



HAL
open science

Super-resolution emulator of cosmological simulations using deep physical models

Doogesh Kodi Ramanah, Tom Charnock, Francisco Villaescusa-Navarro,
Benjamin D. Wandelt

► **To cite this version:**

Doogesh Kodi Ramanah, Tom Charnock, Francisco Villaescusa-Navarro, Benjamin D. Wandelt. Super-resolution emulator of cosmological simulations using deep physical models. *Monthly Notices of the Royal Astronomical Society*, 2020, 495, pp.4227. 10.1093/mnras/staa1428 . hal-02475275

HAL Id: hal-02475275

<https://hal.science/hal-02475275>

Submitted on 28 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Super-resolution emulator of cosmological simulations using deep physical models

Doogesh Kodi Ramanah^{1,2,3★}, Tom Charnock^{1,2}, Francisco Villaescusa-Navarro^{4,5}
and Benjamin D. Wandelt^{2,3,5}

¹*DARK, Niels Bohr Institute, University of Copenhagen, Lyngbyvej 2, DK-2100 Copenhagen, Denmark*

²*CNRS, UMR 7095, Institut d'Astrophysique de Paris, Sorbonne Université, 98 bis bd Arago, F-75014 Paris, France*

³*Institut Lagrange de Paris (ILP), Sorbonne Université, 98 bis bd Arago, F-75014 Paris, France*

⁴*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544-0010, USA*

⁵*Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA*

Accepted 2020 May 19. Received 2020 April 29; in original form 2020 January 24

ABSTRACT

We present an extension of our recently developed Wasserstein optimized model to emulate accurate high-resolution (HR) features from computationally cheaper low-resolution (LR) cosmological simulations. Our deep physical modelling technique relies on restricted neural networks to perform a mapping of the distribution of the LR cosmic density field to the space of the HR small-scale structures. We constrain our network using a single triplet of HR initial conditions and the corresponding LR and HR evolved dark matter simulations from the QUIJOTE suite of simulations. We exploit the information content of the HR initial conditions as a well-constructed prior distribution from which the network emulates the small-scale structures. Once fitted, our physical model yields emulated HR simulations at low computational cost, while also providing some insights about how the large-scale modes affect the small-scale structure in real space.

Key words: methods: numerical – methods: statistical – dark matter – large-scale structure of Universe.

1 INTRODUCTION

One of the most powerful ways to learn about the cosmological model of our Universe relies on extracting the physical information encoded in its large-scale structure. However, the highly non-linear dynamics involved in gravitational structure formation renders this a complex problem, such that numerical simulations are a prerequisite to obtain theoretical predictions in the fully non-linear regime of structure growth. Cosmological inference from next-generation galaxy surveys such as *Euclid* (Laureijs et al. 2011) and the Large Synoptic Survey Telescope (LSST; Ivezić et al. 2008) will require a large number of detailed and high-resolution (HR) simulations to generate mock observations of the Universe. Fast and reliable emulators of these complex dynamics would, therefore, be crucial for data analysis and light-cone production for such survey missions.

By virtue of their versatility and effectiveness, deep generative models have recently been developed for a range of applications in cosmology, especially with recent advances in the field of deep learning. He et al. (2019) constructed a deep neural network, based

on the U-Net (Ronneberger, Fischer & Brox 2015) architecture, to predict the non-linear cosmic structure formation from the linear perturbation theory. Giusarma et al. (2019) employed a similar architecture to map standard N -body simulations to non-standard ones with massive neutrinos. Zhang et al. (2019) developed a two-phase convolutional neural network to predict the galaxy distribution in hydrodynamic simulations from their corresponding 3D dark matter density fields. Various approaches based on 3D convolutional neural networks have been developed to generate mock halo catalogues from the cosmological initial conditions (Berger & Stein 2019; Bernardini et al. 2019). The deep convolutional generative adversarial network architecture (Goodfellow et al. 2014; Radford, Metz & Chintala 2015) and its variants (Arjovsky, Chintala & Bottou 2017) have been used to generate cosmological weak lensing convergence maps with high statistical confidence (Mustafa et al. 2019; Tamosiunas et al. 2020), realistic 2D and 3D realizations of the cosmic web (Rodríguez et al. 2018; Perraudin et al. 2019; Feder, Berger & Stein 2020), 3D gas pressure distributions from dark matter fields (Tröster et al. 2019), and 3D cosmic neutral hydrogen (HI) distributions (Zamudio-Fernandez et al. 2019). Recent deep U-Net models were also successful in generating 3D gas density distributions with dark matter annihilation feedback (List, Bhat & Lewis 2019).

* E-mail: ramanah@nbi.ku.dk

In this work, we present an extension of our recent Wasserstein optimized network, developed for mapping approximately evolved dark matter fields to their corresponding 3D halo count distributions obtained from N -body simulations (Kodi Ramanah, Charnock & Lavaux 2019a). Here, we accurately emulate super-resolution N -body simulations using their respective initial conditions and evolved low-resolution (LR) structures. We find, as with the previous work, that using a well motivated prior and simple physical principles results in high computational efficiency and remarkable performance while requiring only one set of simulations for fitting.

The remainder of this paper is structured as follows. In Section 2, we outline the salient features of the QUIJOTE N -body simulation suite employed in the fitting and validation of our neural network, which is described in Section 3. We briefly review the conceptual foundations of the Wasserstein generative networks in Section 3.1, followed by a description of the network architecture and optimization procedure in Sections 3.2 and 3.3, respectively. We subsequently evaluate the performance of our HR N -body emulator in terms of its capacity to reproduce the complex structures of the cosmic web from a ground truth simulation using various metrics in Section 4. Finally, in Section 5, we provide a summary of the key aspects of our work, possible avenues for future investigations and potential applications where employing such an emulator may yield crucial advantages.

2 QUIJOTE SIMULATIONS

The QUIJOTE suite of simulations (Villaescusa-Navarro et al. 2019) is a set of 43 100 full N -body simulations encompassing over 7000 different cosmological models, specifically designed to provide high-quality data to train machine learning algorithms. Here, we use only one HR simulation, one LR simulation generated with the same initial conditions and those initial conditions as our data for fitting the model. We use a larger set of these triplets from the remainder of the suite to validate and test the model.

The simulations cover a cosmological volume of $1000 h^{-1}$ Mpc and are initialized at a redshift $z = 127$ and evolved to present time using the TreePM code GADGET3 – an improved version of GADGET2 (Springel 2005). The gravitational softening length is set to $1/40$ of the mean interparticle separation, corresponding to 100 and $50 h^{-1}$ kpc, for the LR and HR simulations with 256^3 and 512^3 particles tracing dark matter, respectively. The input matter power spectrum and transfer functions are obtained by rescaling their present-day ($z = 0$) counterparts from CAMB (Lewis, Challinor & Lasenby 2000). The simulations assume a Planck-like Lambda cold dark matter (Λ CDM) cosmology (Planck Collaboration VI 2018) with $\Omega_m = 0.3175$, $\Omega_b = 0.049$, $h = 0.6711$, $\sigma_8 = 0.834$, $n_s = 0.9624$, and $w = -1$. A comprehensive description of the QUIJOTE simulations is provided in Villaescusa-Navarro et al. (2019).

3 DEEP GENERATIVE MODELLING

3.1 Wasserstein generative network

The Wasserstein generative model is an improved variant of generative adversarial networks (GANs; Goodfellow et al. 2014). GANs cast the generative process as a competition between a generator \mathcal{G}_θ and a discriminator \mathcal{D}_φ , where θ and φ correspond to the weights of the generator and discriminator, respectively. The role of the former is to produce some artificial data given a vector of random noise, while the latter must differentiate between the generated samples and the *real* ones. The Wasserstein variant (Arjovsky & Bottou

2017) works similarly with a slight change in the loss function, but with a large change in interpretation of *training*. Instead of training a discriminative classifier, one approximates the optimal transport distance which takes the samples from the generated distribution $\mathbf{g} = \mathcal{G}_\theta(\mathbf{z})$ with $\mathbf{z} \sim \mathbb{P}_z$ to the target distribution \mathbb{P}_r using a critic network \mathcal{C} . This optimal transport distance is known as the Wasserstein-1 distance or Earth mover’s distance. A more in-depth description of the conceptual framework underlying the Wasserstein generative network is provided in our previous work (Kodi Ramanah et al. 2019a, hereafter KCL19).

