MG-NECOLA: Fast Neural Emulators for Modified Gravity Cosmologies

J. Bayron Orjuela-Quintana
Universidad del Valle
Santiago de Cali 760032, Colombia
john.orjuela@correounivalle.edu.co

Mauricio Reyes
Michigan Technological University
Houghton, MI 49931, USA
mrhurtad@mtu.edu

Elena Giusarma
Michigan Technological University
Houghton, MI 49931, USA
egiusarm@mtu.edu

Francisco Villaescusa-Navarro
Center for Computational Astrophysics
162 5th Avenue, New York, NY 10010, USA
fvillaescusa@flatironinstitute.org

Neerav Kaushal
Michigan Technological University
Houghton, MI 49931, USA
kaushal@mtu.edu

César A. Valenzuela-Toledo
Universidad del Valle
Santiago de Cali 760032, Colombia
cesar.valenzuela@correounivalle.edu.co

Abstract

Observations of the large-scale structure (LSS) provide a powerful test of gravity on cosmological scales, but high-resolution N-body simulations of modified gravity (MG) are prohibitively expensive. We present `MG-NECOLA`, a convolutional neural network that enhances fast `MG-PICOLA` simulations to near–N-body fidelity at a fraction of the cost. `MG-NECOLA` reproduces `QUIJOTE-MG` N-body results in the power spectrum and bispectrum with better than 1% accuracy down to non-linear scales ($k \simeq 1\ h\ \mathrm{Mpc}^{-1}$), while reducing computational time by several orders of magnitude. Importantly, although trained only on $f(R)$ models with massless neutrinos, the network generalizes robustly to scenarios with massive neutrinos, preserving accuracy to within 5% at non-linear scales. This combination of precision and robustness establishes `MG-NECOLA` as a practical emulator for producing large ensembles of high-fidelity simulations, enabling efficient exploration of modified gravity and beyond-ΛCDM cosmologies in upcoming surveys.

1 Introduction

Modern cosmology and astrophysics increasingly rely on accurate theoretical predictions to extract information from the wealth of data provided by current and upcoming galaxy surveys. Statistical measures such as the matter power spectrum, correlation functions, and higher-order statistics encode the imprint of fundamental physics on the large-scale structure (LSS) of the Universe. The prevailing framework, known as the Λ Cold Dark Matter (ΛCDM) model, combines General Relativity (GR) with a Universe dominated by cold dark matter and dark energy. While ΛCDM has been highly successful in explaining a broad range of observations, growing evidence suggests possible inconsistencies between different cosmological probes [1, 2, 12]. These observational hints, together with the fact that GR has never been directly tested on the largest cosmic scales, provide strong motivation to explore *modified gravity* (MG) as a compelling extension of the standard model. MG theories alter the growth of cosmic structures by introducing additional degrees of freedom that can enhance or suppress gravitational interactions on cosmological scales. The next generation of

surveys will deliver unprecedentedly precise measurements of LSS, creating a unique opportunity to test these alternatives to GR and potentially uncover new physics.

To exploit this opportunity, fast and accurate theoretical predictions are essential. High-resolution N-body simulations provide the most reliable predictions for non-linear structure formation, but their computational cost makes them impractical for exploring the vast parameter spaces of MG cosmologies. Approximate methods, such as the *COmoving Lagrangian Acceleration* (COLA) [20] approach, are far more efficient and reproduce large-scale clustering well, but their accuracy deteriorates on non-linear scales where MG effects are most pronounced. We address this challenge with `MG-NECOLA`, a deep learning framework based on convolutional neural networks (CNNs) that maps fast COLA simulations into high-fidelity N-body–like realizations. Trained on the `QUIJOTE-MG` suite [5, 21] of extended cosmologies, `MG-NECOLA` recovers key summary statistics such as the matter power spectrum down to $k \simeq 1\ h\ \mathrm{Mpc}^{-1}$ while remaining orders of magnitude faster than full N-body runs. This enables efficient exploration of MG scenarios and provides a scalable tool for large-scale structure analyses in upcoming surveys.

