Fast and Accurate Non-Linear Predictions of Universes with Deep Learning

Renan Alves de Oliveira  
PPGCosmo/Flatiron Institute  
fisica.renan@gmail.com

Yin Li  
Flatiron Institute  
yinli@flatironinstitute.org

Francisco Villaescusa-Navarro  
Princeton University  
villaescusa.francisco@gmail.com

Shirley Ho  
Flatiron Institute  
shirleysurelyho@gmail.com

David N. Spergel  
Flatiron Institute  
dspergel@flatironinstitute.org

Abstract

Cosmologists aim to model the evolution of initially low amplitude Gaussian density fluctuations into the highly non-linear "cosmic web" of galaxies and clusters. They aim to compare simulations of this structure formation process with observations of large-scale structure traced by galaxies and infer the properties of the dark energy and dark matter that make up 95% of the universe. These ensembles of simulations of billions of galaxies are computationally demanding, so that more efficient approaches to tracing the non-linear growth of structure are needed. We build a V-Net based model that transforms fast linear predictions into fully non-linear predictions from numerical simulations. Our NN model learns to emulate the simulations down to small scales and is both faster and more accurate than the current state-of-the-art approximate methods. It also achieves comparable accuracy when tested on universes of significantly different cosmological parameters from the one used in training. This suggests that our model generalizes well beyond our training set.

1 Introduction

Cosmology is at critical stage: the 2019 Nobel Prize recognized that a simple model with only five basic parameters fits a host of astronomical observations on large-scales where the fluctuations are in the linear regime. This model implies that atoms make up only 5% of the universe with the remaining 95% in the form of dark matter and dark energy. Astronomers have embarked on several large-scale surveys and are launching multiple satellites that aim to collect data to understand these mysteries. Analysis of this new data is limited by our computational abilities as numerical simulations are essential for comparing theoretical predictions to observations on the small non-linear scales that contain most of the cosmological information [1,2,3,4,5]. Unfortunately, the computational cost of these simulations is high; a single simulation may require from thousands to millions of CPU hours.

In this work, we aim to build a neural network (NN) model that can learn to emulate these numerical simulations both accurately and efficiently. This paper builds on the work of He et al. [6], which showed that neural networks can indeed learn to run fast approximate simulations. Our main
contributions is improvement of the model by training on full N-body simulations and extending to much smaller scales, where the learning task becomes more difficult due to nonlinearity.

2 Methods

**N-body Simulation.** An N-body simulation involves a large number of massive dark matter particles interacting with each other only through Newtonian gravity. Since our Universe began with matter distributed almost uniformly, the simulation starts with particles only slightly perturbed from a uniform grid and nearly at rest. During the simulation, a particle moves from its initial location \( q \) to its final position \( x = q + \Psi(q) \), where \( \Psi \) is the displacement vector. N-body simulations usually end at our current time, i.e. "today"; this is the time at which we compare them to our ML predictions. The left panel of Fig. 1 shows an example of a simulation output.

We use 210 N-body simulations from the Quijote suite [7] for training (180), validation (20), and testing (10). Each simulation contains \( 512^3 \) particles in a box of \( 1 \) (Gpc/h)^3 volume (nearly 5 billion light years in size). In contrast to [6], the training set data was fast approximation simulations using FastPM [8] (see next section), with \( 32^3 \) particles in a volume of \( 128 \) (Mpc/h)^3. The simulated universes resemble our own, and can be characterized by the value of five cosmological parameters: matter density parameter \( \Omega_m = 0.3175 \) (this means 31.7% of the Universe is made of dark matter), baryon density parameter \( \Omega_b = 0.049 \), dark energy density parameter \( \Omega_k = 0.6825 \), Hubble parameter \( h = 0.6711 \), matter fluctuation within a 8 Mpc/h sphere \( \sigma_8 = 0.834 \), primordial spectral index \( n_s = 0.9624 \). In addition, we also test our model, trained on the above simulations, on three Quijote simulations with different values of the cosmological parameters, but the same configuration otherwise.

**Fast Approximation.** N-body simulations can be very computationally expensive, as they typically solve the dynamics of billions of particles for thousands of time steps. Fast approximate simulations [9, 8] are usually adopted when a large number of simulations are needed. These methods save computation time by only integrating tens of time steps; they are thus less accurate than full N-body simulations. We aim to build an NN model that is more accurate and faster than these approximators. In this work we use the widely used method COLA (COmoving Lagrangian Acceleration) [9], as implemented in the publicly available package L-PICOLA [10], as a benchmark. COLA solves the particle motions relative to predictions by the second order Lagrangian perturbation theory [11]. We setup L-PICOLA using the same configurations as the full N-body simulations, running it for only 10 time steps.

