Learning an Effective Evolution Equation for Particle-Mesh Simulations Across Cosmologies

Nicolas Payot
Département de Physique, Université de Montréal
Mila - Quebec Artificial Intelligence Institute
Ciela - Montreal Institute for Astrophysical Data Analysis and Machine Learning

Pablo Lemos
Département de Physique, Université de Montréal
Mila - Quebec Artificial Intelligence Institute
Ciela - Montreal Institute for Astrophysical Data Analysis and Machine Learning

Laurence Perreault-Levasseur
Département de Physique, Université de Montréal
Mila - Quebec Artificial Intelligence Institute
Ciela - Montreal Institute for Astrophysical Data Analysis and Machine Learning

Carolina Cuesta-Lazaro
The NSF AI Institute for Artificial Intelligence and Fundamental Interactions
Massachusetts Institute of Technology, Cambridge, MA
Center for Astrophysics | Harvard & Smithsonian, Cambridge, MA

Chirag Modi
Center for Computational Astrophysics, Flatiron Institute, NY
Center for Computational Mathematics, Flatiron Institute, NY

Yashar Hezaveh
Département de Physique, Université de Montréal
Mila - Quebec Artificial Intelligence Institute
Ciela - Montreal Institute for Astrophysical Data Analysis and Machine Learning

Abstract

Particle-mesh simulations trade small-scale accuracy for speed compared to traditional, computationally expensive N-body codes in cosmological simulations. In this work, we show how a data-driven model could be used to learn an effective evolution equation for the particles, by correcting the errors of the particle-mesh potential incurred on small scales during simulations. We find that our learnt correction yields evolution equations that generalize well to new, unseen initial conditions and cosmologies. We further demonstrate that the resulting corrected maps can be used in a simulation-based inference framework to yield an unbiased inference of cosmological parameters. The model, a network implemented in Fourier space, is exclusively trained on the particle positions and velocities.

1 Introduction

N-body simulations are a ubiquitous tool in astrophysics for modeling the dynamics of particles under the influence of their collective gravitational potential. While calculating interactions among a small number of particles can be relatively straightforward (e.g., with algorithms like Verlet integration), the computational burden escalates sharply with an $\mathcal{O}(n^2)$ time complexity, where n denotes the number of particles in the simulation [Hockney and Eastwood, 2021].

Machine Learning and the Physical Sciences Workshop, NeurIPS 2023.

In cosmology, N-body simulations are employed to generate theoretical predictions for the large scale structure of the Universe by simulating the evolution of the dark matter distribution. These predictions are then used for comparison with observational data. With the advent of new generations of astronomical surveys probing increasingly large scales with unprecedented precision, there is a pressing need for N-body simulations that are both fast and precise on those scales. This has inspired the development of faster approximate methods, including solvers based on Lagrangian Perturbation Theory (LPT) [Buchert, 1992, Buchert and Ehlers, 1993] and particle-mesh (PM) simulation-based approaches [Hockney and Eastwood, 2021]. Commonly used implementations of these techniques include 2LPT [Crocce et al., 2006], FastPM [Feng et al., 2016], and COLA [Tassev et al., 2013].

Particle-mesh simulations work by binning the particles on a grid depending on their mass and position. The contribution of the particles to other grid points is calculated using a cloud-in-cell (CIC) interpolation scheme. Following this step, Poisson's equation is solved using Fast Fourier Transforms (FFTs) for the mass distribution on the grid. Once the potential is calculated, its inverse Fourier transform is obtained using another FFT operation. Forces acting on each particle can be interpolated using the obtained potential with another CIC. An update to particle positions and velocities are then computed and a time step is taken. Particle-mesh simulations are limited by either the number of particles (since assigning the particles to the grid mesh is of order $\mathcal{O}(n)$, where n is the number of particle) or the potential calculation, which is of order $\mathcal{O}(m \log m)$, where m is the number of grid nodes [Hockney and Eastwood, 2021, Bodenheimer, 2007].

