nabla\phi$ by combining absolute and relative acceleration errors:

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{N} \sum_{i=1}^N \left(\|\mathbf{a}_i - \nabla\phi(\mathbf{x}_i|\boldsymbol{\theta})\| + \lambda_r \frac{\|\mathbf{a}_i - \nabla\phi(\mathbf{x}_i|\boldsymbol{\theta})\|}{\|\mathbf{a}_i\|} \right). \quad (1)$$

The relative term, weighted by λ_r , prevents loss of accuracy at large radii where accelerations decay toward zero. Because the loss is written in terms of the gradient of the learned potential, each optimization step explicitly penalizes violations of the force–potential relation $\mathbf{a} = -\nabla\phi$, tying model capacity to physically consistent solutions.

2.2 Input representation

We adopt the 5D spherical coordinate system of Martin and Schaub [8], which maps the domain to a finite range to improve numerical stability. In this parameterization, the three angular coordinates (s, t, u) specify direction, while two compactified radial coordinates (r_i, r_e) map the interior and exterior regions to bounded intervals. This construction keeps both ϕ and its spatial derivatives within a well-conditioned range across the domain, allowing the physics-informed loss in Eq. (1) to weight violations of $\mathbf{a} = -\nabla\phi$ consistently over all scales.

2.3 Design features: radial scaling and analytic fusing

Expanding on the design of Martin and Schaub [8], we factor out the analytic radial trend from the full potential $\phi(\mathbf{x})$ using a scaling function $n(\mathbf{x})$, and train the network to predict a scaled residual potential $\tilde{\phi}_{\text{NN}}(\mathbf{x}) = \phi(\mathbf{x}) \cdot n(\mathbf{x})$. This compresses the dynamic range of the output, improving numerical conditioning of both ϕ and $\nabla\phi$ and directing the loss in Eq. (1) toward structure that deviates from the dominant radial fall-off.

We also incorporate an analytic baseline (AB) model $\phi_{\text{AB}}(\mathbf{x})$ that captures the known (generally large-scale, low-order) structure of the gravitational field. Rather than relearning these well-understood features, the network predicts only the scaled residual component $\tilde{\phi}_{\text{NN}}(\mathbf{x})$. The full potential is reconstructed as

$$\phi(\mathbf{x}) = \phi_{\text{AB}}(\mathbf{x}) + \tilde{\phi}_{\text{NN}}(\mathbf{x}|\boldsymbol{\theta})/n(\mathbf{x}). \quad (2)$$

This fusion leverages the interpretability and reliability of analytic modeling while reserving the network’s flexibility for capturing higher-order perturbations. Setting $\phi_{\text{AB}} \equiv 0$ recovers the scaling-only variant. In Figure 1, we show the effect of these design features on the reconstruction of the acceleration and potential fields of the Milky Way-LMC system (see Section 3 for details).

2.4 Bayesian framework

To quantify epistemic uncertainty, we adopt a Bayesian neural network (BNN) framework. Unlike conventional neural networks, which produce fixed parameter values, a BNN treats each network

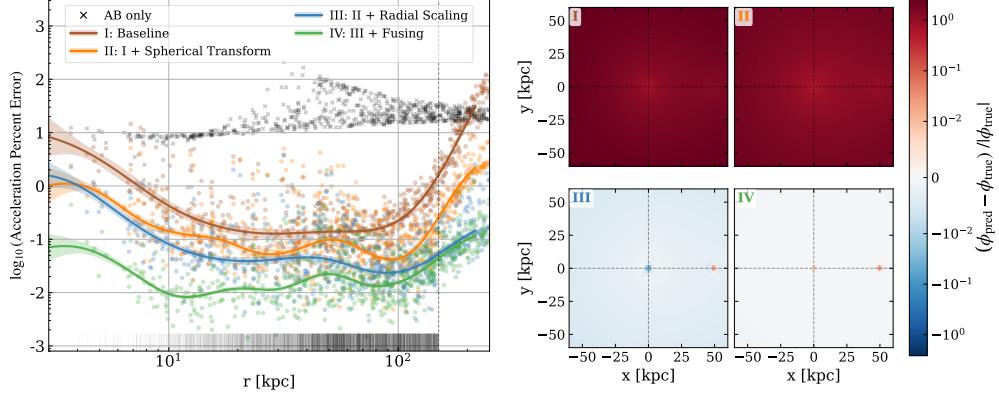


Figure 1: **Impact of model design choices on reconstructing the MW-LMC system.** **Left:** Radial profile of relative acceleration percent error ($100 \cdot \frac{\|\mathbf{a}_{\text{pred}} - \mathbf{a}_{\text{true}}\|}{\|\mathbf{a}_{\text{true}}\|}$). Black dashes mark the training point locations, and black crosses represent the error from an analytic MW model without NN correction. **Right:** Relative residual of the potential field in the x - y plane with $z = 0$.

