Hybrid Physical-Neural Simulator for Fast Cosmological Hydrodynamics

Arne Thomsen
ETH Zürich
Zürich, Switzerland
athomsen@phys.ethz.ch

Tilman Tröster
ETH Zürich
Zürich, Switzerland

François Lanusse
Université Paris-Saclay
CEA, CNRS, AIM
Gif-sur-Yvette, France

Abstract

Cosmological field-level inference requires differentiable forward models that solve the challenging dynamics of gas and dark matter under hydrodynamics and gravity. We propose a hybrid approach where gravitational forces are computed using a differentiable particle-mesh solver, while the hydrodynamics are parametrized by a neural network that maps local quantities to an effective pressure field. We demonstrate that our method improves upon alternative approaches, such as an Enthalpy Gradient Descent baseline, both at the field and summary-statistic level. The approach is furthermore highly data efficient, with a single reference simulation of cosmological structure formation being sufficient to constrain the neural pressure model. This opens the door for future applications where the model is fit directly to observational data, rather than a training set of simulations.

1 Introduction

Analyses of the large-scale structure of the Universe aim to constrain fundamental cosmological parameters. While structure formation at large scales is dominated by dark matter, state-of-the-art observations now probe smaller scales where the contribution of ordinary (baryonic) matter in the form of gas becomes non-negligible, necessitating hydrodynamical modeling.

Explicit field-level inference leverages the full spatial information in cosmic density and velocity fields to jointly constrain cosmological parameters and initial conditions of the matter distribution [Jasche and Wandelt, 2013]. Because the parameter space of initial conditions is extremely high-dimensional, their inference requires gradient-based sampling methods like Hamiltonian Monte Carlo [Betancourt, 2018] and hence differentiable forward models.

We present a step towards such a differentiable simulation jointly evolving gas and dark matter. Recognizing that the majority of cosmic matter is dark matter, whose purely gravitational dynamics can be solved efficiently with differentiable particle mesh (PM) methods [Hockney and Eastwood, 1988], we propose a hybrid approach that maintains physics-driven gravitational dynamics while approximating gas physics through a physically-constrained neural network parametrization. We demonstrate this solver-in-the-loop scheme [Um et al., 2021] by fitting the trainable network parameters to a single fully hydrodynamical reference simulation.

2 Related work

HYPER The hydro PM (HPM) code `HYPER` introduced by He et al. [2022] shares conceptual similarities with our approach, modeling the evolution of both dark matter and gas particles within a PM framework where additional gas forces are derived from a scalar pressure field. However, unlike in our data-driven parametrization, the functional form of the pressure field in HPM is derived

Machine Learning and the Physical Sciences Workshop, NeurIPS 2025.

analytically using the halo model [Asgari et al., 2023], with distinct treatments for the lower-density intergalactic and higher-density intracluster medium.

diffHydro The `diffHydro` code [Horowitz and Lukic, 2025] implements a fully differentiable approach to cosmological hydrodynamical simulations. In contrast to the hybrid method presented in this work, `diffHydro` is purely physics-driven, solving the hydrodynamical Euler equations in comoving coordinates using a finite volume scheme, at significant computational cost.

Enthalpy Gradient Descent (EGD) The EGD method developed in Dai et al. [2018] is a fast, differentiable technique for post-processing dark matter-only simulations to include gas effects. Under the assumption of an effective power-law equation of state $T(\delta) = T_0(1+\delta)^{(\gamma-1)}$, the method constructs an enthalpy field from the matter density contrast δ and displaces a random subset of dark matter particles along the enthalpy gradient to approximate gas dynamics. The free parameters T_0 and γ can be fit to reference hydrodynamical simulations by optimizing the matter power spectrum.

For this work, we use EGD as our primary baseline comparison since neither `HYPER` nor `diffHydro` have publicly available implementations.

3 Method

We extend the open-source dark matter-only PM code `JaxPM`[1] from Lanzieri et al. [2022] by introducing a gas particle species that, in addition to gravity, experiences a learned pressure force. As a PM code, forces derived from gradients of scalar fields can be efficiently computed through point-wise multiplications in Fourier space, with cloud-in-cell interpolation to transition between particle and mesh representations. The `jax` [Bradbury et al., 2018] implementation provides the automatic differentiation, just-in-time compilation, and GPU acceleration necessary to train the neural pressure model.