Unfortunately, these methods trade the small-scale accuracy of full N-body simulations for speed. To mitigate these limitations, PM simulations can be enhanced by techniques designed to correct the small-scale interactions, as done by P3M method [Bodenheimer, 2007], or incorporating machine learning models. However, these methods typically focus on correcting small-scale without considering the specific cosmology being used (see e.g., He et al. [2019] for a neural network correction of LPT within a single cosmology), resulting in poor interpretability and questionable generalizability. As an alternative, the work of e.g. Angulo and White [2010], Ruiz et al. [2011] has focused on developing simple analytical schemes to rescale an N-body simulations at a particle or halo level to different cosmologies. However, such schemes cannot generate simulations with varying initial conditions, and assessing their accuracy is difficult beyond the first order quasi-linear regime.

Inspired by the framework of the effective field theory of large scale structure [Carrasco et al., 2014], a more principled approach is to learn a correction to the evolution equation of the particles, that is to say, to learn an effective evolution equation capturing the error introduced by coarse-graining the gravitational potential on a mesh. This idea was initially explored in Lanzieri et al. [2022], however, their loss function had to include the ratio of the predicted and reference power spectra. Therefore, the accuracy of such correction for different, potentially more informative summary statistics is not guaranteed. In the present work, we show that it is possible to obtain an effective evolution equation that is robust to variations in cosmological parameters and initial conditions in a physically principled manner, by only imposing loss terms at the level of the position and velocities of the particles, effectively enforcing the conservation of global angular and linear momenta.

2 Methods

We used `JaxPM`, a differential PM simulation package written in `JAX` [1]. Following Lanzieri et al. [2022] and Chatziloizos et al. [2022], we adopted a fully connected neural network whose outputs represent the coefficients of a B-spline function with an order of 3. The network is an isotropic filter in Fourier space with sinusoidal activation functions [Zhumekenov et al., 2019] to preserve translational and rotational symmetries. Specifically, the network learns to correct the particle-mesh potential and subsequently applies these corrections to the potential in position space.

In `JaxPM`, the particle-mesh solver consists of a set of ordinary differential equations (ODEs), enabling the back-propagation of the gradient to the initial conditions [Lanzieri et al., 2022]. The ODE is:

$$\frac{dx}{da} = \frac{v}{a^3 E(a)} \tag{1}$$

$$\frac{dv}{da} = \frac{F_\theta(x, a, \Omega_m, \sigma_8)}{a^2 E(a)}\,, \tag{2}$$

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

with the force given by the equation (which includes the neural network correction to the potential):

$$F_\theta(x, a, \Omega_m, \sigma_8) = \frac{3\Omega_m}{2} \nabla \left(\phi_{PM}(x) * \mathcal{F}^{-1}(1 - f_\theta(a, |k|, \Omega_m, \sigma_8)) \right) . \quad (3)$$

Here, x represents the positions of the particles, and a corresponds to the cosmological scale factor. We use $\mathcal{F}^{-1}$ to denote inverse Fourier transform, and $f_\theta(a, |k|, \Omega_m, \sigma_8)$ to represent the neural network. In this study, we employ a total of 32 knots, and the fully connected network consists of 5 hidden layers, each with a size of 64.

For training, we used simulations from the CAMELS suite [Villaescusa-Navarro et al., 2021, Ni et al., 2023, Villaescusa-Navarro et al., 2023]. We used IllustrisTNG dark matter only simulations from the LH set, which contains 1000 simulations, each featuring distinct initial conditions and varying cosmological parameters [Nelson et al., 2019]. Specifically, Ω_m and σ_8 are sampled within the range $[0.1, 0.5]$ and $[0.6, 1]$, respectively, using a Latin Hypercube sampling method. Each simulation consists of 256^3 particles within a periodic comoving volume of $\left(25\,\mathrm{Mpc}\,h^{-1}\right)^3$, spanning redshifts $z = 127$ to $z = 0$ and captured in 34 snapshots. During training, the particle-mesh simulation is initialized at $z = 127$ and subsequently simulated until $z = 0$. The remaining 33 snapshots are then used in the loss function, constructed with the $L2$ norm of the desired positions and velocities of the particles in the simulations:

$$\mathcal{L} = \sum_i \sum_{j=0}^{33} \left(||x_{ij}^{\text{nbody}} - x_{ij}||_2^2 + \lambda ||v_{ij}^{\text{nbody}} - v_{ij}||_2^2 \right) + \gamma \sum_i \beta_i^2 \,. \quad (4)$$

Here, x_{ij} and v_{ij} represent the position and velocity of particle i in snapshot j and β_i denotes the weights of the fully-connected neural network used for the correction, making the second term a simple $L2$ regularization. Two hyperparameters λ and γ adjust the contribution of the different losses and the impact of the regularization, respectively.