2 Method

In the standard cosmological paradigm, the growth of structure is governed by GR, where matter density perturbations source the Newtonian potential. MG theories extend this framework by introducing additional degrees of freedom that rescale the gravitational strength in a scale- and time-dependent way [3, 10]. A widely studied example is the Hu–Sawicki $f(R)$ model [14], where a scalar degree of freedom mediates an additional force, leaving characteristic signatures in the matter power spectrum at non-linear scales [11, 19]. The strength of this force is controlled by the parameter f_{R_0}, with observationally allowed values typically satisfying $|f_{R_0}| < 10^{-5}$ [6, 9]. In this work, we focus on $f(R)$ gravity as a benchmark to evaluate our deep learning framework, `MG-NECOLA`.

2.1 Simulation Data:

Our framework requires both high-fidelity references and fast approximate simulations. As ground truth, we use the `QUIJOTE-MG` suite [5], a collection of large cosmological N-body simulations including extensions beyond ΛCDM. These runs evolve 512^3 CDM particles in a periodic box of $1000\ \mathrm{Mpc}\ h^{-1}$ from redshift $z = 127$ to $z = 0$. In this context, each particle's trajectory is described by its displacement, defined as $\Delta = \mathbf{x}_f - \mathbf{x}_i$, where $\mathbf{x}_i$ are the initial grid positions and $\mathbf{x}_f$ the final positions at the present epoch. With thousands of timesteps, N-body simulations provide accurate non-linear predictions but are computationally prohibitive for systematic parameter exploration.

For fast approximations, we generate simulations with the publicly available `MG-PICOLA` code [22], which achieves efficiency by evolving large-scale modes with second-order Lagrangian perturbation theory (2LPT) while treating smaller scales with N-body dynamics [8]. This reduces runtime by more than an order of magnitude, but accuracy deteriorates in the non-linear regime - precisely where MG effects are most relevant.

2.2 Training and Evaluation:

`MG-NECOLA` is a CNN designed to refine `MG-PICOLA` simulations toward N-body accuracy. The network takes as input the `MG-PICOLA` displacement field, $\Delta_{\mathrm{MG\text{-}COLA}}$, and is trained to predict the residual relative to the high-fidelity `QUIJOTE-MG` reference: $\Delta_{\mathrm{res}} = \Delta_{\mathrm{QUIJOTE\text{-}MG}} - \Delta_{\mathrm{MG\text{-}COLA}}$. The corrected displacement field is then obtained as $\Delta_{\mathrm{MG\text{-}COLA}} + \Delta_{\mathrm{res}}$, providing an enhanced approximation that closely reproduces the N-body result at a fraction of the computational cost.

Each `MG-PICOLA` run is initialized with the same random seeds and cosmological parameters as its `QUIJOTE-MG` counterpart, ensuring particle-by-particle correspondence between the two simulations. The dataset consists of 100 paired simulations with the modified gravity parameter fixed at $f_{R_0} = 5 \times 10^{-7}$, while the cosmological parameters are held constant at $(\Omega_m, \Omega_b, h, n_s, \sigma_8) = (0.3175, 0.049, 0.6711, 0.9624, 0.8340)$.

Performance is evaluated against three references: (1) `MG-PICOLA`, our fast-simulation benchmark; (2) `QUIJOTE-MG`, the high-fidelity N-body target; and (3) `MG-NECOLA`, our CNN-enhanced predictions.

Assessment is carried out at both the particle level and through cosmological summary statistics, with particular emphasis on the matter power spectrum and bispectrum at non-linear scales.

2.3 Network Architecture:

Our network design follows the V-Net–based convolutional architecture introduced in the `NECOLA` framework [15, 16], originally developed for ΛCDM simulations. While the core structure is similar, this work makes two key advances: (i) it extends the framework to modified gravity cosmologies, where non-linear effects are stronger and more scale-dependent, and (ii) it introduces an adapted loss function that balances accuracy in both Lagrangian and Eulerian displacement fields, along with a gradient-matching term to capture small-scale features. Together, these modifications enable `MG-NECOLA` to recover the characteristic signatures of MG models with high fidelity, going beyond the original ΛCDM application. The novelty thus lies not in the architectural form but in the integration of MG-specific priors into the learning objective and its application to non-standard gravity simulations.