3.1 Hybrid physical-neural equations of motion

The simulator evolves equal numbers of cold dark matter and baryonic gas particles with masses proportional to the cosmological density fractions Ω_{cdm} and Ω_{b}, respectively. Similar to He et al. [2022], we integrate the system of ordinary differential equations (ODEs) in comoving coordinates:

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{x}_{\mathrm{dm}}}{\mathrm{d}a} = \dfrac{1}{a^3 E(a)}\mathbf{v}_{\mathrm{dm}} \\ \dfrac{\mathrm{d}\mathbf{v}_{\mathrm{dm}}}{\mathrm{d}a} = -\dfrac{1}{a^2 E(a)}\nabla\Phi_{\mathrm{tot}}, \end{cases} \qquad \begin{cases} \dfrac{\mathrm{d}\mathbf{x}_{\mathrm{gas}}}{\mathrm{d}a} = \dfrac{1}{a^3 E(a)}\mathbf{v}_{\mathrm{gas}} \\ \dfrac{\mathrm{d}\mathbf{v}_{\mathrm{gas}}}{\mathrm{d}a} = -\dfrac{1}{a^2 E(a)}\left(\nabla\Phi_{\mathrm{tot}} + \dfrac{\nabla P}{\rho_{\mathrm{gas}}}\right), \end{cases} \tag{1}$$

where $\mathbf{x}$ and $\mathbf{v}$ denote particle positions and velocities, a is the cosmological scale factor serving as the time variable, and $E(a) = H(a)/H_0$ is the dimensionless Hubble parameter. The scale factor-dependent prefactors account for the background expansion of the universe [Quinn et al., 1997, Schaller et al., 2024]. The force terms proportional to $-\nabla\Phi_{\mathrm{tot}}$ and $-\nabla P/\rho_{\mathrm{gas}}$ are detailed below.

The dynamics of both particle species are coupled through the scalar gravitational potential Φ_{tot}, which is related to the total matter density contrast δ_{tot} via the Poisson equation [Angulo and Hahn, 2022]

$$\nabla^2\Phi_{\mathrm{tot}}(\mathbf{x}) = \frac{3}{2}H_0^2\Omega_{\mathrm{m}}\delta_{\mathrm{tot}}(\mathbf{x}), \tag{2}$$

where $\delta_{\mathrm{tot}} = \delta_{\mathrm{dm}} + \delta_{\mathrm{gas}}$ and $\delta = \rho/\bar{\rho} - 1$ with ρ the local density and $\bar{\rho}$ its spatial mean. Here, $\Omega_{\mathrm{m}} = \Omega_{\mathrm{cdm}} + \Omega_{\mathrm{b}}$ is the total matter density fraction. The density fields are computed on a mesh, enabling efficient solution of eq. (2) in Fourier space. Adopting the approach of Lanzieri et al. [2022], we optionally include a residual neural correction f_θ to the Fourier-transformed potential: $\tilde{\Phi}^{\star}_{\mathrm{tot}} = (1 + f_\theta(a, |\mathbf{k}|))\,\tilde{\Phi}_{\mathrm{tot}}$. This hybrid approach combines the physics-driven potential with a data-driven correction to improve small-scale accuracy.

[1] `https://github.com/DifferentiableUniverseInitiative/JaxPM`

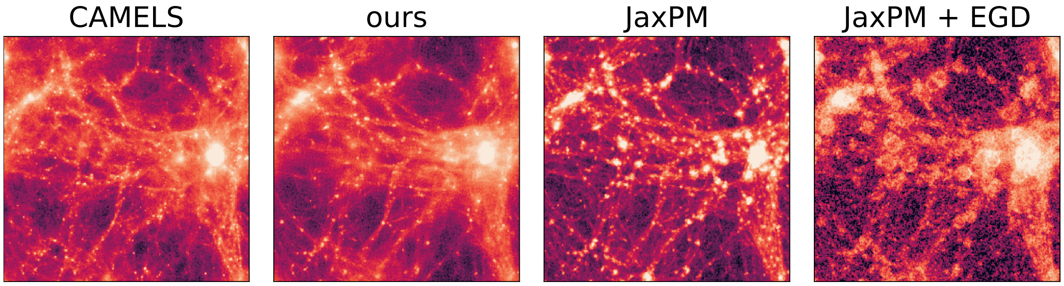

Figure 1: Projected gas density ρ_{gas} (logarithmic scale) comparing four methods: reference hydrodynamical CAMELS simulation, our hybrid simulator with learned gas pressure force, gravity-only `JaxPM`, and its EGD post-processing. Our pressure model suppresses small-scale structure formation, better matching the hydrodynamical reference than purely gravitational evolution while retaining more detail than EGD.