3 Results

We explored with different hyperparameters in the loss function to find models with better performance. With $\lambda = 0$, which effectively removes the contribution of velocities to the loss, the velocities tended to be over-corrected, even with strong γ regularization. With $\lambda = 1$, the $L2$ norm of the velocities tended to be bigger than that of the positions, causing the model to prioritize velocity corrections at the expense of positional accuracy. We found empirically that choosing $\lambda = 0.01$ and $\gamma = 1$ balanced the contributions of the three terms.

After fixing these hyperparameters, we trained two models: one with 64^3 particles and one with 32^3. Both models were trained using a learning rate of 0.001 on simulations ranging from LH100 to LH500, comprising a total of 400 distinct cosmologies. They were trained respectively for 400 and 500 epochs on a single NVIDIA A100. The complete set of simulations could not be utilized because of limitations in storage capacity. Out of the remaining 600 simulations, 100 were used for validation, while the remaining 500 were used for testing.

As shown in Lanzieri et al. [2022] a learnt correction to the evolution equations can generalize well across different cosmologies, even when trained on a single set of cosmological parameters. We observed a similar result with our models; however, conditioning the neural network on the cosmological parameters performed even better, hence we only present the results of the latter strategy. Figure 1 illustrates how the correction can generalize across various cosmologies and even different initial conditions. As opposed to Lanzieri et al. [2022], the neural network does not have access to power spectra during training, making this a valuable tool for assessing the performance of our correction. This summary statistics is therefore a representative quantification of the performance of our correction, whereas as it is difficult to assess the effectiveness of the Lanzieri et al. [2022] correction beyond the power spectrum (e.g. on other summary statistics).

Another key difference between this work and that of Lanzieri et al. [2022] is that they train a different correction for every set of initial conditions, whereas this work trains a single model capable of evolving the simulation across any initial conditions. The distribution of power spectra across the entire range of simulations is depicted in the right column of Figure 1. We find that 85% of all simulations remain within 30% of the N-body reference. The test simulations with the worse performances are the ones for which the power spectrum of the initial conditions is more than 10%

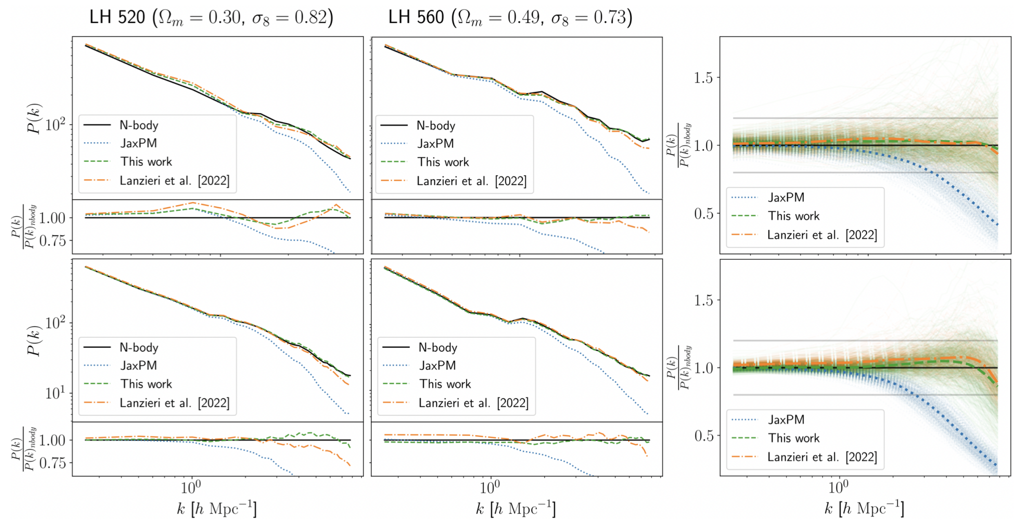

Figure 1: **Top row:** Simulations with 32^3 particles. **Bottom row:** Simulations with 64^3 particles. **Left two columns:** Power spectra of two different simulations for CAMELS (black), JaxPM (blue), this work (green) and from [Lanzieri et al., 2022] (orange). **Right column:** Fractional error of same three methods to the CAMELS simulation for 500 simulations with cosmologies distinct from those seen during training. The thick lines represent the averages. Note that, as opposed to [Lanzieri et al., 2022], the method presented here was not trained with the power spectrum explicitly in the loss.

