Predicting Cosmological Massive Neutrino Simulation with Convolutional Neural Networks

Elena Giusarma
Flatiron Institute, New York City, NY, USA
Michigan Technological University, Houghton, MI, USA
egiusarm@mtu.edu

Mauricio Reyes Hurtado
Michigan Technological University, Houghton, MI, USA

Francisco Villaescusa-Navarro, Siyu He, Shirley Ho
Flatiron Institute, New York City, NY, USA

Abstract

Neutrinos are elusive subatomic particles that play an important role in the evolution of the Universe. Despite they are the second most abundant particle of the Universe, their properties remain still unknown. Cosmological observations provide a powerful probe of neutrino properties, and in particular of their mass, M_ν. In order to maximize the information that can be retrieved from those observations, accurate theoretical predictions are needed. Currently, simulations are the best tool to provide those. In this work, we propose a new method, based on a deep learning network, to quickly generate cosmological simulations with massive neutrinos from standard simulations without neutrinos. We show that our approach can predict accurately cosmological simulations and can generate new massive neutrino simulations 10,000 times faster than traditional methods.

1 Introduction

The discovery of neutrino oscillations, which resulted in the 2015 Nobel Prize in Physics, has robustly established the fact that neutrinos are massive [6, 3, 2]. Cosmology provides an independent tool to constrain the neutrino masses. The light massive neutrinos are relativistic in the early Universe and contribute to the radiation energy density during that epoch. At late times, they became non-relativistic, contributing to the total matter density of the Universe. Relic neutrinos thus leave a characteristic signature on different cosmological observables, affecting the background evolution, the spectra of matter perturbations and Cosmic Microwave Background (CMB) anisotropies. The tightest constraints on M_ν, which combine different cosmological data, range from 0.22 eV to 0.12 eV at 95% confidence level [10, 11, 4, 13, 5, 9].

Upcoming galaxy surveys will provide one of the most important sources of information to constrain the neutrino masses and their mass ordering. To achieve those results, the need for accurate theoretical predictions in the presence of massive neutrino will be crucial. A powerful tool to obtain rigorous predictions in cosmology is by running numerical cosmological simulations. Over the past years, a large number of N-body simulations, including massive neutrinos, have been proposed and developed in the literature, [1, 16, 8, 15]. However, those simulations are computationally expensive since they require a high number of CPU hours to complete – 700 CPU hours for each neutrino mass case.

Second Workshop on Machine Learning and the Physical Sciences (NeurIPS 2019), Vancouver, Canada.

In this paper, we develop a convolutional neural network (U-Net), [12, 7] to establish the mapping between simulations with massless and massive neutrinos. The important advantage of our approach consists in the ability to generate complex numerical simulations from the standard ones in an extremely fast way and with high accuracy compared to the traditional N-body simulation techniques.

2 N-body simulation data

We use two subsets of HADES simulations [14], the precursor of the QUIJOTE N-body simulations [15]: the standard (ΛCDM) simulations (without neutrinos), and simulations with neutrino masses. For neutrino simulations, we assume three degenerate neutrino masses. The value of the cosmological parameters are: $h = 0.6711$, $\Omega_m = 0.3175$, $\Omega_b = 0.049$, $\Omega_\Lambda = 0.6825$, $n_s = 0.9624$ and $A_s = 2.13$ and $\sigma_8 = 0.833$ ($\sigma_8 = 0.798$ when we include massive neutrinos). All the simulations are run in a periodic box size of $1\ h^{-1}\mathrm{Gpc}$ and follow the evolution of the cold dark matter and neutrino particles (only in the case of simulations with massive neutrinos), from $z = 99$ to $z = 0$. We use 100 independent realizations for each model, containing the same number of cold dark matter and neutrino particles ($N_{\mathrm{cdm}} = N_\nu = 512^3$). We focus our analysis at the present time ($z = 0$).

In order to speed up the process of training the neural network and avoid GPUs memory problems, we separate each of the 100 simulations into sub-cubes of size 32^3 voxels, corresponding to a volume of $62.5h^{-1}\mathrm{Mpc}$ on each side. Each realization contains $4,096$ sub-cubes for a total of $409,600$ data for each model. We then split the realization of each model into three chunks: 70% (70 realizations) for training, 20% (20 realizations) for validation and 10% (10 realizations) for testing.