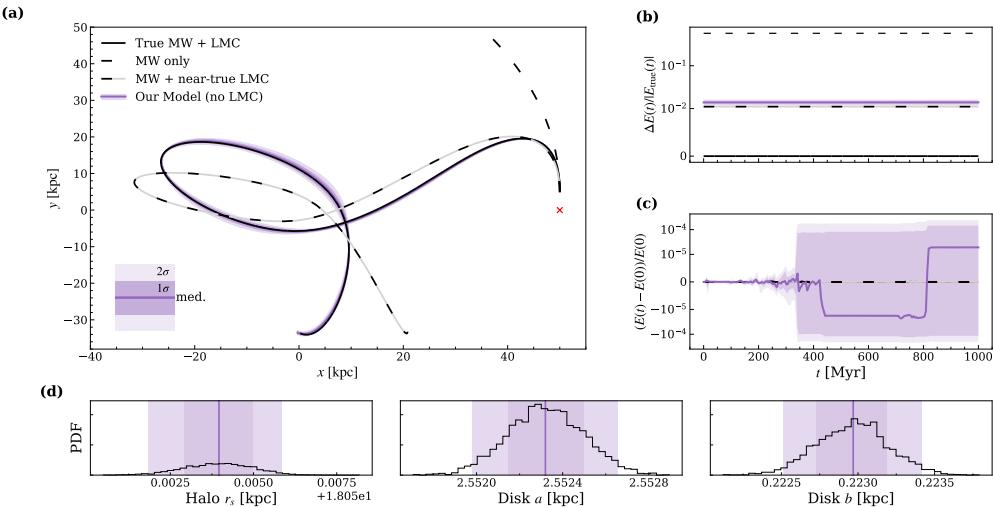


Figure 2: **Dynamics validation for the MW-LMC test system** (a) Orbit initialized at the LMC center with the local circular velocity and integrated for 1 Gyr under four models: the true potential, an analytic MW-only baseline without the LMC, an MW+LMC model with misspecified parameters (LMC center offset by 1 kpc and scale radius r_s misestimated by 2%), and the BNN reconstruction. (b) Relative deviation from true energy, where $\Delta E(t) = E_{\text{pred}} - E_{\text{true}}$. (c) Relative energy drift along the predicted orbit. (d) BNN posteriors of the analytic scale parameters. Uncertainty envelopes in all panels correspond to the 16–84th (1σ) and 2.5–97.5th (2σ) percentiles across 200 posterior draws.

weight and bias as a random variable with an associated probability distribution [9]. We implement this framework in NumPyro and approximate the posterior distributions via stochastic variational inference (SVI). We place truncated-normal priors on the analytic parameters α within physically plausible ranges, and zero-mean Gaussian priors with variance σ_λ^2 on the neural network weights θ . Accelerations are modeled with a Gaussian likelihood $p(a_{\text{obs}} | x, \alpha, \theta) = \mathcal{N}(a_{\text{pred}}(x; \alpha, \theta), \sigma_a^2)$ with fixed noise $\sigma_a = 2 \times 10^{-4}$. The variational posterior $q(\alpha, \theta)$ is a fully factorized Gaussian (AutoNormal) optimized by maximizing the ELBO, which enforces the physics-informed loss (1) through its likelihood term. To stabilize joint inference of analytic and residual components, training proceeds in two stages: an initial phase with a small weight prior ($\sigma_\lambda = 0.01$) that fits the analytic baseline, followed by a phase with tighter analytic priors and relaxed neural priors ($\sigma_\lambda = 1.0$) that allows the network to capture residual structure.

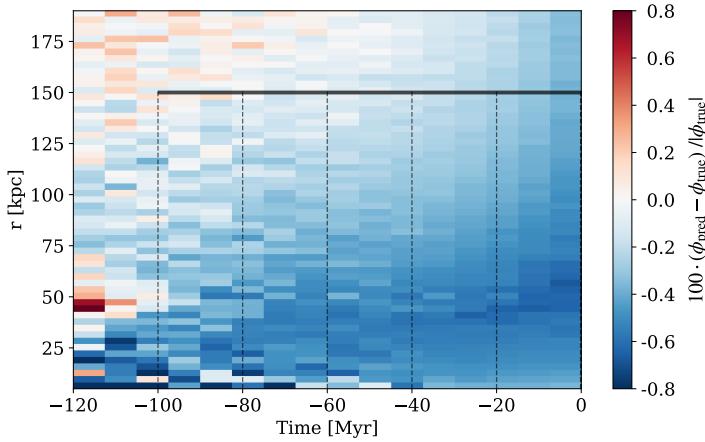


Figure 3: **Relative error of the time-dependent MW-LMC potential reconstruction**, binned by radius and time relative to the present ($t = 0$; $t < 0$ earlier). The dashed vertical lines mark training snapshot times; the horizontal line indicates the maximum training radius.