Computing the Wasserstein-1 distance, in practice, is intractable, but an approximation is feasible for the case where the critic network \mathcal{C}_φ is a 1-Lipschitz function parameterized by weights φ that lie in a compact space. To enforce this constraint, one can introduce a gradient penalty to the loss function, as proposed by Gulrajani et al. (2017), which provides a substantial improvement over previous weight-clipping methods. The augmented loss function for the Wasserstein generative model may be expressed as

$$\mathcal{L} = \mathbb{E}_{\mathbf{z} \sim \mathbb{P}_z} [\mathcal{C}_\varphi(\mathcal{G}_\theta(\mathbf{z}))] - \mathbb{E}_{\mathbf{x} \sim \mathbb{P}_r} [\mathcal{C}_\varphi(\mathbf{x})] + \lambda \mathbb{E}_{\hat{\mathbf{x}} \sim \mathbb{P}_{\hat{\mathbf{x}}}} \left[\left(\|\nabla_{\hat{\mathbf{x}}} \mathcal{C}_\varphi(\hat{\mathbf{x}})\|_2 - 1 \right)^2 \right], \quad (1)$$

where the vector $\mathbf{z} \sim \mathbb{P}_z$ is a sample of the set of LR evolved dark matter density field and corresponding HR initial conditions, which emulates a HR evolved dark matter density simulation $\mathbf{g} = \mathcal{G}_\theta(\mathbf{z})$. $\mathbf{x} \sim \mathbb{P}_r$ is a sample from the true distribution of HR evolved dark matter density fields. The first two terms yield the approximate Wasserstein distance, while the last term ensures that the gradient of the critic network remains close to unity. $\hat{\mathbf{x}} \sim \mathbb{P}_{\hat{\mathbf{x}}}$ is obtained by interpolating between the real and generated samples: $\hat{\mathbf{x}} = \varepsilon \mathbf{x} + (1 - \varepsilon)\mathbf{g}$ for ε sampled randomly and uniformly, $\varepsilon \in [0, 1]$, and λ is an arbitrary penalty coefficient.

3.2 Neural network architecture

A schematic representation of our super-resolution emulator and the critic is illustrated in Fig. 1. As with our halo painting network in KLC19, we consider the *generator* as an *emulator*, i.e. the input to the network is not a flat array of noise, as usually considered for generative models, but rather involves 3D boxes of LR density field and corresponding HR initial conditions. As such, the super-resolution emulator performs a physical mapping from the distribution of LR density field to HR structures, conditional on the initial conditions. The critic compares the difference between a scalar summary of the features present in the real and emulated HR fields. By minimizing the difference between these scalar summaries, we fit the function that transports the set of features in the emulated field to the real simulation. This is the approximate Wasserstein distance. Once the parameters of this critic have converged, it can be used as the metric (loss function) that is minimized to fit our super-resolution N -body emulator.

The choice of architecture for our HR emulator and the critic is depicted in Fig. 1. We employ residual Inception blocks (cf. fig. 4 in KLC19 for a schematic illustration), as outlined below, in designing the emulator, with the additional inclusion of 3D convolutions with kernel size of $7 \times 7 \times 7$. All the convolutional layers, including those in the Inception blocks, have six filter channels, except for the output layer, which has a single filter ReLU activation. The leaky ReLU activation function with a leaky parameter of $\alpha = 0.1$ encodes the required approximate non-linearity in the initial layers. We used the same critic architecture as in KLC19, except

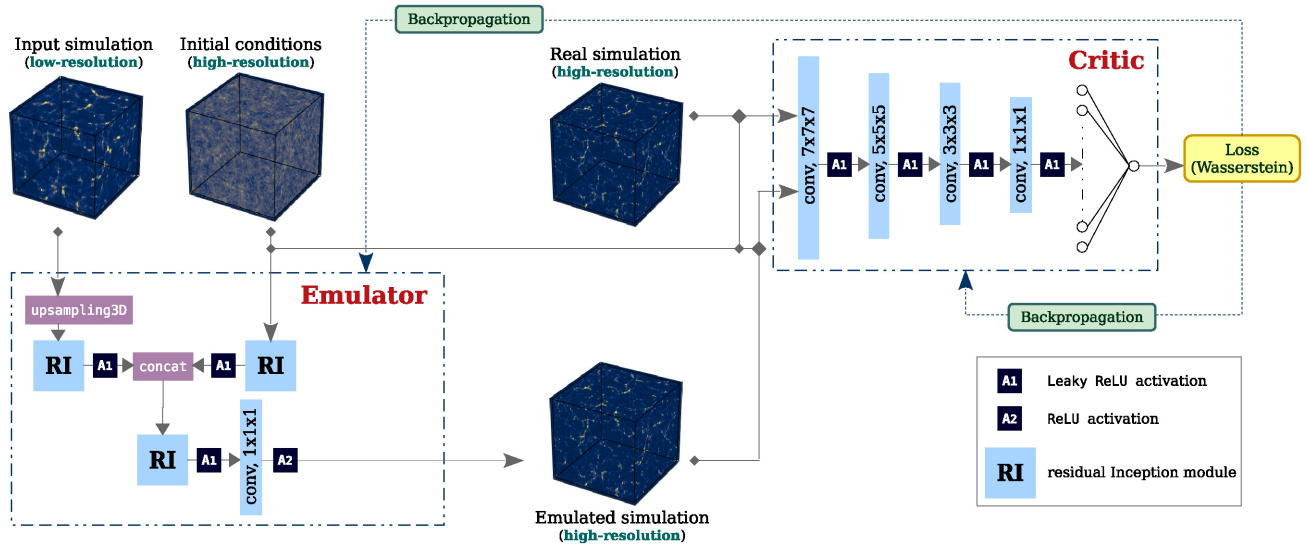


Figure 1. Schematic representation of the super-resolution emulator implemented in this work. The emulator approximates the underlying mapping of the distribution of LR density field to HR structures, with the input initial conditions providing an informative prior distribution from which the emulator constructs the fine structures, to yield a super-resolution field. The difference between the output of the critic for the real and emulated density fields, conditional on the initial conditions, is the approximate Wasserstein distance, which is minimized to fit the super-resolution N -body emulator.

that the initial conditions are concatenated with the input real HR simulation. The critic contains four 3D convolutional layers with leaky ReLU activation, with an initial kernel of $7 \times 7 \times 7$, and gradually reducing the kernel size to $1 \times 1 \times 1$, with the subsequent output flattened and fed to a fully connected layer with linear activation.

The residual Inception (Szegedy et al. 2017) module combines the Inception (Szegedy et al. 2015, 2016) architecture, which performs parallel convolutions with kernels of distinct sizes, and residual connections (He et al. 2016) between the input to the module to its subsequent output, thereby facilitating the flow of small-scale information. This residual Inception block allows feature extraction on different scales of the density field simultaneously and significantly improves computational efficiency.

The super-resolution emulator is designed to perform a physical mapping that encodes some fundamental symmetries as motivated by basic physical principles and the cosmological principle. In particular, we enforce translational invariance via the use of 3D convolutions and the random flip of the inputs during training, and rotational invariance by performing a random rotation of the inputs while optimizing the network. The (non-) locality of the super-resolving procedure is inbuilt in the largest convolutional kernel in the Inception module, whose size is motivated by causal transport arguments outlined in the next paragraph. Finally, the non-linear activations provide the adequate non-linearity as required by this physical mapping.