3.2 Effective neural pressure

The effective pressure force acting on gas particles takes the physically motivated Euler form $-\nabla P/\rho_{\text{gas}}$ as defined in eq. (1). Under an ideal gas law assumption, we express the learned pressure field at particle position $\mathbf{x}$ as

$$P_\varphi(a, \mathbf{x}) \propto \rho_{\text{gas}}(\mathbf{x})\, U_\varphi(a, \mathbf{h}(\mathbf{x})), \tag{3}$$

where U_φ is a neural network with trainable parameters φ that maps the scale factor a and feature vector of local quantities $\mathbf{h}(\mathbf{x})$ to the non-negative internal energy. This decomposition reduces the dynamic range compared to predicting P directly, significantly improving training stability. We implement U_φ as a fully convolutional neural network (CNN) operating on the mesh interpolation, though the framework supports alternative architectures such as multilayer perceptrons (MLPs) acting directly on particles. Both architectures operate locally, decoupling the box sizes of reference and hybrid simulations and hence enabling future application to larger cosmological volumes.

For computational efficiency within the PM framework, we construct input features $\mathbf{h}(\mathbf{x})$ that can be rapidly evaluated on the mesh. The resulting feature vector comprises

$$\mathbf{h}(\mathbf{x}) = \left(\rho_{\text{gas}}(\mathbf{x}),\ f_{\text{scalar}}(\mathbf{x}),\ \nabla_{\mathbf{v}}(\mathbf{x}),\ \sigma^2_{\mathbf{v}}(\mathbf{x})\right), \tag{4}$$

where ρ_{gas} is the gas density, $\nabla_{\mathbf{v}} = \nabla \cdot \mathbf{v}$ is the velocity divergence, and $\sigma^2_{\mathbf{v}} = |\mathbf{v} - \langle \mathbf{v} \rangle|^2$ is the velocity dispersion with $\langle \mathbf{v} \rangle$ denoting the local mean velocity. As in He et al. [2022], we define

$$f_{\text{scalar}}(\mathbf{x}) \propto \rho_{\text{gas}}(\mathbf{x}) * \frac{1}{|\mathbf{x}|^2} \Rightarrow \tilde{f}_{\text{scalar}}(\mathbf{k}) \propto \frac{2\pi^2 \tilde{\rho}_{\text{gas}}(\mathbf{k})}{|\mathbf{k}|}, \tag{5}$$

which we efficiently compute in Fourier space. The input fields are visualized in Fig. D.1. We empirically find that extending $\mathbf{h}(\mathbf{x})$ by the rotationally invariant eigenvalues of the tidal field tensor (or Hessian of the gravitational potential) does not significantly improve performance and therefore do not include them.

We integrate eq. (1) using `diffrax`[2] [Kidger, 2022], which allows automatic differentiation with respect to model parameters φ using the recursive checkpoint adjoint [Wang et al., 2009, Stumm and Walther, 2010]. Following Lanzieri et al. [2022], we exploit this differentiability to optimize our hybrid solver-in-the-loop [Um et al., 2021] model directly against a reference simulation (see Appendix A) with particle positions $\mathbf{x}^{\text{ref}}$ and velocities $\mathbf{v}^{\text{ref}}$. We minimize the loss function

$$\mathcal{L} = \sum_s \left[H_\delta\left(\mathbf{r}_s\right) + \lambda H_{\delta'}\left(\mathbf{v}_s - \mathbf{v}_s^{\text{ref}}\right) + \mu \left\| \frac{P_s(|\mathbf{k}|)}{P_s^{\text{ref}}(|\mathbf{k}|)} - 1 \right\|_2^2 \right], \tag{6}$$

where $H_\delta(r) = \min(r^2/2, \delta|r| - \delta^2/2)$ is the robust Huber loss that reduces sensitivity to outlier particles, $\mathbf{r}_s = \left(\left(\mathbf{x}_s - \mathbf{x}_s^{\text{ref}} + L/2\right) \bmod L\right) - L/2$ denotes the particle displacement under periodic boundary conditions in a simulation box of side length L, $P(|\mathbf{k}|)$ is the power spectrum of ρ_{gas}, and $\{\lambda, \mu, \delta, \delta'\}$ are hyperparameters. The index s runs over simulation snapshots (or a sparse subset thereof) taken at different scale factors a_s during cosmic evolution. Example reference snapshots are shown in the top row of Fig. D.3.

[2] `https://docs.kidger.site/diffrax/`

4 Results

We demonstrate the data efficiency of our hybrid approach by training the neural pressure model on a single fully hydrodynamical reference simulation from the SIMBA[3] subset [Davé et al., 2019] of the CAMELS[4] suite [Villaescusa-Navarro et al., 2021], detailed in Appendix A. We select SIMBA because its strong gas feedback provides a more challenging test than ASTRID [Bird et al., 2022, Ni et al., 2022], whose weaker feedback produces negligible deviations from gravity-only dynamics at our resolution.