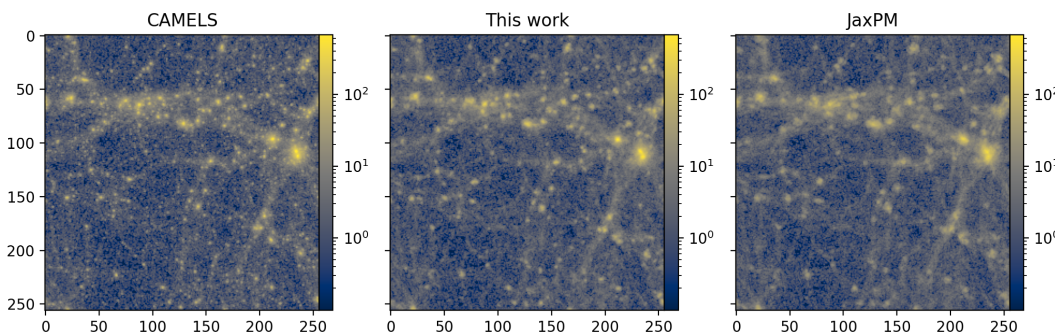

Figure 2: Comparison of z=0 density fields from the CAMELS simulation, our correction to `JaxPM`, and the pure `JaxPM` simulation.

away from those seen during training. We therefore attribute this to the relatively small size of the training set, and believe that better performances could be achieve with more training examples.

4 Cosmological parameter inference

One of the key questions we wish to address is whether our learnt correction to `JaxPM` is close enough to a full N-body simulation that it could be used to perform unbiaised inference of cosmological parameters (specifically, Ω_M and σ_8) in a simulation-based inference (SBI) framework [Cranmer et al., 2020]. As a first test of this, we use the power spectrum of our produced dark matter-only maps as the compressed statistics to train an Sequential Neural Posterior Estimator (SNPE) [Deistler et al., 2022] using an ensemble of 5 masked auto-regressive flows (MAFs) to model the density, by making use of the `sbi` package[2].

Once trained, we use a full N-body simulation obtained through the CAMELS dataset as the data and compute its power spectrum to infer Ω_M and σ_8. The results are presented as the red contours in Fig 3. As a point of comparison, we train another similar SNPE model using `JaxPM` as the simulator to obtain the power spectra to perform the same inference, and obtain the blue contours presented in

[2]https://github.com/mackelab/sbi

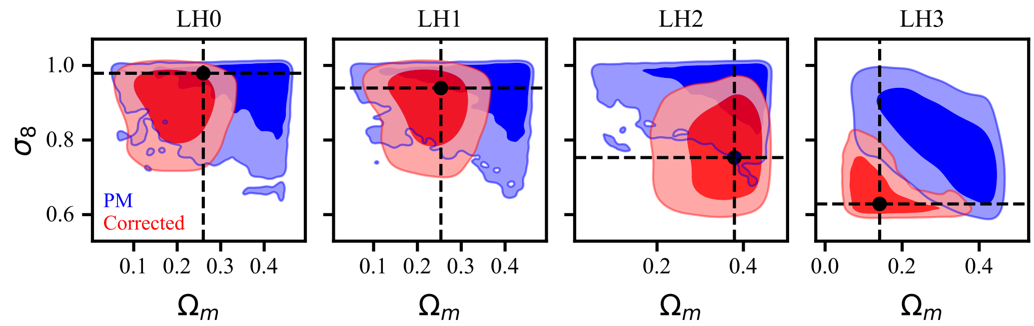

Figure 3: 1 and 2σ constraints on Ω_m and σ_8 obtained in a SBI framework (with SNPE) using `JaxPM` as the simulator (blue contours) and the simulator in this work (red contours). The correction we propose significantly alleviates the bias otherwise induced by `JaxPM` in the inference.

Fig 3. As can be seen, the learnt correction significantly alleviates the biases that otherwise plague the inference.