The typical displacement field, derived from the initial conditions, is $\sim 5 h^{-1}$ Mpc in a Λ CDM Universe with Planck cosmological parameters (Planck Collaboration VI 2018), with an upper limit of $\sim 20 h^{-1}$ Mpc for the fastest moving objects, as substantiated in Lavaux, Jasche & Leclercq (2019). The $7 \times 7 \times 7$ kernel in the first residual Inception block has a receptive field of $\sim 24 h^{-1}$ Mpc and ensures that the network is capable of encapsulating the causal displacements between the HR initial conditions and evolved density field. This is necessary to transport the initial conditions to the correct regions of the emulated field. This transportation is provided by the second Inception module, after concatenating the subsequent feature maps. The effective receptive patch of nearly

$50 h^{-1}$ Mpc is sufficiently large to pull all the relevant information from the respective inputs. The largest scale modes of the emulated density field are provided by the evolved LR simulation. Although the network is never fit with the large modes explicitly (with extremely large convolutional kernels), they are incorporated by fitting with small sub-patches of the LR field that includes these modes on average. In Section 4.6, we examine a selection of kernels and feature maps in an attempt to introspect our super-resolution simulation emulator.

3.3 Fitting methodology

We make use of a single set of three 3D gridded dark matter fields for inferring the parameters of the emulator, corresponding to the HR initial conditions, and LR and HR density fields evolved from these initial conditions (at $z = 0$). The LR and HR simulation boxes have 256^3 and 512^3 voxels, respectively, and they both have a size of $1000 h^{-1}$ Mpc. We utilize most of the boxes for fitting, with a independent portion kept for validation. The optimization rationale entails minimizing the approximate Wasserstein distance between the true and emulated HR density fields, conditional on the initial conditions, such that the super-resolution simulation emulator approximates the correct mapping from LR to HR density fields.

The optimization routine proceeds as follows: a set of two randomly chosen 3D patches of HR initial conditions and the corresponding patch of LR density field, with respective sizes of 40^3 and 20^3 voxels, are passed through the first network to emulate a HR density volume, with the input patches encoding a sufficiently large number of informative features for optimizing the network. The physical voxel sizes for the LR and HR fields are ~ 4 and $\sim 2 h^{-1}$ Mpc, respectively. As such, the sub-volume involved during the optimization procedure is $\sim 80 h^{-1}$ Mpc. The convolutional layers in the residual Inception module do not utilize any padding so as not to induce any numerical artefacts arising from the boundaries of the input 3D slices. Hence, in order to eliminate the need for padding, the input size is conveniently chosen to be

larger, such that the network output has the desired box size. For our particular network architecture, as depicted in Fig. 1, since there are two Inception modules, with the largest convolutional kernel being $7 \times 7 \times 7$, the input slices must be larger by $(7 - 1) \times 2 = 12$ voxels on each side. During optimization, the emulator, therefore, physically maps a sub-volume of ~ 80 to $\sim 55 h^{-1}$ Mpc. Once the model is optimized, the inputs may be of any arbitrary box size, exploiting the translational invariance of convolutional kernels, to predict the HR density field of corresponding desired size.

To artificially augment the data manifold, the 3D volumes are randomly rotated and mirrored along the three axes on entry to the emulator. The emulated super-resolution field and the corresponding true HR density volume (rotated and mirrored in the same way as the initial conditions and the LR simulation) are then passed through the critic. The parameters of the critic are optimized to minimize the difference between the summary of the true density field and the emulated one via equation (1), which equates to finding the optimal transport of features conditional on the architecture of the critic. At this step, the parameters of the emulator are kept fixed so that the critic properly approximates the difference between the features of the emulated samples $\mathbf{g} = \mathcal{G}_\theta(\mathbf{z})$ from distribution \mathbb{P}_z and true distribution \mathbb{P}_r via the optimization through several steps until convergence. Once convergence of the critic is reached, a single weight update of the emulator is made using

$$\mathcal{L}_{\text{emulator}} = \mathbb{E}_{z \sim \mathbb{P}_z} [\mathcal{C}_\psi(\mathcal{G}_\theta(z))]. \quad (2)$$

The batch size is initialized at unity and is doubled for every subsequent one hundred thousand weight updates to make better estimates of the expectation value of the distribution once each individual emulated simulation is close to its corresponding true simulation (Smith et al. 2017). This iterative optimization routine is repeated until convergence of the two distributions.

The neural network and optimization procedure are implemented in TENSORFLOW (Abadi et al. 2016). We set the number of iterations for the critic updates to $n_{\text{critic}} = 10$ and the coefficient for the gradient penalty to $\lambda = 10$. We employ the *Adam* (Kingma & Ba 2014) optimizer, with a learning rate of $\eta = 10^{-4}$ and first- and second-moment exponential decay rates of $\beta_1 = 0.5$ and $\beta_2 = 0.999$, respectively. We optimize the emulator for $\sim 5 \times 10^5$ weight updates, requiring around 120 h on a NVIDIA Quadro P6000. As with the halo painting network from our previous work (cf. KLC19), building a physically motivated neural engine, in contrast to conventional extremely deep black box approaches, yields crucial advantages in terms of relatively low network complexity. As a consequence, much less data are needed to fit the emulator. Our HR emulator has $\mathcal{O}(10^4)$ network parameters, and is fit with a data set of only one each of the initial conditions, LR and HR density fields. This is extremely cost efficient in comparison to recent deep generative models developed in cosmology with $\mathcal{O}(10^7)$ parameters, which require several thousands of numerical simulations for training (e.g. Rodríguez et al. 2018; Giusarma et al. 2019; He et al. 2019). Physically motivated neural networks (cf. KLC19 and this work) deliver results of at least comparable accuracy, as we will show in the following.

4 RESULTS

We assess the performance of the super-resolution emulator using a series of both qualitative and quantitative diagnostics. Fig. 2 depicts the 3D visualization of the model prediction and the ground truth for a 3D slice of length $500 h^{-1}$ Mpc. The corresponding input initial conditions and LR density field are also provided for completeness.

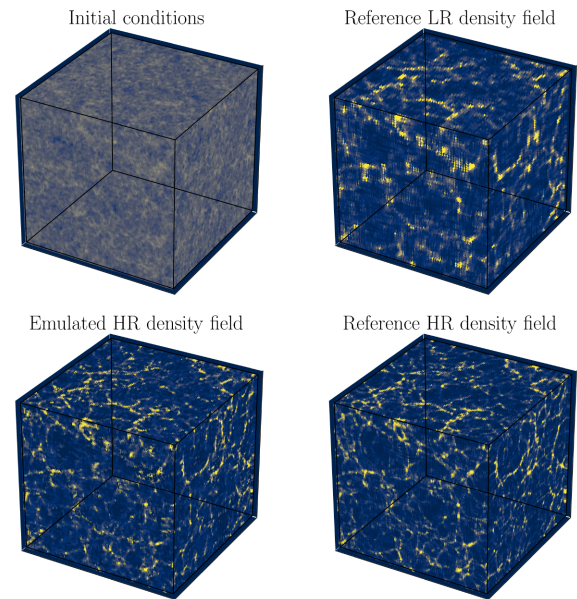


Figure 2. HR prediction (bottom left-hand panel) of the non-linearly evolved density field by our super-resolution emulator for a 3D slice of length $500 h^{-1}$ Mpc. For comparison, the reference density field from the HR simulation is depicted in the bottom right-hand panel, with the corresponding input initial conditions and LR density field illustrated in the top panels. This provides a qualitative assessment of the efficacy of the HR emulator.

Visually, there is extremely good agreement between the network prediction and reference HR simulation.

For a more quantitative evaluation, we employ four summary statistics, namely the 1D probability distribution function (PDF), power spectrum, bispectrum, and the void size function, to compare the properties of the emulated HR 3D density field to their corresponding (reference) numerical simulations. Our test set consists of 10 distinct (unseen) simulations and the following plots indicate the means and respective 1σ confidence regions. Finally, in an attempt to introspect the inner workings of the emulator, we visualize the filters of the convolutional kernels and the resulting feature maps and comment on their meaning.