We downsample this reference to 128^3 particles each for dark matter and gas, with an equally sized mesh for PM force computation, enabling training on an individual A100 GPU. The system in eq. (1) is initialized using positions and velocities from the first CAMELS snapshot ($s = 0$) at scale factor $a_0 = 0.14$, corresponding to redshift $z_0 = 6$. Additional implementation details are provided in Appendix B.

To further emphasize data efficiency, we evaluate the loss in eq. (6) at only 4 of the 34 available snapshots with scale factors $a_s \in \{0.30, 0.44, 0.65, 1.0\}$. Training remains feasible despite this sparse supervision due to the physically constrained parametrization of the gas pressure force in eqs. (1) and (3), the large number of simulated particles per snapshot, and the mesh size (128 cells per dimension) substantially exceeding the network's receptive field (17 cells). Furthermore, the physics-driven gravitational dynamics in our hybrid approach provide a data-independent backbone, with gas physics acting as a learned correction.

All results are evaluated on held-out test simulations from CAMELS with different random initial conditions, ensuring model performance reflects generalization to cosmic variance rather than memorization of a specific realization.

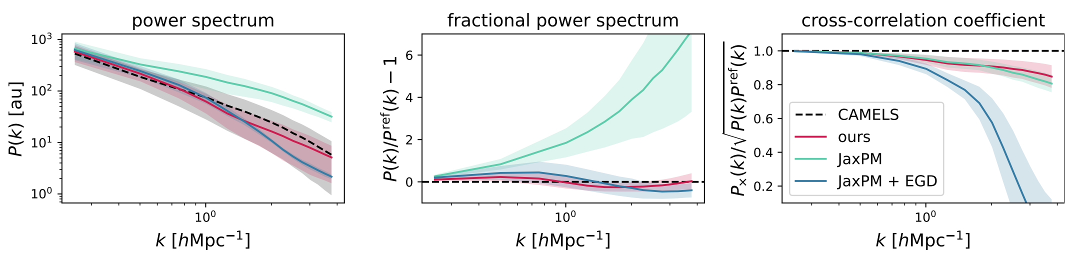

Figure 2: Two-point statistics of gas densities ρ_{gas} from Fig. 1. Solid lines and shaded bands indicate means and standard deviations over random initial conditions (cosmic variance), respectively. *Left and middle:* Power spectra demonstrate that both our method and EGD suppress small-scale power relative to gravity-only evolution (`JaxPM`), achieving better agreement with the CAMELS reference. *Right:* Our hybrid simulator shows consistently higher cross-correlation with the reference than EGD. These quantitative results confirm the visual intuition from Fig. 1.

We present results at $a = 1$ (present day), showing the density fields in Fig. 1 and their corresponding two-point statistics measured across varying random initial conditions in Fig. 2. We compare the reference CAMELS simulation against our hybrid method, exclusively gravitational `JaxPM` evolution, and the EGD baseline (see Appendix C for implementation details). A key distinction exists between our method and this baseline: EGD post-processes gravity-only simulations at a single target scale factor, whereas our approach self-consistently evolves the system throughout cosmic history, even when trained on only a small subset of simulation snapshots. Results at different points in the evolution are shown in Fig. D.3.

We find that our hybrid extension significantly enhances gravity-only `JaxPM` evolution at both the field and two-point levels. Compared to EGD, our approach achieves similar improvement in the power spectrum (measuring Fourier mode amplitudes) while demonstrating superior field-level performance, as shown in Fig. 1 and in the cross-correlation coefficient (capturing Fourier phase coherence).

Additionally, we show the dark matter distributions in Fig. D.2, demonstrating negligible differences in baryonic back-reaction on dark matter across methods.

[3] `http://simba.roe.ac.uk/`

[4] `https://camels.readthedocs.io/en/latest/index.html`

5 Conclusion

We present a computationally efficient hybrid simulator that jointly evolves gas and dark matter particles by computing gravitational forces through fast PM methods while modeling gas pressure forces with an embedded physically constrained neural network.

This fully differentiable approach enables field-level inference, jointly constraining high-dimensional initial conditions with cosmological parameters. The method's data efficiency, requiring only a few snapshots of a single reference simulation for training, in principle opens the possibility of fitting directly to observational data rather than relying on extensive simulation training sets. For example, observations of gas, such as Sunyaev-Zeldovich effects, together with tomographic weak gravitational lensing could provide similar information on the gas and dark matter density as the sparse sampling of the simulations employed here. This could go towards addressing the recognized challenge of model misspecification between different hydrodynamical codes [e.g. Ni et al., 2023, Akhmetzhanova et al., 2025].