3 Method

Network Architecture We use a modified version of the deep neural network, Deep Density Displacement Model (D^{3}M), introduced by [7]. D^{3}M is a generalization of the standard U-Net, first proposed in [12] for bio-medical image segmentation, to work with three-dimensional data. It consists of a *contracting path* and an *expansive path*. The contracting path follows the typical architecture of a convolutional neural network, and it includes two blocks. Each block consists of the repeated application of two convolutions with stride 1 and a down-sampling convolution with stride 2 each followed by batch normalization (BN) and a rectified linear unit (ReLU). For each convolutional layer, we use $3 \times 3 \times 3$ filters, and we apply a zero padding with size 1. At each down-sampling step, we reduce the spatial information by half and we increase by double the number of feature channels. The bottom of the U-Net consists of two convolutional neural networks followed by the same padding, BN and ReLU. The expansive path includes two repeated expansion blocks each of which consists of one up-sampling convolution with stride 1/2 that halves the number of feature channels, a concatenation, and two successive convolution with stride 1, each followed by zero padding, BN and ReLU. At the final layer, a $1 \times 1 \times 1$ convolution without padding is used to map 64 features to the final 3-D data. In total the network present 15 convolutional layers.

Training We train our model using the displacement field, defined as the difference between the final position of the particles (at redshift $z = 0$), $\mathbf{\Psi}_f$, and the Largangian position of the same particles on a uniform grid, $\mathbf{\Psi}_i$. The input of our deep neural network and the target are the displacement field from N-body simulations without and with massive neutrinos. Our benchmark model consists of Zel'dovich Approximation (hereafter ZA) with massive neutrinos. The ZA is a significantly faster approximation to N-body simulations produced by first-order perturbation theory with neutrinos.

We train our neural network by using the mean absolute scaled error loss function that is given by the sum of all the absolute differences between the target displacement field and the predicted values.

4 Results

We quantify the agreement between our model and numerical simulation by considering four different summary statistics: the 1-D probability distribution function (PDF), the power spectrum, the bispectrum, and the void size function.

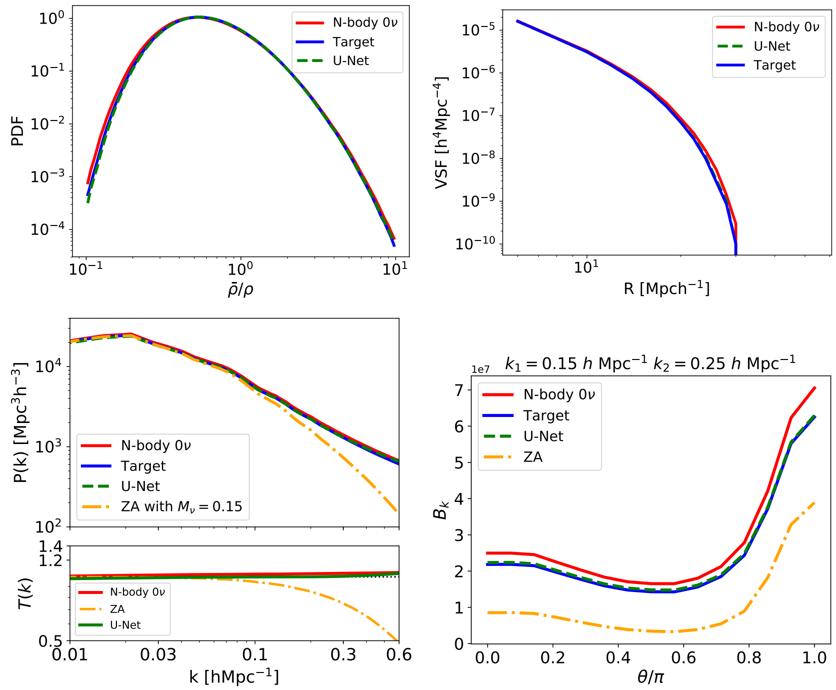

Figure 1: The plot shows the four summary statistics performed in the analysis: the 1D PDF (top-left panel), the power spectrum and the transfer function (bottom-left panel), the bispectrum (bottom-right panel), and the void abundance (top-right panel). In all panels, the red and blue lines refer to the cosmological simulations without and with massive neutrinos, the green dashed line indicates the result of our U-Net model and the orange dashed-dot line denotes the ZA approximation.

1-D Probability distribution function We compute the 1-D PDF of the Cold Dark Matter (CDM) for the N-body simulations and U-Net model. The PDF provides information on the distribution of CDM overdensities in the cells of the particles grid.

The results are depicted in Figure 1, top-left panel, where we show the comparison of the 1-D PDF among the standard simulations (the input, red line),massive neutrino simulations (the target, blue line) and our U-Net model (green dashed line). We can clearly see the perfect agreement between our U-Net model and N-body simulations with massive neutrinos.

Power Spectrum The most widely statistic used in cosmology to extract information from observation is the two-point correlation function $\xi(r)$, which is defined as the excess probability, compared with a random distribution of galaxies, of finding a pair of galaxies at a given separation. The Fourier transform of the two-point correlation function is the power spectrum, $P(k)$:

$$P(k) = \int d^3\vec{r}\xi(r)e^{i\vec{k}\cdot\vec{r}} \,, \tag{1}$$

where k is wavenumber of a fluctuation. The power spectrum is the quantity predicted directly by theories for the formation of large-scale structure, and in the case of a Gaussian density field, it provides a complete statistical description of the fluctuations.