4.1 1D probability distribution function

We verify whether the emulated HR simulations manage to reproduce the 1D PDF of the real simulations. The 1D PDF describes the dark matter distribution across the voxels of the grid. We first apply a smoothing to the density field with a top-hat filter on a scale of $10 h^{-1}$ Mpc and compute the 1D PDF by binning the dark matter density contrast. The resulting distributions for the emulated and reference LR and HR density fields are illustrated in Fig. 3. As can be seen, there is an almost identical agreement between the emulated simulations and the ground truth.

4.2 Power spectrum

Summary statistics are widely used in cosmology to compare model predictions with actual observations or to extract information from the latter. We first consider the two-point correlation function, $\xi(r)$, as a reliable metric to evaluate the capacity of our super-resolution emulator to encode the essential Gaussian information of the true HR density field. Its Fourier transform, the power spectrum, denoted by $P(k)$, is defined as follows:

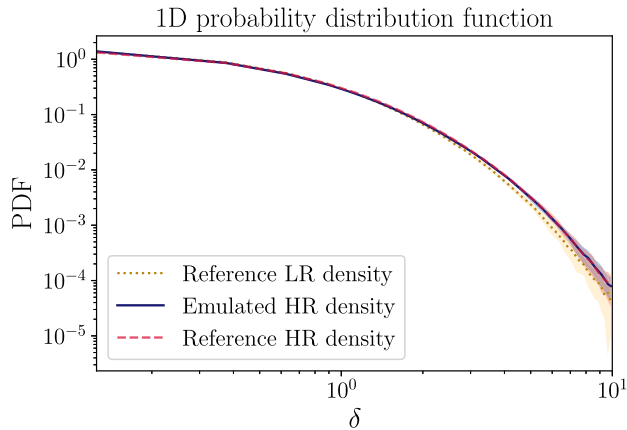


Figure 3. Normalized probability distribution function (PDF) of the dark matter density contrast. The solid blue line indicates the mean for the 10 emulated HR simulations, which matches very closely that of the reference simulations depicted by the dashed red line. The shaded regions indicate their respective 1σ regions. For completeness, the 1D PDF of the corresponding LR simulations is also shown.

$$\xi(|\mathbf{r}|) = \langle \delta(\mathbf{r}')\delta(\mathbf{r}' + \mathbf{r}) \rangle, \quad (3)$$

$$P(|\mathbf{k}|) = \int d^3\mathbf{r} \xi(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (4)$$

where k is the wavenumber of the fluctuation, i.e. $k = 2\pi/\lambda$, for given wavelength λ . The matter density distribution is typically described as a dimensionless overdensity field (or density field contrast) via $\delta(\mathbf{r}) = \rho(\mathbf{r})/\bar{\rho} - 1$, where $\rho(\mathbf{r})$ is the matter density at position \mathbf{r} and $\bar{\rho}$ is the mean density. The above two-point statistics provide a sufficient statistical description of Gaussian random fields. Since the cosmic density field on large scales, or in the initial stages of structure growth, resemble a Gaussian field, such metrics are predominantly employed in standard cosmological analyses. The transfer function, $T(k)$, defined as the square root of the ratio of the emulated power spectrum, $P_g(k)$, to the reference power spectrum, $P_{\text{ref}}(k)$,

$$T(k) \equiv \sqrt{\frac{P_g(k)}{P_{\text{ref}}(k)}}, \quad (5)$$

more adequately characterizes the agreement between the amplitudes, as a function of Fourier modes.

Fig. 4 illustrates the power spectrum and transfer function for the emulated HR simulations and corresponding reference simulations. In the top panel, the amplitude and shape of the emulated power spectra match the ground truth extremely, well implying that the super-resolution emulated density field has the correct statistical properties at the level of two-point statistics. The transfer function, depicted in the bottom panel, is close to unity up to a scale of $k \sim 1.0 h^{-1} \text{ Mpc}$, with less than 1.5 per cent deviation, further substantiating the accuracy of the emulation even in the non-linear regime of cosmic structure formation.

4.3 Bispectrum

Gravitational evolution yields a non-Gaussian component to the cosmic matter structures on the small scales. These are key probes of the nature of gravity, dark matter, and dark energy. To access the statistical information encoded in the non-Gaussian features,

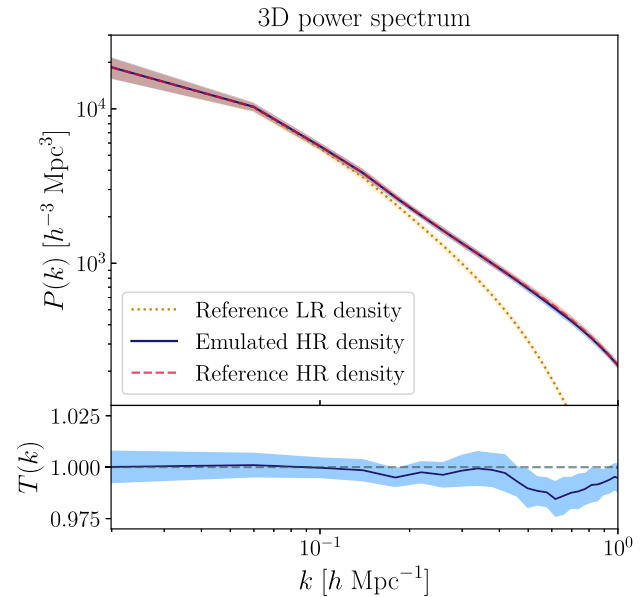


Figure 4. Summary statistics of the 3D power spectra for the emulated HR simulations and corresponding reference HR and LR realizations. The solid lines indicate the mean for 10 emulated HR realizations, with the shaded areas depicting the corresponding 1σ confidence regions. Top panel: the power spectrum diagnostics demonstrate that the emulated HR density fields reproduce extremely well the characteristic two-point statistics of the reference HR fields. Bottom panel: the transfer function better quantifies the deviation, as a function of Fourier modes, of the emulated power spectra to the ground truth. This further illustrates the fidelity of the HR simulations from our super-resolution emulator at the level of two-point summary statistics.

such as the peaks, filaments, and voids, of the cosmic web, higher order statistics (going beyond the power spectrum) are essential. We employ the bispectrum, the Fourier transform of the three-point correlation function, to quantify the spatial distribution of the cosmic structures, defined as

$$(2\pi)^3 B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)\delta_{\text{D}}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) = \langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3) \rangle, \quad (6)$$

where δ_{D} is the Dirac delta.

The bispectra reconstructed from the emulated HR density fields and the reference simulations are depicted in Fig. 5. We have chosen a particular configuration, $k_1 = 0.15 h \text{ Mpc}^{-1}$, $k_2 = 0.25 h \text{ Mpc}^{-1}$, and studied the variation with the angle between the two vectors. This choice, as justified in Giusarma et al. (2019), corresponds to the smallest scales typically used for cosmological parameter inference from measured galaxy power spectra. As can be seen by comparing the blue and red lines, the super-resolution emulator performs extremely well in reproducing the non-linear structures of the cosmic web.

4.4 Void abundance

We finally investigate the capacity of our super-resolution emulator to also replicate the void statistics in the actual HR simulations. Voids are the underdense regions of the cosmic web and encode a substantial amount of cosmological information (e.g. Lavaux & Wandelt 2012; Hamaus et al. 2016) by virtue of their ubiquity in the Universe. Two useful diagnostics, therefore, are the 3D void distribution and the void size function (VSF), describing the variation of the void abundance with void radius.