Our neural parametrization of the pressure field uses only instantaneous local quantities derived from the particle positions and velocities, ignoring their history. Complete thermodynamic treatment requires tracking an additional state variable like internal energy, entropy, or temperature. This limitation could be overcome by augmenting eq. (1) with latent variables whose time derivatives are predicted by the embedded network, yielding a neural ODE [Chen et al., 2019] that enables history-aware pressure predictions.

References

Aizhan Akhmetzhanova, Carolina Cuesta-Lazaro, and Siddharth Mishra-Sharma. Detecting Model Misspecification in Cosmology with Scale-Dependent Normalizing Flows, August 2025.

Raul E. Angulo and Oliver Hahn. Large-scale dark matter simulations. *Living Reviews in Computational Astrophysics*, 8(1):1, February 2022. ISSN 2365-0524. doi: 10.1007/s41115-021-00013-z.

Marika Asgari, Alexander J. Mead, and Catherine Heymans. The halo model for cosmology: A pedagogical review. *The Open Journal of Astrophysics*, 6, November 2023. doi: 10.21105/astro.2303.08752.

Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E. Hinton. Layer Normalization, July 2016.

Michael Betancourt. A Conceptual Introduction to Hamiltonian Monte Carlo, July 2018.

Simeon Bird, Yueying Ni, Tiziana Di Matteo, Rupert Croft, Yu Feng, and Nianyi Chen. The ASTRID simulation: Galaxy formation and reionization. *Monthly Notices of the Royal Astronomical Society*, 512(3):3703–3716, May 2022. ISSN 0035-8711. doi: 10.1093/mnras/stac648.

James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: Composable transformations of Python+NumPy programs, 2018.

Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud. Neural Ordinary Differential Equations, December 2019.

Biwei Dai, Yu Feng, and Uroš Seljak. A gradient based method for modeling baryons and matter in halos of fast simulations. *Journal of Cosmology and Astroparticle Physics*, 2018(11):009, November 2018. ISSN 1475-7516. doi: 10.1088/1475-7516/2018/11/009.

Romeel Davé, Daniel Anglés-Alcázar, Desika Narayanan, Qi Li, Mika H Rafieferantsoa, and Sarah Appleby. Simba: Cosmological simulations with black hole growth and feedback. *Monthly Notices of the Royal Astronomical Society*, 486(2):2827–2849, June 2019. ISSN 0035-8711. doi: 10.1093/mnras/stz937.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, Las Vegas, NV, USA, June 2016. IEEE. ISBN 978-1-4673-8851-1. doi: 10.1109/CVPR.2016.90.

Yizhou He, Hy Trac, and Nickolay Y. Gnedin. A Hydro-Particle-Mesh Code for Efficient and Rapid Simulations of the Intracluster Medium. *The Astrophysical Journal*, 925(2):134, February 2022. ISSN 0004-637X, 1538-4357. doi: 10.3847/1538-4357/ac3bcb.

R. W. Hockney and J. W. Eastwood. *Computer Simulation Using Particles*. CRC Press, Boca Raton, 1988. ISBN 978-0-367-80693-4. doi: 10.1201/9780367806934.

Benjamin Horowitz and Zarija Lukic. Differentiable Cosmological Hydrodynamics for Field-Level Inference and High Dimensional Parameter Constraints, February 2025.

Jens Jasche and Benjamin D. Wandelt. Bayesian physical reconstruction of initial conditions from large-scale structure surveys. *Monthly Notices of the Royal Astronomical Society*, 432(2):894–913, June 2013. ISSN 0035-8711. doi: 10.1093/mnras/stt449.

Patrick Kidger. On Neural Differential Equations, February 2022.

Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization, January 2017.

Denise Lanzieri, François Lanusse, and Jean-Luc Starck. Hybrid Physical-Neural ODEs for Fast N-body Simulations, July 2022.

Dylan Nelson, Volker Springel, Annalisa Pillepich, Vicente Rodriguez-Gomez, Paul Torrey, Shy Genel, Mark Vogelsberger, Ruediger Pakmor, Federico Marinacci, Rainer Weinberger, Luke Kelley, Mark Lovell, Benedikt Diemer, and Lars Hernquist. The IllustrisTNG simulations: Public data release. *Computational Astrophysics and Cosmology*, 6(1):2, May 2019. ISSN 2197-7909. doi: 10.1186/s40668-019-0028-x.

Yueying Ni, Tiziana Di Matteo, Simeon Bird, Rupert Croft, Yu Feng, Nianyi Chen, Michael Tremmel, Colin DeGraf, and Yin Li. The ASTRID simulation: The evolution of supermassive black holes. *Monthly Notices of the Royal Astronomical Society*, 513(1):670–692, June 2022. ISSN 0035-8711. doi: 10.1093/mnras/stac351.