5 Conclusion

In this work, we have demonstrated how a model trained solely on particle positions and velocities can learn a data-diven correction to the equations of motion of dark matter particles in fast Particle-Mesh simulations, and effectively correct the power spectra for different initial conditions and cosmologies. Furthermore, employing these corrected simulations for cosmological parameter inference effectively mitigates bias arising from small scales inaccuracies in standard `JaxPM` simulations.

Through our choice of loss function, our work emphasizes the importance of imposing the preservation of known conserved quantities. We have found that this, together with learning a correction at the level of the evolution equation rather than directly in the data space, greatly improves robustness of the learnt simulator to different initial conditions and cosmologies. Moreover, this opens the door to more interpretability, as a scheme such as symbolic regression could be used to extract an analytical expression from the learnt correction.

Acknowledgments and Disclosure of Funding

This work is supported by the Simons Collaboration on "Learning the Universe". The Flatiron Institute is supported by the Simons Foundation. The work is in part supported by computational resources provided by Calcul Quebec and the Digital Research Alliance of Canada. Y.H. and L.P. acknowledge support from the Canada Research Chairs Program, the National Sciences and Engineering Council of Canada through grants RGPIN-2020- 05073 and 05102, and the Fonds de recherche du Québec through grants 2022-NC-301305 and 300397. P.L acknowledges support from the Simons Foundation.

References

R.W. Hockney and J.W. Eastwood. *Computer Simulation Using Particles*. CRC Press, 2021. ISBN 978-1-4398-2205-0. URL `https://books.google.ca/books?id=SVslEAAAQBAJ`.

Thomas Buchert. Lagrangian theory of gravitational instability of Friedman-Lemaitre cosmologies and the 'Zel'dovich approximation'. *Monthly Notices of The Royal Astronomical Society*, 254: 729–737, February 1992. doi: 10.1093/mnras/254.4.729.

Thomas Buchert and Jürgen Ehlers. Lagrangian theory of gravitational instability of Friedman–Lemaître cosmologies – second-order approach: an improved model for non-linear clustering. *Monthly Notices of the Royal Astronomical Society*, 264(2):375–387, 09 1993. ISSN 0035-8711. doi: 10.1093/mnras/264.2.375. URL `https://doi.org/10.1093/mnras/264.2.375`.

M. Crocce, S. Pueblas, and R. Scoccimarro. Transients from Initial Conditions in Cosmological Simulations. *Monthly Notices of the Royal Astronomical Society*, 373(1):369–381, November

2006. ISSN 0035-8711, 1365-2966. doi: 10.1111/j.1365-2966.2006.11040.x. URL http://arxiv.org/abs/astro-ph/0606505. arXiv:astro-ph/0606505.

Yu Feng, Man-Yat Chu, Uroš Seljak, and Patrick McDonald. FASTPM: a new scheme for fast simulations of dark matter and haloes. *Monthly Notices of the Royal Astronomical Society*, 463: 2273–2286, December 2016. ISSN 0035-8711. doi: 10.1093/mnras/stw2123. URL https://ui.adsabs.harvard.edu/abs/2016MNRAS.463.2273F. ADS Bibcode: 2016MNRAS.463.2273F.

Svetlin Tassev, Matias Zaldarriaga, and Daniel Eisenstein. Solving Large Scale Structure in Ten Easy Steps with COLA. *J. Cosmol. Astropart. Phys.*, 2013(06):036–036, June 2013. ISSN 1475-7516. doi: 10.1088/1475-7516/2013/06/036. URL http://arxiv.org/abs/1301.0322. arXiv:1301.0322 [astro-ph].

Peter Bodenheimer, editor. *Numerical methods in astrophysics: an introduction*. Series in astronomy and astrophysics. Taylor & Francis, New York, 2007. ISBN 978-0-7503-0883-0. OCLC: ocm70775801.

Siyu He, Yin Li, Yu Feng, Shirley Ho, Siamak Ravanbakhsh, Wei Chen, and Barnabás Póczos. Learning to predict the cosmological structure formation. *Proc. Natl. Acad. Sci. U.S.A.*, 116(28): 13825–13832, July 2019. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1821458116. URL https://pnas.org/doi/full/10.1073/pnas.1821458116.