2.5 Modeling time dependence

We adopt a NODE formulation in which the network predicts the time derivative $f \equiv \frac{d}{dt} \tilde{\phi}_{\text{NN}}$ (parameterized by θ_2) of the scaled residual potential field and the initial spatial correction (parameterized by θ_1). The full potential is then reconstructed as:

$$\phi(\mathbf{x}, t) = \phi_{AB}(\mathbf{x}, t) + \frac{1}{n(\mathbf{x}, t)} \cdot \left[\tilde{\phi}_{\text{NN}}(\mathbf{x}, t | \theta) = \tilde{\phi}_{\text{NN}0}(\mathbf{x} | \theta_1) + \int_0^t f(\mathbf{x}, t' | \theta_2) dt' \right], \quad (3)$$

where $\tilde{\phi}_{\text{NN}0}(\mathbf{x} | \theta_1)$ denotes the initial correction at $t = 0$. The time integral is evaluated numerically via Gauss–Legendre quadrature.

A naive approach to modeling temporal variation is to append time as a fourth coordinate in the network input. However, this approach treats time as an unordered feature and does not enforce the causal structure of the evolution. By contrast, the NODE approach constrains the evolution to follow a continuous trajectory, ensuring temporal consistency across the learned evolution

2.6 Implementation details

We implement the network in JAX with optax for optimization; training uses the Adam optimizer with an initial learning rate of 3×10^{-3} and an exponential schedule that halves the rate every 1000 epochs. All experiments reported in Section 3 were run on a standard Apple M2 CPU; wall-clock training times are summarized alongside accuracy in Figure 4. We make scripts for data generation, training, and evaluation available on GitHub.

3 Performance

We assess the model through a series of controlled tests, the most challenging of which models the Large Magellanic Cloud (LMC) — the most massive satellite of the Milky Way (MW) — as a perturbation to an analytic baseline MW potential.

Static field reconstruction: The network is tasked with learning the LMC-induced perturbation as well as residual substructure omitted from the analytic baseline (e.g., the bulge and nucleus). The global scale parameters of the host potential—the halo scale radius r_s and the disk scale lengths a and b —are inferred within a Bayesian framework.

In this MW–LMC configuration (with the LMC placed 50 kpc from the Galactic center; [10]), we train a 4×128 dense network for 10,000 epochs using Adam optimization, and setting $\lambda_r = 0.1$. The training set consists of 4096 samples drawn via density-based rejection sampling to reflect the

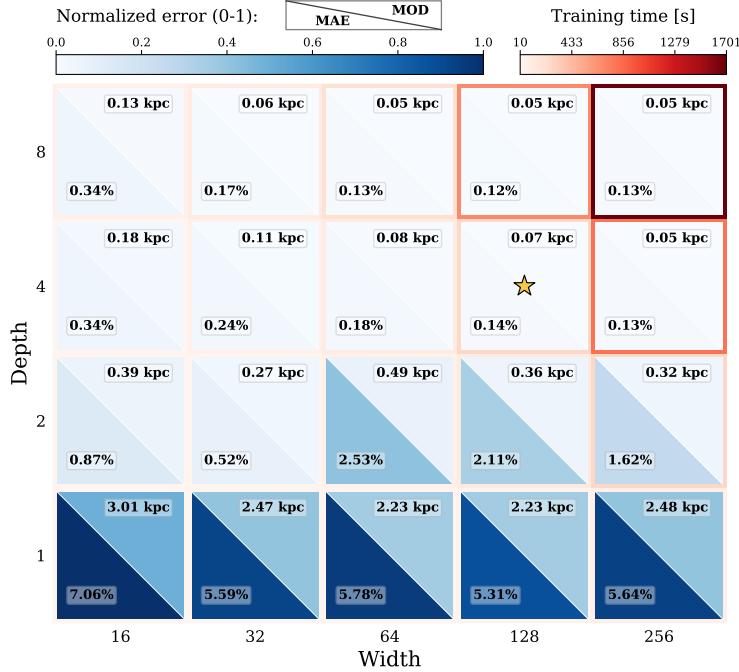


Figure 4: Effect of network size on performance. Each square corresponds to one network configuration, where the diagonal separates the two metrics: the lower-left shows the mean acceleration error (MAE, %), evaluated on 32,768 testing points; the upper-right shows the mean orbit deviation (MOD, kpc), evaluated on a random sample of 100 orbits. Training time is encoded by the color of each cell border. The star marks the configuration selected for follow-up tests.

underlying mass distribution. To assess dynamical consistency, we integrate test-particle orbits in the reconstructed potential and compare them with ground-truth trajectories. For an orbit initialized at the LMC center with the local circular velocity, the posterior mean trajectory deviates from the true trajectory by a maximum of 1.1 kpc over 1 Gyr, a substantial improvement over the 108 kpc error of the analytic MW-only baseline (Figure 2a). The learned orbits maintain stable energy evolution with minimal fluctuations relative to both the true orbit and the initial energy (Figure 2b–c).