Yueying Ni, Shy Genel, Daniel Anglés-Alcázar, Francisco Villaescusa-Navarro, Yongseok Jo, Simeon Bird, Tiziana Di Matteo, Rupert Croft, Nianyi Chen, Natalí S. M. de Santi, Matthew Gebhardt, Helen Shao, Shivam Pandey, Lars Hernquist, and Romeel Dave. The CAMELS Project: Expanding the Galaxy Formation Model Space with New ASTRID and 28-parameter TNG and SIMBA Suites. *The Astrophysical Journal*, 959(2):136, December 2023. ISSN 0004-637X. doi: 10.3847/1538-4357/ad022a.

Ethan Perez, Florian Strub, Harm De Vries, Vincent Dumoulin, and Aaron Courville. FiLM: Visual Reasoning with a General Conditioning Layer. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), April 2018. ISSN 2374-3468, 2159-5399. doi: 10.1609/aaai.v32i1.11671.

Thomas Quinn, Neal Katz, Joachim Stadel, and George Lake. Time stepping N-body simulations, October 1997.

Matthieu Schaller, Josh Borrow, Peter W Draper, Mladen Ivkovic, Stuart McAlpine, Bert Vandenbroucke, Yannick Bahé, Evgenii Chaikin, Aidan B G Chalk, Tsang Keung Chan, Camila Correa, Marcel van Daalen, Willem Elbers, Pedro Gonnet, Loïc Hausammann, John Helly, Filip Huško, Jacob A Kegerreis, Folkert S J Nobels, Sylvia Ploeckinger, Yves Revaz, William J Roper, Sergio Ruiz-Bonilla, Thomas D Sandnes, Yolan Uyttenhove, James S Willis, and Zhen Xiang. Swift: A modern highly parallel gravity and smoothed particle hydrodynamics solver for astrophysical and cosmological applications. *Monthly Notices of the Royal Astronomical Society*, 530(2):2378–2419, May 2024. ISSN 0035-8711. doi: 10.1093/mnras/stae922.

Philipp Stumm and Andrea Walther. New algorithms for optimal online checkpointing. *SIAM Journal on Scientific Computing*, 32(2):836–854, 2010. doi: 10.1137/080742439.

Kiwon Um, Robert Brand, Yun, Fei, Philipp Holl, and Nils Thuerey. Solver-in-the-Loop: Learning from Differentiable Physics to Interact with Iterative PDE-Solvers, January 2021.

Francisco Villaescusa-Navarro, Daniel Anglés-Alcázar, Shy Genel, David N. Spergel, Rachel S. Somerville, Romeel Dave, Annalisa Pillepich, Lars Hernquist, Dylan Nelson, Paul Torrey, Desika Narayanan, Yin Li, Oliver Philcox, Valentina La Torre, Ana Maria Delgado, Shirley Ho, Sultan Hassan, Blakesley Burkhart, Digvijay Wadekar, Nicholas Battaglia, Gabriella Contardo, and Greg L. Bryan. The CAMELS project: Cosmology and Astrophysics with MachinE Learning Simulations. *The Astrophysical Journal*, 915(1):71, July 2021. ISSN 0004-637X, 1538-4357. doi: 10.3847/1538-4357/abf7ba.

Qiqi Wang, Parviz Moin, and Gianluca Iaccarino. Minimal repetition dynamic checkpointing algorithm for unsteady adjoint calculation. *SIAM Journal on Scientific Computing*, 31(4):2549–2567, 2009. doi: 10.1137/080727890.

Acknowledgments

AT thanks Adrian Bayer and Matthew Ho for helpful discussions, and Alexandre Refregier for his backing of the project. FL thanks Chihway Chang and Yuuki Omori for early conversations, and the Flatiron Institute for its support. This work was supported by an ETH Zurich Doc.Mobility Fellowship, which funded AT's research visit to the Flatiron Institute's Center for Computational Astrophysics in New York. TT acknowledges funding from the Swiss National Science Foundation under the Ambizione project PZ00P2_193352.

A Reference simulation data

Because our approach treats the pressure field P in eq. (1) as a data-driven component without explicit physical assumptions, we require a ground truth reference to train it. We employ the fully hydrodynamical Cosmology and Astrophysics with MachinE Learning Simulations (CAMELS) suite [Villaescusa-Navarro et al., 2021] for this purpose.

We utilize the cosmic variance (CV) subsets of CAMELS, which comprise 27 simulations with varying random initial conditions at $z = 127$ while maintaining fixed fiducial hydrodynamical and ΛCDM cosmological parameters: $\Omega_{\rm cdm} = 0.251$, $\Omega_{\rm b} = 0.049$, $\sigma_8 = 0.8$, $h = 0.6711$, and $n_s = 0.9624$. Each simulation evolves 256^3 dark matter particles in a periodic box of comoving volume $(25\,{\rm Mpc}/h)^3$.