In order to quantify the performance of our model against the target, we compute the transfer function $T(k)$, defined as the square root of the ration of the two power spectra:

$$T(k) = \sqrt{\frac{P_{\text{pred}}(k)}{P_{\text{target}}(k)}} \,, \tag{2}$$

where $P_{\text{pred}}(k)$ is the density power spectrum predicted by our U-Net model, and $P_{\text{target}}(k)$ is the equivalent from N-body simulations with massive neutrinos.

Figure 1 bottom-right panel, shows the average power spectrum and the transfer function of the density field over ten simulations. The red and blue lines of the $P(k)$'s panel correspond to the power spectrum of N-body simulations without and and with massive neutrinos simulations respectively. We can note that massive neutrinos induce a suppression of the power spectrum at small scales (larger k). The green and orange lines depict the power spectrum from U-Net and benchmark model (Zel'dovich approximation). Notice that the latter performs accurately the power spectrum on very large-scales ($k < 0.05$ h/Mpc). We find that the density distribution from the U-Net sample has, on large-scales ($k < 0.1$ h/Moc), an average power spectrum which is very close in amplitude and in shape to that of N-body simulations with neutrinos (the target). On smaller scales, the power spectrum of our model departs from that of the target. To quantify this deviation, we focus on the $T(k)$'s panel, that shows the transfer function as a function of scales for the benchmark model, the U-Net and the input. For a perfectly accurate prediction, the transfer function is expected to be 1. The density transfer function of the ZA approximation is close to one up to $k < 0.05$ h/Mpc and then decreases as expected. On the other hand, the density transfer function of the U-Net prediction is approximately 1 for $0.03 < k < 0.5$ h/Mpc and differs from the unity by a 3.5% at scale $k \approx 0.55$ h/Mpc. This discrepancy increases to 6.3% as k increases around 0.7 h/Mpc. Those results suggest that our model manages to predict accurately the power spectrum with massive neutrinos from large to intermediate scales. On smaller scales ($k > 0.6$ h/Mpc) the prediction starts to deviate from the target. This is not a surprise since those are the scales where the effects of non-linear evolution are important and very difficult to solve accurately.

Bispectrum The inflationary scenario predicts Gaussian initial conditions, characterized by the two-point correlation function (2PCF) or the power spectrum. However, on small scales, nonlinear gravitational instability induces non-Gaussian signatures in the mass distribution, which contain information on the nature of gravity and the dark matter. The lowest order statistical tool to describe a non-Gaussian field is the three-point correlation function (3PCF), or equivalently, its Fourier transform, the bispectrum, defined as: $B(k1, k2, k3) = \langle \delta_{\mathbf{k_1}} \delta_{\mathbf{k_2}} \delta_{\mathbf{k_3}} \rangle$.

This observable plays an important role in cosmology because it carries information about the spatial coherence of large-scale structures and can place strong constraints on models of structure formation. The presence of massive neutrinos induces a suppression of the amplitude of the bispectrum. Figure 1 bottom-right panel, shows the average bispectrum over 10 simulations for the U-Net model (green-dashed line), the benchmark model (orange line), the input (red line) and the target (blue line). We find that the mean relative bispectrum residual of our model compared to the target is 0.4%. This suggests that the U-Net model can generate neutrino simulations with correct high-order statistics in the mildly non-linear regime. Notice that on the other hand, the benchmark model is not able to reproduce the bispectrum from the target. This is not surprising as the ZA approximation we are using as benchmark model, is only valid on larger linear scales.

Voids Abundance Voids are the most under-dense regions of our Universe and together with halos, filaments, and walls, constitute the large-scale structure of the Universe, known as the cosmic web. Voids enclose a large amount of cosmological information since they fill most of the volume of the Universe. Massive neutrinos modify the properties of cosmic voids, such as their number, size and shape.

The last statistical tool we consider here is the void size function, defined as the abundance of voids as a function of the voids radius. In presence of massive neutrinos the abundance of large voids is suppressed [8]. In the upper-right panel of Figure 1, we compare the average void size function over 10 realizations for the U-Net model (green line), the input (red line), and the target (blue line). We can clearly see the suppression of the voids abundance induced by massive neutrinos at larger radius. We can also note the good agreement in the void size function among the the U-Net model and the target simulations. This indicates that our U-Net model is also able to capture the voids' information completely.