The V-Net-based architecture [4, 7, 13, 18] follows a symmetric encoder–decoder ("U-shaped") design with residual connections. The model consists of two downsampling and two upsampling stages, each comprising residual blocks of two 3^3 convolutions with 1^3 shortcut paths. Batch normalization is applied after every convolution except for the first and last two, and leaky ReLU activations (slope 0.01) follow each normalization as well as the first and penultimate convolutions. The hidden layers use 64 channels, the input and output layers have three channels (corresponding to displacement components), and the concatenated layers use 128 channels.

Residual learning is applied by adding the input displacement field, $\Delta_{\text{MG-COLA}}$, directly to the network output, so that the model learns only the correction toward the target $\Delta_{\text{QUIJOTE-MG}}$. Downsampling and upsampling are implemented using 3^3 convolutions and 2^3 transposed convolutions with stride 2. All 3^3 convolutions are unpadded to preserve translational equivariance; encoder features are therefore cropped before concatenation with decoder layers. The training loss combines three contributions:

$$\mathcal{L} \equiv w_{\text{lag}} \, \mathcal{L}_{\text{lag}} + w_{\text{eul}} \, \mathcal{L}_{\text{eul}} + w_{\text{grad}} \, \mathcal{L}_{\text{grad}}, \tag{1}$$

where $\mathcal{L}_{\text{lag}}$ measures the root-mean-squared error (RMSE) on the displacement field, $\mathcal{L}_{\text{eul}}$ computes the RMSE on the overdensity field, and $\mathcal{L}_{\text{grad}}$ evaluates the RMSE on the displacement field gradient. This gradient term penalizes both overly smooth and excessively abrupt displacement predictions, ensuring better recovery of small-scale structure particularly important in MG scenarios. We set $(w_{\text{lag}}, w_{\text{eul}}, w_{\text{grad}}) = (1.0, 3.0, 2.0)$ so that the three terms contribute comparably in the validation set. This loss formulation represents a significant departure from the `NECOLA` approach, which used $\mathcal{L}_{\text{NECOLA}} \equiv \log(\mathcal{L}_{\text{eul}} \mathcal{L}_{\text{lag}}^{\lambda})$ with hyperparameter λ, focusing primarily on particle number density and displacements in ΛCDM. Our three-term loss reflects the need to capture MG-specific small-scale signatures with higher accuracy, as demonstrated in our ablation studies (see bottom left panel of Fig. 1).

Given the large data size (3×512^3), inputs are split into subcubes of 3×128^3 (250 Mpc h^{-1} per side). To preserve translational equivariance, no padding is applied in the 3^3 convolutions; instead, inputs are periodically padded with 20 voxels per side, yielding effective inputs of 3×164^3. Data augmentation includes random rotations and parity transformations, ensuring equivariance of the displacement field.

Training uses Adam optimizer, [17], with a learning rate 10^{-4}, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and a learning rate halving after three stagnant validation epochs. The model is trained for 100 epochs using 70% of the simulations, with 20% reserved for validation and 10% for testing.