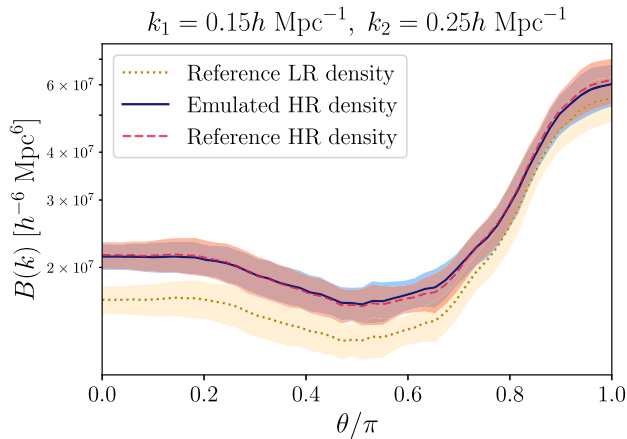


Figure 5. Summary statistics of the 3D bispectra, for a given configuration ($k_1 = 0.15 h \text{ Mpc}^{-1}$, $k_2 = 0.25 h \text{ Mpc}^{-1}$), for the emulated HR simulations and corresponding reference HR and LR realizations. The close agreement between the predictions and the ground truth implies that our super-resolution emulator reproduces with high fidelity the complex morphology, such as the filaments and voids, of the cosmic web.

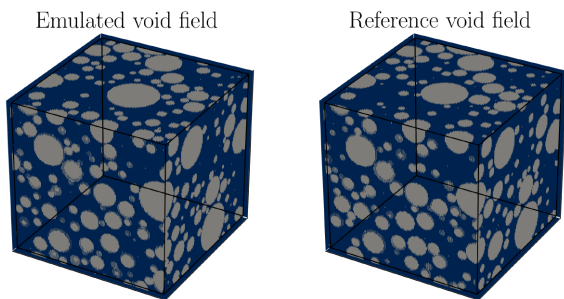


Figure 6. 3D void fields from the emulated HR simulations and the corresponding reference density field for a 3D slice of length $500 h^{-1} \text{ Mpc}$. Our super-resolution emulator adequately reproduces the properties of the void distribution. Whilst the general size and positions of the voids agree extremely well, the positions of smaller voids are occasionally slightly off.

To compute the VSF, we must identify the voids in the density fields. To that end, we adopt the void-finding algorithm devised by Banerjee & Dalal (2016), which involves smoothing the density field with a top-hat filter on a given scale R , initially chosen to be rather large, and selecting the voxels whose density is below a chosen threshold. The voxel, with the lowest density among the selected voxels, is then identified as the void centre with associated radius R . This procedure is repeated for subsequently smaller values of R , yielding a hierarchical identification of larger to smaller voids.

Fig. 6 displays the reconstructed 3D void fields from the emulated HR density field and the reference simulation in Fig. 2. From a visual inspection, we find that the network adequately reproduces the 3D distribution of voids, although the positions of smaller sized voids are slightly off in some places. This may be due to the model predictions on the smallest scales being limited by the grid resolution. The overall network performance is quantitatively verified by the respective VSF variations as shown in Fig. 7. This implies that the emulator accurately reproduces the void abundance in the actual HR simulation.

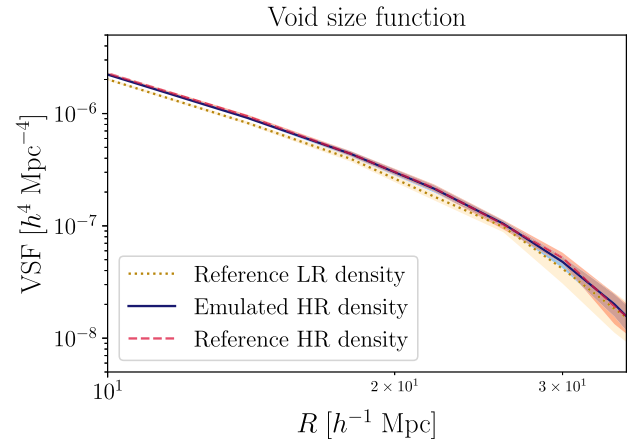


Figure 7. Abundance of voids as a function of void radius, as described by the void size function, for the emulated super-resolution simulations and corresponding reference realizations. Our emulator accurately replicates the void abundance from the actual HR simulations and, as such, fully captures the complex structures in the cosmic web.

4.5 Dependence on cosmology

We also verify the cosmological dependence of the super-resolution emulator, optimized at a given fiducial cosmology of $\Omega_m = 0.3175$, by applying it to two additional sets of 10 LR simulations of cosmologies with $\Omega_m = 0.3075$ and 0.3275 . These two variants are obtained assuming different values of Ω_m , but with the initial random phases kept the same, such that only the physical effects induced by different cosmological parameters are introduced. We find that the super-resolution emulator is fairly robust to slight variations in cosmology, in contrast to the halo painting emulator from KCL19. Fig. 8 illustrates the various summary diagnostics with respect to their corresponding ground truth. The fractional deviations of the power spectrum and bispectrum are at the level of 2 and 5 per cent, respectively, with no substantial bias induced for the 1D PDF and VSF of the emulated simulations. This slight bias may result from the emulator applying an overall local transformation to the density field. One potential solution to mitigate this issue is to pass only the small-scale modes of the HR initial conditions during the optimization procedure, such that the large-scale modes of the emulated field emanate directly from the LR density field. This would render the emulator more robust to variations in cosmology, while also improving its overall performance. We defer this potential extension of our emulator to a future investigation.

The execution time for the HR simulation (512^3 voxels) is around 500 CPU hours, while the LR one (256^3 voxels) required roughly 45 CPU hours. This implies a speed-up factor of ~ 11 when using the super-resolution emulator rather than running the HR simulation. In terms of storage, using the emulator to generate the HR simulation from the LR one on the fly provides a storage gain by a factor of 8. This gain in computation time and storage would be even more significant when mapping to higher resolutions (e.g. 1024^3 voxels), which is feasible using the super-resolution emulator. The network architecture may, however, require some tuning to optimize its performance for other resolution settings.

4.6 Visualization of filters and feature maps

In an attempt to introspect the inner-workings of the super-resolution emulator and gain some physical understanding of the

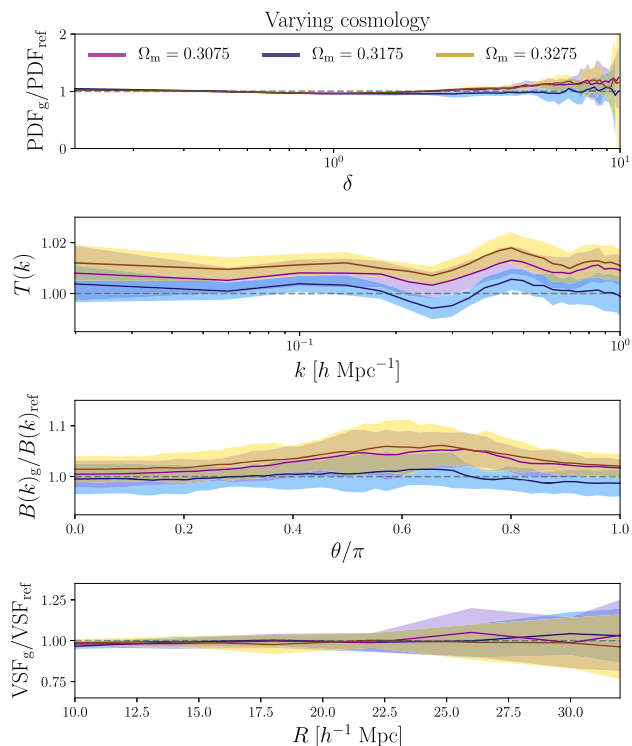


Figure 8. Fractional deviations of the model predictions for the 1D PDF, power spectrum, bispectrum, and VSF from their respective ground truth for emulated simulations at different cosmologies labelled in the top panel. The super-resolution emulator, optimized at a given fiducial cosmology, displays a slight bias at the level of the power spectrum and bispectrum when emulating simulations at varying cosmologies.

function the emulation is approximating, we visualize the filters of the convolutional kernels and their resulting feature maps. These indicate the informative features as inferred by the neural network and consequently provide some insights about the influence of the large-scale density distribution on the small-structures, although the conclusions that may be drawn from such visualizations are rather limited. We examine the filters and feature maps at the first layer of convolutions in the emulator where the majority of the information on the combination of different scales comes from.