R. E. Angulo and S. D. M. White. One simulation to fit them all - changing the background parameters of a cosmological N-body simulation. *Monthly Notices of The Royal Astronomical Society*, 405(1): 143–154, June 2010. doi: 10.1111/j.1365-2966.2010.16459.x.

Andrés N. Ruiz, Nelson D. Padilla, Mariano J. Domínguez, and Sofía A. Cora. How accurate is it to update the cosmology of your halo catalogues? *Monthly Notices of the Royal Astronomical Society*, 418(4):2422–2434, 12 2011. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2011.19635.x. URL https://doi.org/10.1111/j.1365-2966.2011.19635.x.

John Joseph M. Carrasco, Simon Foreman, Daniel Green, and Leonardo Senatore. The Effective Field Theory of Large Scale Structures at two loops. *Journal of Cosmology and Astroparticle Physics*, 2014(7):057, July 2014. doi: 10.1088/1475-7516/2014/07/057.

Denise Lanzieri, François Lanusse, and Jean-Luc Starck. Hybrid Physical-Neural ODEs for Fast N-body Simulations. In *International Conference on Machine Learning*, Machine Learning for Astrophysics workshop. 2022.

Georgios Markos Chatziloizos, Tristan Cazenave, and François Lanusse. Deep Learning Modeling of Subgrid Physics in Cosmological N-body Simulations. In *Advances in Neural Information Processing Systems*, Machine Learning and the Physical Sciences workshop, December 2022. URL https://ml4physicalsciences.github.io/2022/files/NeurIPS_ML4PS_2022_45.pdf.

Abylay Zhumekenov, Malika Uteuliyeva, Olzhas Kabdolov, Rustem Takhanov, Zhenisbek Assylbekov, and Alejandro J. Castro. Fourier Neural Networks: A Comparative Study, February 2019. URL http://arxiv.org/abs/1902.03011. arXiv:1902.03011 [cs].

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. *ApJ*, 915(1):71, July 2021. ISSN 0004-637X, 1538-4357. doi: 10.3847/1538-4357/abf7ba. URL https://iopscience.iop.org/article/10.3847/1538-4357/abf7ba.

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. 2023. doi: 10.48550/ARXIV.2304.02096. URL https://arxiv.org/abs/2304.02096. Publisher: arXiv Version Number: 1.

Francisco Villaescusa-Navarro, Shy Genel, Daniel Anglés-Alcázar, Lucia A. Perez, Pablo Villanueva-Domingo, Digvijay Wadekar, Helen Shao, Faizan G. Mohammad, Sultan Hassan, Emily Moser, Erwin T. Lau, Luis Fernando Machado Poletti Valle, Andrina Nicola, Leander Thiele, Yongseok Jo, Oliver H. E. Philcox, Benjamin D. Oppenheimer, Megan Tillman, ChangHoon Hahn, Neerav Kaushal, Alice Pisani, Matthew Gebhardt, Ana Maria Delgado, Joyce Caliendo, Christina Kreisch, Kaze W. K. Wong, William R. Coulton, Michael Eickenberg, Gabriele Parimbelli, Yueying Ni, Ulrich P. Steinwandel, Valentina La Torre, Romeel Dave, Nicholas Battaglia, Daisuke Nagai, David N. Spergel, Lars Hernquist, Blakesley Burkhart, Desika Narayanan, Benjamin Wandelt, Rachel S. Somerville, Greg L. Bryan, Matteo Viel, Yin Li, Vid Irsic, Katarina Kraljic, and Mark Vogelsberger. The CAMELS project: public data release. *ApJS*, 265(2):54, April 2023. ISSN 0067-0049, 1538-4365. doi: 10.3847/1538-4365/acbf47. URL `http://arxiv.org/abs/2201.01300`. arXiv:2201.01300 [astro-ph].

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. doi: 10.1186/s40668-019-0028-x.

Kyle Cranmer, Johann Brehmer, and Gilles Louppe. The frontier of simulation-based inference. *Proceedings of the National Academy of Science*, 117(48):30055–30062, December 2020. doi: 10.1073/pnas.1912789117.

Michael Deistler, Pedro J Goncalves, and Jakob H Macke. Truncated proposals for scalable and hassle-free simulation-based inference. *arXiv e-prints*, art. arXiv:2210.04815, October 2022. doi: 10.48550/arXiv.2210.04815.