**Linear Theory.** Fast and accurate solutions exist when matter distribution is close to uniform, and density fluctuations are small. For instance, at linear order, particles move along straight lines, with the distance determined by the growth function \( D \) of time: \( \Psi_{\text{lin}}(q, t) = D(t)/D(t_0) \Psi_{\text{lin}}(q, t_0) \). At early times \( t_0 \to 0 \), such as the starting time of the simulations, \( \Psi_{\text{lin}}(t_0) \) agrees very well with the simulation prediction for \( \Psi(t_0) \) due to uniformity.

The linear theory prediction is a very good approximation on large scales, where density fluctuations are small. However, on small scales, the density contrast increases drastically and structure formation becomes non-linear, limiting the validity of linear theory predictions. Therefore, we use \( \Psi_{\text{lin}}(q, t) \) as the input to our NN model, which predicts the fully non-linear target \( \Psi(q, t) \) given by the N-body simulations at the same \( t \). By design, our NN will make accurate predictions on large scales. This is true even if we test the model on universes with cosmological parameters different from the one used for training, as we show below.

**Neural Network Model.** Both the input (linear theory) \( \Psi_{\text{lin}} \)'s and target (N-body simulation) \( \Psi \)'s of our NN model are functions of \( q \)'s, that form a uniform grid. So each of them is a displacement field that can be viewed as a 3D image, with three channels being the three Cartesian components of the displacement vectors. This allows us to apply many computer vision models to our problem.

In this work we adopt a simple U-Net / V-Net [12, 13] type architecture similar to that in [6]. The model works on 3 levels of resolution connected in a "V" shape by 2 downsampling layers and 2 upsampling layers, achieved by stride-2 2\(^3\) convolutions and stride-\(\frac{1}{2}\) 2\(^3\) transposed convolutions, respectively. Blocks of 2 3\(^3\) convolutions connect the input, the resampling, and the output layers. As in V-Net, a residual connection, in this case, 1\(^3\) convolutions instead of identity, are added over
We consider ten Quijote simulations from the test set (see the left panel of Fig. 1 for an example),

We define a transfer function,

$$T$$

we find that a value of

To preserve the rotational equivariance of the simulation box, we implement data augmentation that

from that in the training simulations. In particular, we varied all five cosmological parameters:

$$k$$

produces percent-level accurate results down to scales as small as

Results

Accuracy We quantify model accuracy by compare the simulation power spectra with the

power spectra estimated by the NN and by the benchmark method. The density power spectrum

quantifies the correlation of density fluctuations as a function of scale: $P_δ(k_i) = \langle N_i V \rangle^{-1} \sum_{k_i < k \leq k_{i+1}} \delta(k)\delta(-k)$, where $V$ is the volume, $δ(k)$ is the Fourier transform of $δ(x)$, $k_i$’s are bin edges for the wavevector $k$, and $N_i$ is the number of $k$’s falling in $[k_i, k_{i+1}]$. The wavenumber $k ≡ |k|$ describes the scale, with low/high $k$ representing large/small scales. The cross-power spectrum, $δ_{pred}(k)δ_{true}(-k)$ measures the covariance between the predicted and target density fields.

We define a transfer function, $T(k) = \sqrt{P_{pred}(k)/P_{true}(k)}$, where $P_{pred}(k)$ and $P_{true}(k)$ are the power spectra of the prediction and the simulation, respectively. If $T(k) = 1$, the predictions accurately capture the amplitude of the density field. The cross-correlation coefficient, $r$, measures the phase correlations: $r(k) = P_{pred×true}(k)/\sqrt{P_{pred}(k)P_{true}(k)}$, with the numerator being the cross-power spectrum between the prediction and the simulation. $T$ and $r$ are good estimators to quantify the accuracy of the model, because when they are both 1 the prediction and target are identical as proved in Ref. [6]. Instead of $r$, we use $1 - r^2$, which gives the amount of unexplained variance between the two fields. Similar to $δ$, we can also compute $P(k), T(k)$, and $r(k)$ for $Ψ$ by simply replacing the scalar product with the dot vector product.

We consider ten Quijote simulations from the test set (see the left panel of Fig. 1 for an example),

together with their fast approximate and NN counterparts. From each of them, we compute their

displacement and density power spectra, as well as the corresponding $T$ and $1 - r^2$. We show the

results in the central and right panels of Fig. [1]. We find that the NN outperforms the benchmark

model in accuracy on all scales and for all considered quantities. We emphasize that our NN model

produces percent-level accurate results down to scales as small as $k ≃ 1 \text{ h/Mpc}$.