Fig. 9 depicts the convolutional kernels and corresponding feature maps in the input layers of the residual Inception block for the LR density field in the top and bottom rows, respectively, of each panel. The distinct kernel sizes in the Inception module are displayed in the different panels. Similarly, Fig. 10 illustrates the kernels and associated feature maps for the input initial conditions. The feature maps, in essence, yield the fields that the neural network has found most informative about the input initial conditions and LR density distribution. As expected, we find that the larger kernels extract the large-scale features while the smaller sized kernels are pulling out the finer structures from the respective inputs. These feature maps are combined in non-linear fashion to provide the final emulated HR density field. As a consequence of this non-linearity, the subsequent feature maps deeper into the network contain a high level of abstraction and are not illustrated here.

The larger sized $7 \times 7 \times 7$ kernels, for both the input LR field and initial conditions, depicted in the top rows of Figs 9 and 10, respectively, appear to have a radial distribution, thereby showing that the network has learned the rotational symmetry on the larger scales to a reasonable extent, as per the underlying fitting rationale

outlined in Section 3.3, driven by the cosmological principle. This, therefore, motivates the use of the multipole expansion of kernels, as employed in Charnock et al. (2019) to use progressively stronger breaking of rotational symmetry as a principle for ordering kernel complexity, which would potentially further reduce the number of model parameters by orders of magnitude. In the case of the initial conditions, the weights are smeared over a broader distribution, highlighting the causal information encoded over the scales relevant for growth of cosmic structures. In general, the convolutional kernels try to enhance the contrast between the denser regions, such as the filaments, and the underdense ones, such as the voids.

5 CONCLUSIONS AND OUTLOOK

We have presented a super-resolution N -body emulator¹ using a deep physical modelling approach, with inspiration from our recent halo painting network (KCL19). Our emulator, once optimized, allows us to populate dark matter simulations with HR structures in a fraction of a second on a modern GPU. We showcased the performance of our method in reproducing the intricate filamentary pattern of the cosmic web from a ground truth simulation using several reliable diagnostics. It should be clearly noted that this emulator is designed to provide high-quality, deterministic approximations of the true HR simulations that are not necessarily realizations of the true distribution of the simulations as could, in the ideal case, be emulated using generative techniques. We note that we have not attempted such work here since there is no known current methods that guarantee that generated realizations are really samples from the true distribution of data, and we preferred high-quality, interpretable emulation over arbitrary generation of samples from black boxes.

The network design and architecture are driven by physical principles, in contrast to a conventional black box approach. We have focused on translational and causal information with a good approximation to rotational symmetry to finesse our network. As such, our physically motivated emulator has relatively few network parameters and may be fit sufficiently well using a single set of simulations. Recent deep generative models, as proposed in the cosmology community, rely on roughly three orders of magnitude more network parameters, which require several thousands of simulations for training. Moreover, since convolutional kernels are translationally invariant, our neural network can emulate HR simulations of arbitrary box size, for a given physical voxel size, allowing us to cheaply generate extremely large, HR simulations. The large-scale modes that are not present explicitly in the data used for the fit will still be present in the super-resolution large size box since they are provided by the LR simulation. We have yet to implement true rotational symmetry that has now been considered in Charnock et al. (2019), which could potentially lead to orders of magnitude fewer model parameters for the fit.

Such an emulator of cosmic dynamics may be employed as a means to accelerate the statistical inference framework of algorithms based on Bayesian forward modelling approaches (e.g. Jasche & Wandelt 2013; Jasche & Lavaux 2019; Porqueres et al. 2019; Kodi Ramanah et al. 2019b), to render HR analyses of upcoming galaxy surveys feasible. In this context, combining neural physical engines with forward modelling techniques is an extremely promising avenue for next-generation cosmological analyses (Charnock et al. 2019). Another practical use of such an

¹The source code repository is available at https://github.com/doogesh/super_resolution_emulator.

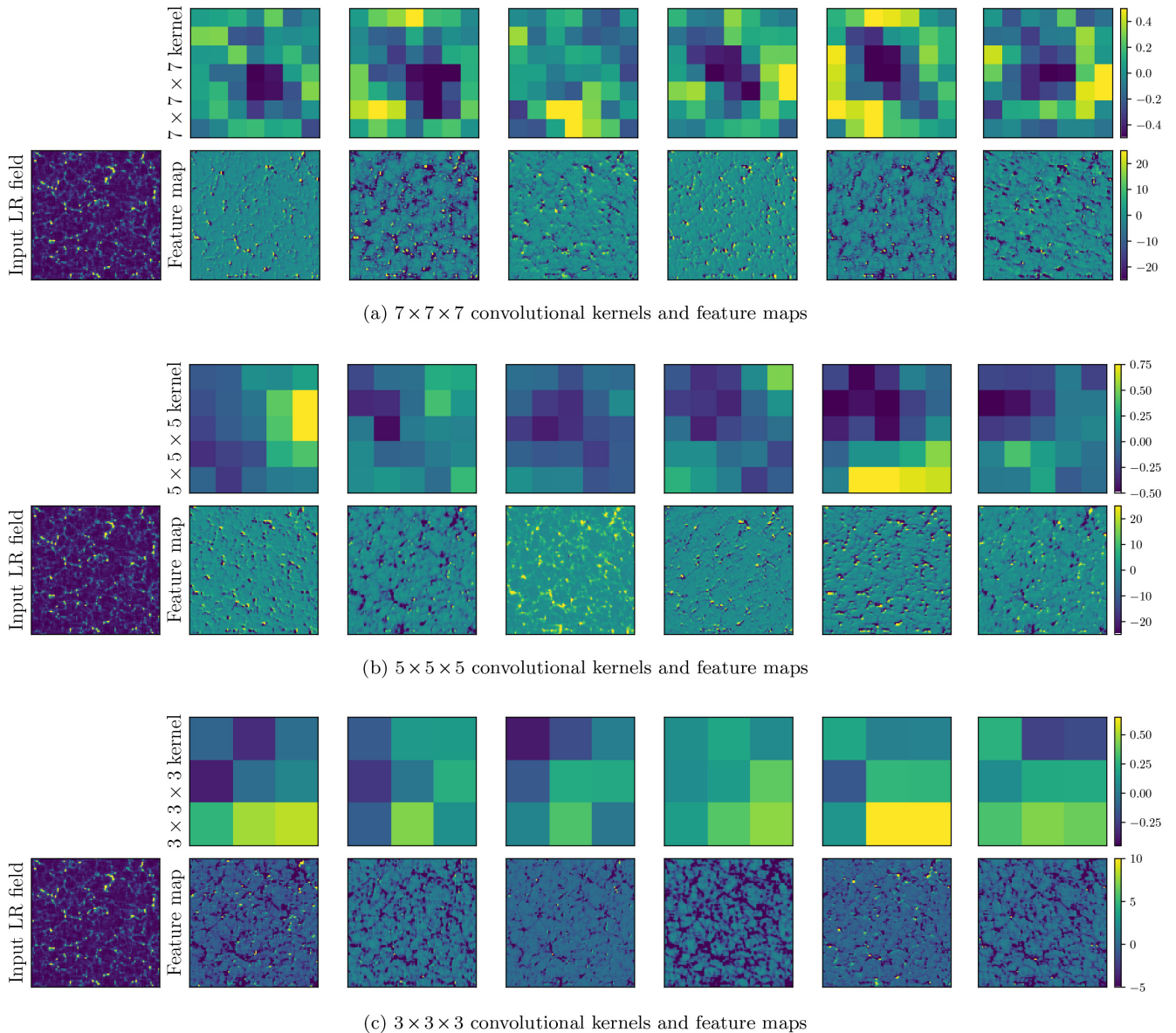


Figure 9. Convolutional kernels and corresponding feature maps in the top layer of the first residual Inception module for the LR density field (cf. Fig. 1) for kernel sizes of (panel a) $7 \times 7 \times 7$, (panel b) $5 \times 5 \times 5$, and (panel c) $3 \times 3 \times 3$, from the top to bottom, respectively. Note that the central slices of the 3D kernels and feature maps are illustrated, with the colour scale anchored for a given row. The feature maps in the bottom rows result from the convolution of the input LR field shown with the kernels in the top rows. The kernels, in particular the $7 \times 7 \times 7$ ones, have an approximately radial distribution, implying that the network has learned a certain degree of rotational symmetry, as expected from the cosmological principle. The feature maps from the larger sized kernels contain finer structures and in general encode the information about the peaks in the density field, whilst the smaller kernels try to enhance the contrast between the filaments and underdense regions.