Time-dependent field reconstruction: We extend the static model to a time-dependent setting by modeling the LMC as a perturbation to a time-evolving MW potential. Training uses six time snapshots with 1024 samples each, and the evolution is learned in reverse from the present-day ($t = 0$). As shown in Figure 3, reconstruction errors remain under 1% across a wide span of radii and epochs for both interpolation and extrapolation, with interpolation consistently more accurate.

Network size: Figure 4 summarizes how depth and width affect training time and two performance metrics: the mean acceleration error (MAE), averaged over evaluation points, and the mean orbit deviation (MOD), averaged over a sample of test orbits. Physics-informed constraints keep architectures compact: even small networks achieve sub-percent MAE and sub-kiloparsec MOD, with training converging in approximately three minutes on a standard M2 CPU.

4 Discussion and future work

We show in this paper that embedding physics-informed constraints into the training process allows neural networks to reconstruct the acceleration and potential fields of galactic systems with high accuracy. A critical next step is validation on realistic, noisy simulations to assess robustness and generalization beyond idealized test cases. Moreover, because density fields are more directly observable than acceleration fields [11], future extensions may benefit from adopting density-based loss functions, enabling more direct comparisons with observational data.

References

- [1] James Binney and Scott Tremaine. *Galactic Dynamics: Second Edition*. Princeton University Press, 2008.
- [2] Illés, Eduárd, Jánosi, Dániel, and Kovács, Tamás. Orbital dynamics in galactic potentials under mass transfer. *Astronomy & Astrophysics*, 692:A240, 2024. doi: 10.1051/0004-6361/202348274. URL <https://doi.org/10.1051/0004-6361/202348274>.
- [3] Louis E. Strigari. Galactic searches for dark matter. *Physics Reports*, 531(1):1–88, 2013. ISSN 0370-1573. doi: <https://doi.org/10.1016/j.physrep.2013.05.004>. URL <https://www.sciencedirect.com/science/article/pii/S0370157313001622>. Galactic searches for dark matter.
- [4] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. A universal density profile from hierarchical clustering. *The Astrophysical Journal*, 490(2):493–508, December 1997. ISSN 1538-4357. doi: 10.1086/304888. URL <http://dx.doi.org/10.1086/304888>.
- [5] Prasenjit Saha. Designer basis functions for potentials in galactic dynamics. *Monthly Notices of the Royal Astronomical Society*, 262(4):1062–1064, 06 1993. ISSN 0035-8711. doi: 10.1093/mnras/262.4.1062. URL <https://doi.org/10.1093/mnras/262.4.1062>.
- [6] Juan Diego Toscano, Vivek Oommen, Alan John Varghese, Zongren Zou, Nazanin Ahmadi Daryakenari, Chenxi Wu, and George Em Karniadakis. From pinns to pikans: Recent advances in physics-informed machine learning, 2024. URL <https://arxiv.org/abs/2410.13228>.
- [7] M. Raissi, P. Perdikaris, and G.E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019. ISSN 0021-9991. doi: <https://doi.org/10.1016/j.jcp.2018.10.045>. URL <https://www.sciencedirect.com/science/article/pii/S0021999118307125>.
- [8] John Martin and Hanspeter Schaub. The physics-informed neural network gravity model: Generation iii, 2025. URL <https://arxiv.org/abs/2312.10257>.
- [9] Julyan Arbel, Konstantinos Pitas, Mariia Vladimirova, and Vincent Fortuin. A primer on bayesian neural networks: Review and debates, 2023. URL <https://arxiv.org/abs/2309.16314>.
- [10] Lilia Correa Magnus and Eugene Vasiliev. Measuring the milky way mass distribution in the presence of the lmc. *Monthly Notices of the Royal Astronomical Society*, 511(2):2610–2630, 12 2021. ISSN 0035-8711. doi: 10.1093/mnras/stab3726. URL <https://doi.org/10.1093/mnras/stab3726>.
- [11] Hamish Silverwood and Richard Easter. Stellar accelerations and the galactic gravitational field. *Publications of the Astronomical Society of Australia*, 36:e038, 2019. doi: 10.1017/pasa.2019.25.