Generalization to Different Universes. We test the NN model on 2000 universes with drastically different cosmological parameters, where structure formation proceeds in a very different way from that in the training simulations. In particular, we varied all five cosmological parameters:
Figure 1: The left panel shows an example of the projected density field in a Universe as predicted by an N-body simulation. We quantify the accuracy of the NN against the fast approximator and the N-body simulation using the power spectrum of the displacement field (middle panel) and the density contrast (right panel). The NN model outperforms the fast simulator in all cases.

Figure 2: Accuracy comparison between predictions by the fast approximator (blue dot-dashed) and our NN (green dashed) for universes with cosmological parameters very different to those used in training: $\{\Omega_m, \Omega_b, h, n_s, \sigma_8\}$ equal to $\{0.1755, 0.0668, 0.7737, 0.8849, 0.6641\}$ (left), $\{0.3889, 0.0363, 0.6275, 0.9513, 0.9137\}$ (middle), $\{0.4291, 0.0337, 0.5511, 1.1687, 0.6619\}$ (right). This demonstrates that our NN generalizes very well.

$\{\Omega_m, \Omega_b, h, n_s, \sigma_8\}$ that are relevant to the simulations. We show the prediction comparisons for the density field in Fig. 2 and find that the NN generalizes very well for cosmologies with low $\Omega_b/\Omega_m$ and achieves a similar accuracy to that in Fig. 1 (same conclusion holds for the displacement field). It also outperforms the fast approximator COLA, even though COLA depends explicitly on the varied cosmological parameters. This result is in concordance with that in Ref. 6, where $\Omega_m$ and $\sigma_8$ were varied individually.

**Efficiency.** We now compare the computational cost of the NN versus the fast approximation benchmark, as well as that of the N-body simulations. We show in Table 1 the total inference time to obtain the whole 512$^3$ displacement field using 64 128$^3$ cubes.

The fast approximation uses 20 CPU cores (Intel Xeon CPU E5-2640 v4), and the NN inference uses PyTorch [16] on 1 GPU (NVIDIA Tesla P100-PCIE-16GB). The N-body simulation requires 500 CPU hours on 48 cores, or $\sim 10^5$ seconds if we scale it to 20 cores to compare with the fast approximation. Our NN model is two times faster than the fast approximation benchmark, and more than $10^5$ times faster than the N-body simulation.
Table 1: Runtime benchmark. The NN model is $2 \times$ faster than the fast approximator, and more than $10^3 \times$ faster than the $N$-body simulation.

<table>
<thead>
<tr>
<th></th>
<th>QUIOTE</th>
<th>L-PICOLA</th>
<th>PyTorch-GPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wall Time (s)</td>
<td>$10^9$</td>
<td>112</td>
<td>59</td>
</tr>
</tbody>
</table>

4 Conclusions

In this paper, we have shown that neural networks are able to accurately emulate the output of expensive N-body simulations. We have shown that our model reproduces the results of full $N$-body simulations down to scales as small as $k = 1 \, h/\text{Mpc}$ at present time to within $\sim 1\%$ accuracy. When compared to the state-of-the-art N-body simulation approximator in cosmology, our V-Net based network achieves better accuracy in less time. Furthermore, we have demonstrated that our network generalizes extremely well, by showing a comparable level of accuracy for universes that it has not been trained on. Our method represents a big step forward in the direction of reducing the computational time needed to provide theory predictions in the non-linear regime. This paves the way to maximizing the scientific return of current and upcoming billion dollar astronomical missions.

Broader Impact

We think the ML methods used in this work may be useful to motivate and enlighten students from different fields, especially those who are interested in astronomy, and to use new tools to help us for a better understanding of our Universe. Moreover, because of the upcoming astronomical probes for the next ten years, we believe that the outcomes of this work might help in future scientists run fast simulations of universes with high precision using a regular computer.

Acknowledgments and Disclosure of Funding

RAO thanks Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Simons Foundation. YL thanks the Simons Foundation for support of fellowship. The Flatiron Institute is supported by the Simons Foundation. We acknowledge that our work was partly performed using the Princeton Research Computing resources at Princeton University which is consortium of groups led by the Princeton Institute for Computational Science and Engineering (PICSciE) and Office of Information Technology’s Research Computing, and made use of the CHE cluster, managed and funded by the COSMO/CBPF/MCTI, with financial support from FINEP and FAPERJ, and operating at Javier Magnin Computing Center/CBPF.

References