emulator is to create accurate approximate realizations of HR simulations on the fly using stored LR simulations and generating only the HR initial conditions. For the given LRs and HRs considered in this work, the gain in execution time and storage capacity is by roughly an order of magnitude, which is significant for applications requiring several thousands of simulations.

Upcoming galaxy surveys, such as LSST and *Euclid*, will cover unprecedented volumes of the sky. While it is computationally cheap to run N -body simulations extending over such cosmological volumes, it is not feasible to run such simulations at a sufficiently HR to resolve the smallest haloes containing the galaxies expected to be observed by these surveys. Our super-resolution emulator,

therefore, constitutes the first step in addressing this computational bottleneck using the latest advances in generative modelling techniques. Achieving this ultimate goal would naturally require an emulator capable of mapping to higher resolutions than considered in this work. This would also lead to other practical applications involving the estimation of covariance matrices and light-cone production from mock observations of the Universe. The emulator may also be used to generate fast HR simulations to be fed to the recently proposed information maximizing neural network (IMNN; Charnock, Lavaux & Wand-elt 2018) to learn the optimal function that summarizes the data, with the trained IMNN subsequently being utilized in a likelihood-free inference setting. By feeding the IMNN

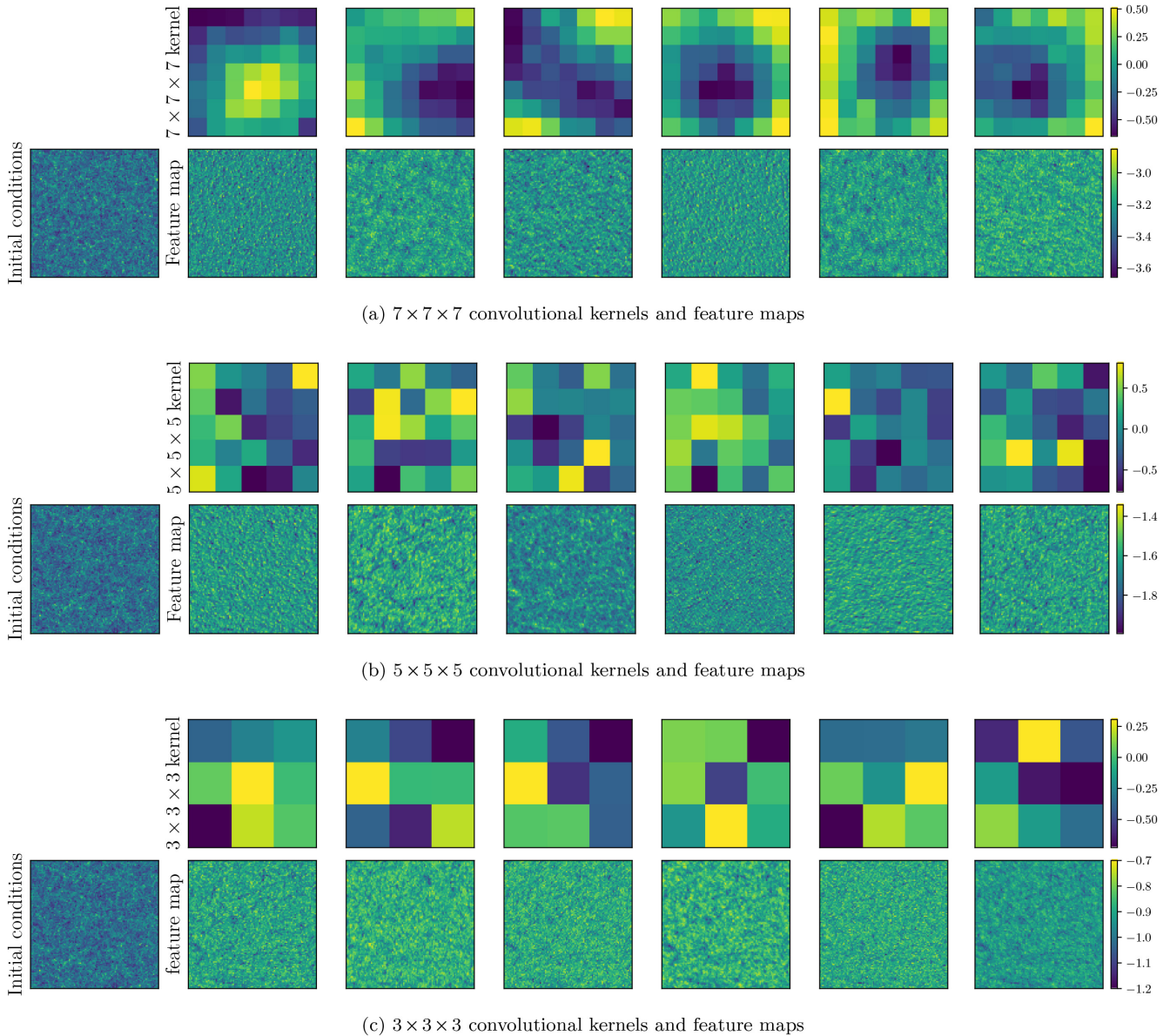


Figure 10. Convolutional kernels and corresponding feature maps in the top layer of the first residual Inception module for the input initial conditions for kernel sizes of (panel a) $7 \times 7 \times 7$, (panel b) $5 \times 5 \times 5$, and (panel c) $3 \times 3 \times 3$, from the top to bottom, respectively. Note that the central slices of the 3D kernels and feature maps are illustrated, with the colour scale anchored for a given row. The initial conditions are also depicted for comparison. As for the LR density field, the radial features are most informative for the larger kernels, but the weights are spread over a wider distribution, which characterizes the causal information encoded over the scales relevant for structure formation.

with HR simulations emulated from the same LR field, but different initial conditions, the network will learn to ignore the small-scale fluctuations and produce summaries that are robust to small-scale non-linearities (analogous to the nuisance-hardened summaries in Alsing & Wandelt 2019).

Finally, we have so far only considered using the emulator to give us approximations to the data via a simple physically motivated fitting procedure. It would be interesting to incorporate a probabilistic model of how likely features are to exist in the data, which would allow us to really sample from the known distribution of data. In this procedure, one would fit a conditional pixelwise probability estimator (Lanusse, Melchior & Moolekamp 2019), but rather than using the output as a probabilistic function, the uncertainties in the fitting could be pushed back to the feature kernels. By analysing how

likely any kernel is, we will be able to understand physically how the interactions between the LR density field and HR initial conditions provide us with the structures that appear in the HR simulations. Since the number of weights in this model is still relatively large, we would propose using variational inference-type fitting, which is becoming very promising using frameworks such as TENSORFLOW PROBABILITY (Dillon et al. 2017). Note that here we are looking for indications of uncertainty in the kernels to learn about the physical processes, rather than trusting the probabilistic interpretation of the network for true inference, for which we would require more direct sampling schemes, such as those used in Charnock et al. (2019), and a good comprehension of the likelihood of the weights in the model. However, once fitted, realizations of the approximate HR images could be obtained for many samples from the uncertain kernels,

and the mean and variance HR maps could be produced, describing where best to trust and not to trust the approximate function.