Since our neural pressure model is particle-based following the dynamics of eq. (1), we focus on hydrodynamical codes that maintain approximately constant gas particle counts and masses throughout the simulation. This requirement excludes mesh-based codes like ILLUSTRISTNG [Nelson et al., 2019] and restricts our analysis to the mesh-free SIMBA [Davé et al., 2019] and smoothed-particle hydrodynamics ASTRID [Bird et al., 2022, Ni et al., 2022] subsets of CAMELS. For this work, we focus on the SIMBA simulations, as the strong feedback processes at the fiducial parameters in its CV set provide a more challenging test for our method. We also tested the neural pressure model on ASTRID, but the weak feedback in its CV set produces minimal differences between the gravity-only solution and one including gas physics at our PM resolution.

The SIMBA simulations are stored as 34 snapshots ($s \in [0, 33]$ in eq. (6)) spanning cosmological redshifts $z = 6$ to the present day ($z = 0$), with scale factor $a(t) = 1/(1 + z)$. For the results presented in this work, we train our models using only a subset of four ($a_s \in \{0.30, 0.44, 0.65, 1.0\}$) of these snapshots from a single simulation.

B Neural pressure implementation details

We implement neural networks in `flax`[5] and train them using `optax`[6]. Our default fully convolutional architecture incorporates feature-wise linear modulation (FiLM) conditioning [Perez et al., 2018] on the scale factor a, residual connections [He et al., 2016], layer normalization [Ba et al., 2016], and circular padding to respect the periodic boundary conditions of the simulation box. Throughout this work, we use a network comprising 6 hidden layers with 16 channels each and (3,3,3) kernels with

[5] `https://flax.readthedocs.io/en/latest/`

[6] `https://optax.readthedocs.io/en/latest/`

unit stride, totaling 47 265 trainable parameters. We optimize using Adam [Kingma and Ba, 2017] with a constant learning rate of 10^{-4} for 1 000 steps and gradient clipping at global norm 1. Training stability is generally sensitive to these choices, which we identified via random hyperparameter search. In the loss $\mathcal{L}$ from eq. (6), we set $\lambda = 0.01$, $\mu = 0.1$, $\delta = M/8$, and $\delta' = 4\delta$, where M is the mesh resolution (number of cells per dimension).

We integrate eq. (1) over 67 fixed steps in a that exactly include the reference scale factors a_s. To enforce non-negativity and reduce dynamic range, the network predicts $\log U_\varphi$ in eq. (3). During training, we apply random 90° rotations and axis-aligned flips as data augmentation to encourage approximate equivariance to these symmetries.

C EGD implementation details

In the EGD method, the effective temperature field $T(\delta) = T_0(1+\delta)^{(\gamma-1)}$ is smoothed by a Gaussian kernel $\hat{\mathbf{O}}(k) = \exp(-(kr_J)^2/2)$ in Fourier space before enthalpy gradients are computed. We determine the free parameters $T_0 = 0.012$, $\gamma = 1.034$, and $r_J = 8.127$ by optimizing only the power spectrum term from eq. (6). This restriction is necessary because the randomly sampled dark matter particles serving as gas particle proxies lack direct counterparts with particles in the reference hydrodynamical simulation, preventing evaluation of the position and velocity loss terms.

D Additional figures

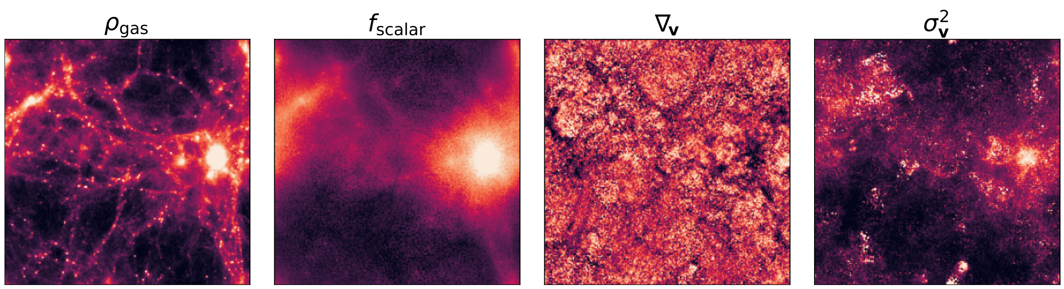

Figure D.1: Projections of the four local features $\mathbf{h}(\mathbf{x})$ used to predict the effective pressure P_φ: gas density, scalar force (see eq. (5)), velocity divergence, and velocity dispersion (left to right). For visualization, each field h is separately scaled by $\operatorname{arcsinh}(h/\sigma_h)$.