5 Conclusion

In this work, we have presented a new a deep learning approach to predict fast non-standard cosmological simulations with massive neutrinos from standard N-body simulations. We have shown that our model manages to learn the mapping between the displacement field in simulations with massless and massive neutrinos. Moreover, it allows to generate massive neutrino simulations five order of magnitude faster than the traditional methods and dramatically reduce the computational costs necessary of implementing simulations with massive neutrinos. The use of such approach will be particularly useful in the near future. The need for fast N-body simulations will increase with the upcoming large cosmological surveys such as Euclid and LSST. Our methods will thus be crucial for producing fast and large simulation data for cosmological analyses. In future work we will explore the possibility of generating fast neutrino simulations with an arbitrary mass from standard simulations or more complex non-standard cosmological simulations. A further direction is also to study whether the same technique can be used to learn the mapping between the second-order Lagrangian perturbation theory (2LPT) and the Zel'dovich approximation without massive neutrinos.

Acknowledgements: We thank Gabriella Contardo, Barnabas Pócsoz, Siamak Ravanbakhsh, and David Spergel for useful discussions. This project is supported by the Simons Foundation.

References

[1] Simeon Bird, Matteo Viel, and Martin G. Haehnelt. Massive Neutrinos and the Non-linear Matter Power Spectrum. *Mon. Not. Roy. Astron. Soc.*, 420:2551–2561, 2012.

[2] Ivan Esteban, M. C. Gonzalez-Garcia, Michele Maltoni, Ivan Martinez-Soler, and Thomas Schwetz. Updated fit to three neutrino mixing: exploring the accelerator-reactor complementarity. *J. High Energy Phys.*, 2017.

[3] D. V. Forero, M. Tórtola, and J. W.F. Valle. Neutrino oscillations refitted. *Phys. Rev. D - Part. Fields, Gravit. Cosmol.*, 2014.

[4] Elena Giusarma, Martina Gerbino, Olga Mena, Sunny Vagnozzi, Shirley Ho, and Katherine Freese. Improvement of cosmological neutrino mass bounds. *Phys. Rev.*, D94(8):083522, 2016.

[5] Elena Giusarma, Sunny Vagnozzi, Shirley Ho, Simone Ferraro, Katherine Freese, Rocky Kamen-Rubio, and Kam-Biu Luk. Scale-dependent galaxy bias, CMB lensing-galaxy cross-correlation, and neutrino masses. *Phys. Rev.*, D98(12):123526, 2018.

[6] M. C. Gonzalez-Garcia, Michele Maltoni, and Thomas Schwetz. Updated fit to three neutrino mixing: status of leptonic CP violation. *JHEP*, 11:052, 2014.

[7] Siyu He, Yin Li, Yu Feng, Shirley Ho, Siamak Ravanbakhsh, Wei Chen, and Barnabás Póczos. Learning to Predict the Cosmological Structure Formation. *Proc. Nat. Acad. Sci.*, 116(28):13825–13832, 2019.

[8] Elena Massara, Francisco Villaescusa-Navarro, Matteo Viel, and P. M. Sutter. Voids in massive neutrino cosmologies. *JCAP*, 1511(11):018, 2015.

[9] Nathalie Palanque-Delabrouille, Christophe Yeche, Julien Baur, Christophe Magneville, Graziano Rossi, Julien Lesgourgues, Arnaud Borde, Etienne Burtin, Jean-Marc LeGoff, James Rich, Matteo Viel, and David Weinberg. Neutrino masses and cosmology with {Lyman}-alpha forest power spectrum. *J. Cosmol. Astropart. Phys.*, 2015(11):11, nov 2015.

[10] Planck Collaboration and et al. Aghanim, N. Planck 2018 results. VI. Cosmological parameters. *arXiv e-prints*, Jul 2018.

[11] N. et al. Planck Collaboration, Aghanim. Planck 2018 results. {V}. {CMB} power spectra and likelihoods. *arXiv e-prints*, 2019.

[12] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 2015.

[13] Sunny Vagnozzi, Elena Giusarma, Olga Mena, Katherine Freese, Martina Gerbino, Shirley Ho, and Massimiliano Lattanzi. Unveiling \${}nu$ secrets with cosmological data: neutrino masses and mass hierarchy. *Phys. Rev. D*, 96(12):123503, dec 2017.

[14] Francisco Villaescusa-Navarro, Arka Banerjee, Neal Dalal, Emanuele Castorina, Roman Scoccimarro, Raul Angulo, and David N. Spergel. The imprint of neutrinos on clustering in redshift-space. *Astrophys. J.*, 861(1):53, 2018.

[15] Francisco Villaescusa-Navarro et al. The Quijote simulations. *arXiv e-prints*, 2019.

[16] Francisco Villaescusa-Navarro, Mark Vogelsberger, Matteo Viel, and Abraham Loeb. Neutrino Signatures on the High Transmission Regions of the Lyman-alpha Forest. *Mon. Not. Roy. Astron. Soc.*, 431:3670, 2013.