3 Results

Figure 1 summarizes the performance of `MG-NECOLA` on the test set of 10 realizations, evaluated with standard large-scale structure statistics. The top left and right panels show the matter power spectrum, $P(k)$, and the bispectrum, $B(k)$. The power spectrum quantifies the two-point clustering of matter fluctuations as a function of scale, while the bispectrum captures three-point correlations in Fourier space and provides sensitivity to non-linear effects. For the bispectrum, we present the configuration with $k_1 = 0.15 \, h \, \text{Mpc}^{-1}$ and $k_2 = 0.25 \, h \, \text{Mpc}^{-1}$ as a function of the angle θ between k_1 and k_2.
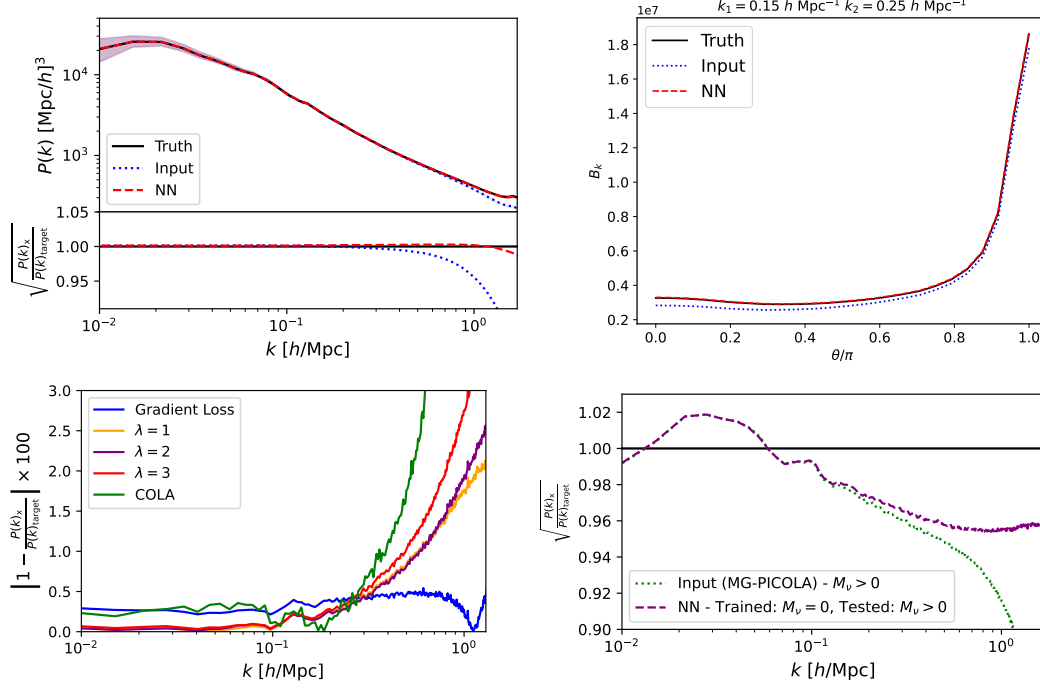
Figure 1: Performance of `MG-NECOLA` (NN) compared to `MG-PICOLA` (input benchmark) and `QUIJOTE-MG` (truth). Top left: power spectrum (top) and transfer function (bottom). Top right: bispectrum. Bottom left: ablation study comparing our loss function (Eq. (1), blue) with `NECOLA` loss variants, demonstrating improved accuracy on non-linear scales. Bottom right: transfer function for the generalization to massive neutrinos after training on $M_\nu = 0$. `MG-NECOLA` (red dashed) substantially improves upon `MG-PICOLA` (blue dotted), recovering results in much closer agreement with the `QUIJOTE-MG` reference (black solid).

The lower panels show the transfer function, $T(k)$, defined as the square root of the ratio between the predicted and target power spectra, and the fractional percentage accuracy, $\%\mathrm{Acc}$, for one test realization. Values of $T(k) = 1$ and $\%\mathrm{Acc} = 0$ indicate perfect recovery of the N-body clustering amplitude.

Across both the power spectrum and bispectrum, `MG-NECOLA` (red dashed) provides substantially improved agreement with the `QUIJOTE-MG` N-body reference (black) compared to `MG-PICOLA` (blue dotted), consistently reducing discrepancies from large to deeply non-linear small scales. The transfer function further demonstrates that `MG-NECOLA` effectively corrects the scale-dependent amplitude biases present in `MG-PICOLA`, yielding predictions that match the N-body results to better than 1% accuracy across the full range of probed modes. Furthermore, the bottom left panel of Fig. 1 shows a small ablation experiment, where it is patented that removing the gradient-matching term ($\mathcal{L}_{\mathrm{grad}}$) reverting to the original `NECOLA` loss significantly degrades performance on non-linear scales, particularly in the $k > 0.3\,h\,\mathrm{Mpc}^{-1}$ regime where MG effects become pronounced. This confirms that our proposed loss formulation, with its explicit emphasis on small-scale structure through gradient matching, provides tangible benefits for capturing MG signatures. These results confirm that `MG-NECOLA` can successfully enhance fast `MG-PICOLA` simulations, recovering the accuracy of `QUIJOTE-MG` N-body runs, particularly at non-linear scales where modified gravity signatures are most significant.