ACKNOWLEDGEMENTS

We thank the reviewer for their constructive feedback that helped to improve the quality of this paper. We express our appreciation to Guilhem Lavaux for his valuable insights. We acknowledge financial support from the ILP LABEX, under reference ANR-10-LABX-63, which is financed by French state funds managed by the ANR within the Investissements d’Avenir programme under reference ANR-11-IDEX-0004-02. DKR is a DARK fellow supported by a Semper Ardens grant from the Carlsberg Foundation (reference CF15-0384). Part of the work of FVN and BDW has been supported by the Simons Foundation. TC is supported by the ANR BIG4 project, grant ANR-16-CE23-0002 of the French Agence Nationale de la Recherche. TC would like to thank NVIDIA for the Quadro P6000 used in this research. The bispectrum and void finding computations were performed using the publicly available PYLIANS² code. This work is done within the Aquila Consortium.³

REFERENCES

- Abadi M. et al., 2016, preprint ([arXiv:1603.04467](https://arxiv.org/abs/1603.04467))
- Alsing J., Wandelt B., 2019, *MNRAS*, 488, 5093
- Arjovsky M., Bottou L., 2017, preprint ([arXiv:1701.04862](https://arxiv.org/abs/1701.04862))
- Arjovsky M., Chintala S., Bottou L., 2017, preprint ([arXiv:1701.07875](https://arxiv.org/abs/1701.07875))
- Banerjee A., Dalal N., 2016, *J. Cosmol. Astropart. Phys.*, 2016, 015
- Berger P., Stein G., 2019, *MNRAS*, 482, 2861
- Bernardini M., Mayer L., Reed D., Feldmann R., 2019, preprint ([arXiv:1912.04299](https://arxiv.org/abs/1912.04299))
- Charnock T., Lavaux G., Wandelt B. D., 2018, *Phys. Rev. D*, 97, 083004
- Charnock T., Lavaux G., Wandelt B. D., Sarma Boruah S., Jasche J., Hudson M. J., 2019, *MNRAS*, 494, 50
- Dillon J. V. et al., 2017, preprint ([arXiv:1711.10604](https://arxiv.org/abs/1711.10604))
- Feder R. M., Berger P., Stein G., 2020, preprint ([arXiv:2005.03050](https://arxiv.org/abs/2005.03050))
- Giusarma E., Reyes Hurtado M., Villaescusa-Navarro F., He S., Ho S., Hahn C., 2019, preprint ([arXiv:1910.04255](https://arxiv.org/abs/1910.04255))
- Goodfellow I. J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y., 2014, preprint ([arXiv:1406.2661](https://arxiv.org/abs/1406.2661))
- Gulrajani I., Ahmed F., Arjovsky M., Dumoulin V., Courville A., 2017, preprint ([arXiv:1704.00028](https://arxiv.org/abs/1704.00028))
- Hamaus N., Pisani A., Sutter P. M., Lavaux G., Escoffier S., Wandelt B. D., Weller J., 2016, *Phys. Rev. Lett.*, 117, 091302
- He K., Zhang X., Ren S., Sun J., 2016, Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 770
- He S., Li Y., Feng Y., Ho S., Ravanbakhsh S., Chen W., Póczos B., 2019, *Proc. Natl. Acad. Sci.*, 116, 13825
- Ivezic Z. et al., 2008, *AJ*, 873, 44
- Jasche J., Lavaux G., 2019, *A&A*, 625, A64
- Jasche J., Wandelt B. D., 2013, *MNRAS*, 432, 894
- Kingma D. P., Ba J., 2014, preprint ([arXiv:1412.6980](https://arxiv.org/abs/1412.6980))
- Kodi Ramanah D., Charnock T., Lavaux G., 2019a, *Phys. Rev. D*, 100, 043515 (KCL19)
- Kodi Ramanah D., Lavaux G., Jasche J., Wandelt B. D., 2019b, *A&A*, 621, A69
- Lanusse F., Melchior P., Moolekamp F., 2019, preprint ([arXiv:1912.03980](https://arxiv.org/abs/1912.03980))
- Laureijs R. et al., 2011, preprint ([arXiv:1110.3193](https://arxiv.org/abs/1110.3193))
- Lavaux G., Wandelt B. D., 2012, *ApJ*, 754, 109
- Lavaux G., Jasche J., Leclercq F., 2019, preprint ([arXiv:1909.06396](https://arxiv.org/abs/1909.06396))
- Lewis A., Challinor A., Lasenby A., 2000, *AJ*, 538, 473
- List F., Bhat I., Lewis G. F., 2019, *MNRAS*, 490, 3134
- Mustafa M., Bard D., Bhimji W., Lukić Z., Al-Rfou R., Kratochvil J. M., 2019, *Comp. Astrophys. Cosmol.*, 6, 1
- Perraudin N., Srivastava A., Lucchi A., Kacprzak T., Hofmann T., Réfrégier A., 2019, *Comput. Astrophys. Cosmol.*, 6, 5
- Planck Collaboration VI, 2018, preprint ([arXiv:1807.06209](https://arxiv.org/abs/1807.06209))
- Porqueres N., Jasche J., Lavaux G., Enßlin T., 2019, *A&A*, 630, A151
- Radford A., Metz L., Chintala S., 2015, preprint ([arXiv:1511.06434](https://arxiv.org/abs/1511.06434))
- Rodríguez A. C., Kacprzak T., Lucchi A., Amara A., Sgier R., Fluri J., Hofmann T., Réfrégier A., 2018, *Comp. Astrophys. Cosmol.*, 5, 4
- Ronneberger O., Fischer P., Brox T., 2015, in Navab N., Hornegger J., Wells W., Frangi A., eds, Lecture Notes in Computer Science, Vol. 9351, U-Net: Convolutional Networks for Biomedical Image Segmentation, Springer, Cham, p. 234
- Smith S. L., Kindermans P.-J., Ying C., Le Q. V., 2017, preprint ([arXiv:1711.00489](https://arxiv.org/abs/1711.00489))
- Springel V., 2005, *MNRAS*, 364, 1105
- Szegedy C. et al., 2015, preprint ([arXiv:1409.4842](https://arxiv.org/abs/1409.4842))
- Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z., 2016, preprint ([arXiv:1512.00567](https://arxiv.org/abs/1512.00567))
- Szegedy C., Ioffe S., Vanhoucke V., Alemi A. A., 2017, AAAI Conf. Artif. Intell., 4, 12
- Tamosiunas A., Winther H. A., Koyama K., Bacon D. J., Nichol R. C., Mawdsley B., 2020, preprint ([arXiv:2004.10223](https://arxiv.org/abs/2004.10223))
- Tröster T., Ferguson C., Harnois-Déraps J., McCarthy I. G., 2019, *MNRAS*, 487, L24
- Villaescusa-Navarro F. et al., 2019, preprint ([arXiv:1909.05273](https://arxiv.org/abs/1909.05273))
- Zamudio-Fernandez J., Okan A., Villaescusa-Navarro F., Bilaloglu S., Derin Cengiz A., He S., Perreault Levasseur L., Ho S., 2019, preprint ([arXiv:1904.12846](https://arxiv.org/abs/1904.12846))
- Zhang X., Wang Y., Zhang W., Sun Y., He S., Contardo G., Villaescusa-Navarro F., Ho S., 2019, preprint ([arXiv:1902.05965](https://arxiv.org/abs/1902.05965))

²<https://github.com/franciscovillaescusa/Pylians>.

³<https://aquila-consortium.org>.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.