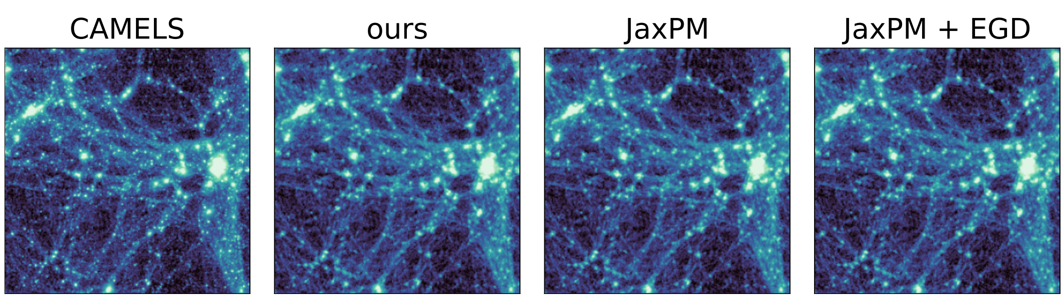

Figure D.2: Like Fig. 1, but for the dark matter density ρ_{dm}. At this resolution, differences between methods are minimal, demonstrating that our neural pressure model does not adversely impact dark matter evolution. The large-scale structure resembles that in Fig. 1 because gas traces the underlying dark matter distribution up to scales where gas feedback effects become significant.

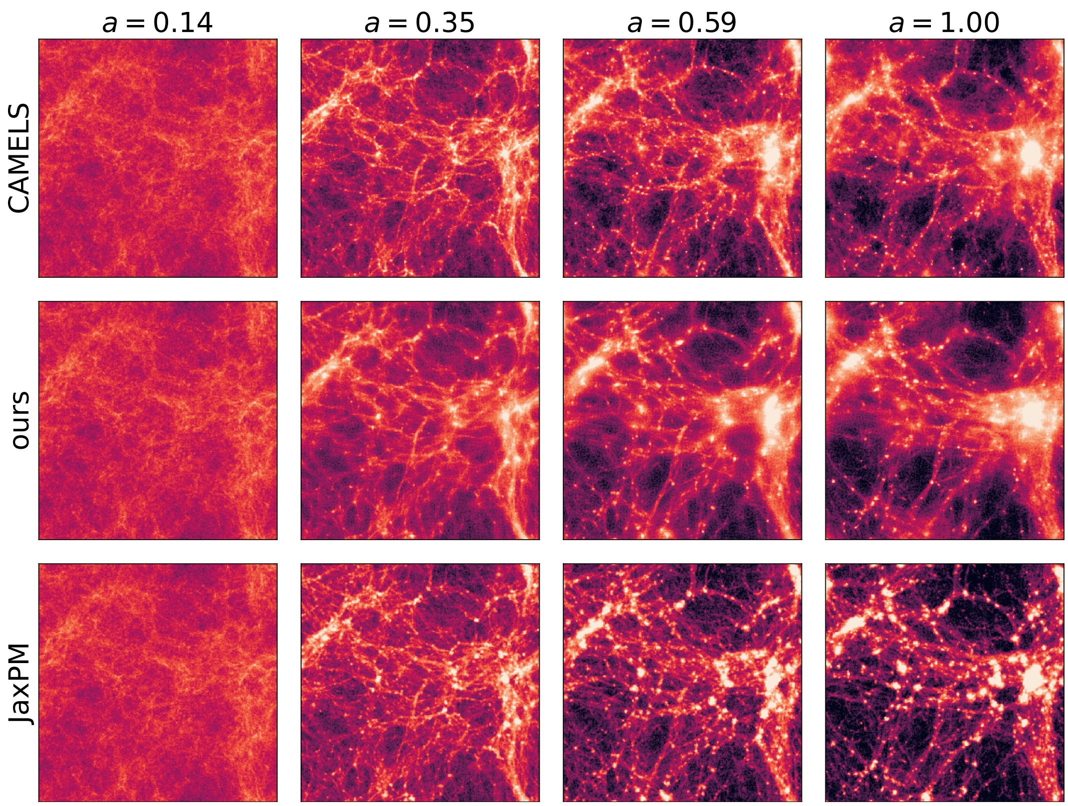

Figure D.3: Evolution of gas density ρ_{gas} from our initial conditions ($a = 0.14$) to the present day ($a = 1$). For this example, the penalized scale factors in the loss function eq. (6) are $a_s \in \{0.30, 0.44, 0.65, 1.0\}$. As a post-processing method, EGD does not model a self-consistent evolution and is therefore not shown here.