To further assess robustness, we evaluated `MG-NECOLA` on $f(R)$ cosmologies with massive neutrinos, despite training only on massless cases. Figure 1 (bottom right) shows the transfer function for $f_{R_0} = -5.42 \times 10^{-5}$ with $M_\nu = 0.6152\,\mathrm{eV}$. Remarkably, `MG-NECOLA` (purple dashed) continues to provide accurate enhancements over `MG-PICOLA` (green dotted), closely following the `QUIJOTE-MG` N-body reference across both large and small scales. Averaged across all modes, `MG-NECOLA` reduces `MG-PICOLA` mean error from 25% down to 9%, while maintaining absolute deviations below 5% even

at the most non-linear scales ($k > 0.3\,h\,\mathrm{Mpc}^{-1}$). These tests are particularly significant because they involve parameters unseen during training, demonstrating that the network has learned transferable features of non-linear structure formation, capturing the interplay of modified gravity and massive neutrinos without explicit exposure.

A typical `QUIJOTE-MG` N-body simulation requires approximately 500 CPU-hours, or 10^6 CPU seconds to complete, while a single `MG-PICOLA` run takes only 10^3 CPU seconds. Once trained, `MG-NECOLA` adds just a lightweight enhancement stage, processing each realization in approximately 180 seconds on a single GPU. Thus, `MG-NECOLA` achieves orders-of-magnitude speed-up over N-body runs while retaining near-`QUIJOTE-MG` accuracy, enabling large ensembles of high-fidelity MG simulations for surveys pipelines.

4 Conclusions

In this work, we introduced `MG-NECOLA`, a convolutional neural network that upgrades fast `MG-PICOLA` simulations to the accuracy of `QUIJOTE-MG` N-body runs with significantly reduced computational cost. Using standard large-scale structure statistics, we showed that `MG-NECOLA` reproduces N-body results with better than 1% accuracy down to non-linear scales of $k \sim 1\,h\,\mathrm{Mpc}^{-1}$, significantly outperforming `MG-PICOLA`. Moreover, although trained only on a single f_{R_0} value with massless neutrinos, the network generalizes robustly to cosmologies with massive neutrinos and varying cosmological parameters, demonstrating strong transferability to unseen physics.

Looking ahead, the next step is to extend `MG-NECOLA` to a broader family of non-standard cosmological models. In particular, we aim to incorporate diverse values of f_{R_0} together with varying neutrino masses, ultimately building a framework capable of mapping across the space of modified gravity and beyond-ΛCDM scenarios. Such a tool would allow systematic exploration of cosmological signatures across a wide range of theories, providing an essential resource for interpreting data from the next generation of large-scale structure surveys.

References

[1] Elcio Abdalla et al. Cosmology intertwined: A review of the particle physics, astrophysics, and cosmology associated with the cosmological tensions and anomalies. *JHEAp*, 34:49–211, 2022. doi: 10.1016/j.jheap.2022.04.002.

[2] M. Abdul Karim et al. DESI DR2 Results II: Measurements of Baryon Acoustic Oscillations and Cosmological Constraints. 3 2025.

[3] Yashar Akrami et al. *Modified Gravity and Cosmology. An Update by the CANTATA Network.* Springer, 2021. ISBN 978-3-030-83714-3, 978-3-030-83717-4, 978-3-030-83715-0. doi: 10.1007/978-3-030-83715-0.

[4] Renan Alves de Oliveira, Yin Li, Francisco Villaescusa-Navarro, Shirley Ho, and David N. Spergel. Fast and Accurate Non-Linear Predictions of Universes with Deep Learning. *arXiv e-prints*, art. arXiv:2012.00240, November 2020.

[5] Marco Baldi and Francisco Villaescusa-Navarro. Quijote-mg simulations. In preparation, 2025.

[6] Spyros Basilakos, Savvas Nesseris, and Leandros Perivolaropoulos. Observational constraints on viable f(R) parametrizations with geometrical and dynamical probes. *Phys. Rev. D*, 87(12): 123529, 2013. doi: 10.1103/PhysRevD.87.123529.

[7] Philippe Berger and George Stein. A volumetric deep Convolutional Neural Network for simulation of mock dark matter halo catalogues. *Mon. Not. Roy. Astron. Soc.*, 482(3):2861–2871, 2019. doi: 10.1093/mnras/sty2949.

[8] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro. Large-scale structure of the universe and cosmological perturbation theory. *Physics Reports*, 367(1-3):1–248, Sep 2002. ISSN 0370-1573. doi: 10.1016/s0370-1573(02)00135-7. URL http://dx.doi.org/10.1016/S0370-1573(02)00135-7.

[9] Matteo Cataneo, David Rapetti, Fabian Schmidt, Adam B. Mantz, Steven W. Allen, Douglas E. Applegate, Patrick L. Kelly, Anja von der Linden, and R. Glenn Morris. New constraints on $f(R)$ gravity from clusters of galaxies. *Phys. Rev. D*, 92(4):044009, 2015. doi: 10.1103/PhysRevD.92.044009.

[10] Timothy Clifton, Pedro G. Ferreira, Antonio Padilla, and Constantinos Skordis. Modified Gravity and Cosmology. *Phys. Rept.*, 513:1–189, 2012. doi: 10.1016/j.physrep.2012.01.001.

[11] Antonio De Felice and Shinji Tsujikawa. f(R) theories. *Living Rev. Rel.*, 13:3, 2010. doi: 10.12942/lrr-2010-3.

[12] Eleonora Di Valentino et al. Cosmology Intertwined III: $f\sigma_8$ and S_8. *Astropart. Phys.*, 131: 102604, 2021. doi: 10.1016/j.astropartphys.2021.102604.

[13] 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. doi: 10.1073/pnas.1821458116.

[14] Wayne Hu and Ignacy Sawicki. A Parameterized Post-Friedmann Framework for Modified Gravity. *Phys. Rev. D*, 76:104043, 2007. doi: 10.1103/PhysRevD.76.104043.

[15] Neerav Kaushal, Francisco Villaescusa-Navarro, Elena Giusarma, Yin Li, Conner Hawry, and Mauricio Reyes. Necola: Toward a universal field-level cosmological emulator. In *35th conference on Neural Information Processing Systems*, 12 2021. URL `https://ml4physicalsciences.github.io/2021/`.

[16] Neerav Kaushal, Francisco Villaescusa-Navarro, Elena Giusarma, Yin Li, Conner Hawry, and Mauricio Reyes. NECOLA: Toward a universal field-level cosmological emulator. *The Astrophysical Journal*, 930(2):115, May 2022. doi: 10.3847/1538-4357/ac5c4a. URL `https://doi.org/10.3847/1538-4357/ac5c4a`.

[17] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017.

[18] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation, 2016.

[19] J. Bayron Orjuela-Quintana and Savvas Nesseris. Tracking the validity of the quasi-static and sub-horizon approximations in modified gravity. *JCAP*, 08:019, 2023. doi: 10.1088/1475-7516/2023/08/019.

[20] Svetlin Tassev, Matias Zaldarriaga, and Daniel J Eisenstein. Solving large scale structure in ten easy steps with COLA. *Journal of Cosmology and Astroparticle Physics*, 2013(06):036–036, jun 2013. doi: 10.1088/1475-7516/2013/06/036. URL `https://doi.org/10.1088`.

[21] Francisco Villaescusa-Navarro, ChangHoon Hahn, Elena Massara, Arka Banerjee, Ana Maria Delgado, et al. The quijote simulations. *The Astrophysical Journal Supplement Series*, 250 (1):2, August 2020. doi: 10.3847/1538-4365/ab9d82. URL `https://doi.org/10.3847/1538-4365/ab9d82`.

[22] Bill S. Wright, Hans A. Winther, and Kazuya Koyama. COLA with massive neutrinos. 2017 (10):054, October 2017. doi: 10.1088/1475-7516/2017/10/054.