

# Aletheia: Emulating the non-linear matter power spectrum in the context of evolution mapping

Ariel G. Sánchez,<sup>1,2\*</sup> Andrés N. Ruiz,<sup>3,4</sup> Facundo Rodriguez,<sup>3,4</sup> Carlos Correa,<sup>1</sup> Andrea Fiorilli,<sup>1</sup> Matteo Esposito,<sup>1</sup> Jenny Gonzalez-Jara,<sup>5</sup> Nelson D. Padilla,<sup>3,4</sup> Alejandro Pérez-Fernández,<sup>1</sup> and Sofia Contarini<sup>1</sup>

<sup>1</sup>Max-Planck-Institut für extraterrestrische Physik, Postfach 1312, Giessenbachstr., 85748 Garching, Germany

<sup>2</sup>Universitäts-Sternwarte München, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstrasse 1, 81679 München, Germany

<sup>3</sup>Instituto de Astronomía Teórica y Experimental, CONICET-UNC, Laprida 854, X5000BGR, Córdoba, Argentina

<sup>4</sup>Observatorio Astronómico, Universidad Nacional de Córdoba, Laprida 854, X5000BGR, Córdoba, Argentina

<sup>5</sup>Instituto de Astrofísica, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Santiago, Chile

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

We present Aletheia, a new emulator of the non-linear matter power spectrum,  $P(k)$ , built upon the evolution mapping framework. This framework addresses the limitations of traditional emulation by focusing on  $h$ -independent cosmological parameters, which can be separated into those defining the linear power spectrum shape ( $\Theta_s$ ) and those affecting only its amplitude evolution ( $\Theta_e$ ). The combined impact of evolution parameters and redshift is compressed into a single amplitude parameter,  $\sigma_{12}$ . Aletheia uses a two-stage Gaussian Process emulation: a primary emulator predicts the non-linear boost factor as a function of ( $\Theta_s$ ) and  $\sigma_{12}$  for fixed evolution parameters, while a second one applies a small linear correction based on the integrated growth history. The emulator is trained on shape parameters spanning  $\pm 5\sigma$  of Planck constraints and a wide clustering range  $0.2 < \sigma_{12} < 1.0$ , providing predictions for  $0.006 \text{ Mpc}^{-1} < k < 2 \text{ Mpc}^{-1}$ . We validate Aletheia against N-body simulations, demonstrating sub-percent accuracy. When tested on a suite of dynamic dark energy models, the full emulator's predictions show a variance of approximately 0.2%, a factor of five smaller than that of the state-of-the-art EuclidEmulator2 (around 1% variance). Furthermore, Aletheia maintains sub-percent accuracy for the best-fit dynamic dark energy cosmology from recent DESI data, a model whose parameters lie outside the training ranges of most conventional emulators. This demonstrates the power of the evolution mapping approach, providing a robust and extensible tool for precision cosmology.

**Key words:** cosmology: theory – large-scale structure of Universe – methods: numerical – methods: statistical

## 1 INTRODUCTION

The large-scale structure (LSS) of the Universe, as traced by the spatial distribution of galaxies, provides one of the most powerful avenues for constraining cosmological models and understanding the fundamental nature of gravity, dark matter, and dark energy. Analyses of LSS, particularly through measurements of baryon acoustic oscillations (BAO) and redshift-space distortions (RSD), have become cornerstone techniques in modern cosmology (e.g., Eisenstein et al. 2005; Cole et al. 2005; Sánchez et al. 2006; Anderson et al. 2014; Alam et al. 2017; Alam et al. 2021). The recent results from the Dark Energy Spectroscopic Instrument (DESI, DESI Collaboration et al. 2025c,a,b), have further underscored the constraining power of these methods. Together with the *Euclid* mission (Euclid Collaboration et al. 2025), these datasets promise to usher in an era of unprecedented precision. These advancements in galaxy surveys are complemented by a wealth of new data from several cosmological probes, including cosmic microwave background (CMB) measurements from the *Planck* satellite (Planck Collaboration et al. 2020;

Tristram et al. 2024) and ground-based experiments like the Atacama Cosmology Telescope (ACT, Louis et al. 2025), type Ia Supernovae (SN) compilations (e.g., Scolnic et al. 2018, 2022; Brout et al. 2022), and weak lensing shear data from surveys such as the Dark Energy Survey (DES, Abbott et al. 2022), the Kilo-Degree Survey (KiDS, Wright et al. 2025), and the Hyper Suprime-Cam Subaru Strategic Program (HSC, Dalal et al. 2023; Li et al. 2023), marking a truly data-rich era in cosmology.

This influx of high-quality data is beginning to reveal intriguing, albeit tentative, hints of physics beyond the standard Lambda cold dark matter ( $\Lambda$ CDM) model. For instance, the combination of DESI data with Planck and SN observations shows a preference for dynamic dark energy models over a cosmological constant (DESI Collaboration et al. 2025b,a). To rigorously test such claims and fully exploit the potential of forthcoming datasets, we require analysis tools that are not only highly accurate but also robust across a wide range of cosmological scenarios.

At their core, the current state-of-the-art LSS analyses, including those from DESI, build upon methodologies developed for previous surveys (e.g., Sánchez et al. 2017; Grieb et al. 2017; Tröster et al. 2020; d'Amico et al. 2020; Ivanov et al. 2020). These rely on per-

\* E-mail: arielsan@mpe.mpg.de

turbation theory to model the clustering of galaxies in the mildly non-linear regime. However, these recipes inherently struggle as the analyses are pushed to smaller scales, where non-linear evolution becomes dominant and perturbative approaches become invalid.

An alternative is to use N-body simulations directly as theoretical predictions. While ideal for capturing non-linear evolution, the computational expense of running such simulations for every point in a vast cosmological parameter space is prohibitive for most analyses. Fortunately, non-linear clustering statistics, like the matter power spectrum,  $P(k)$ , typically exhibit a smooth dependence on cosmological parameters. This characteristic enables the development of emulators: sophisticated interpolation schemes trained on a set of N-body simulations. Once calibrated, emulators can deliver rapid and accurate predictions for non-linear quantities (e.g., Heitmann et al. 2010, 2016; Garrison et al. 2018; DeRose et al. 2019; Bocquet et al. 2020; Euclid Collaboration et al. 2019, 2021; Angulo et al. 2021; Yang et al. 2025).

Despite their utility, emulators are inherently limited by the parameter space and redshifts sampled during their calibration. The high computational cost of the underlying simulations means that the number of training nodes must be kept manageable. The vast parameter space of possible cosmological models means that current emulators are often focused on narrow regions around a fiducial  $\Lambda$ CDM cosmology to maintain accuracy, limiting their applicability to explorations of more exotic models or broader parameter variations.

These challenges are exacerbated by the very language traditionally used to describe cosmic structure, which is commonly based on quantities that depend explicitly on the dimensionless Hubble parameter  $h$ . Sánchez (2020) highlighted the complications associated with this common practice, which obscures the information content of different cosmological observables.

Building directly upon these insights, the evolution mapping framework introduced in Sánchez et al. (2022) offers a new framework that reframes how we model the non-linear Universe. Its core principle is to adopt a parametrisation free from any explicit dependence on the dimensionless Hubble parameter. This enables a separation of cosmological parameters into two distinct classes: shape parameters, which determine the shape of the linear matter power spectrum, and evolution parameters, which govern the subsequent growth of structure. At the linear level, the impact of different evolution parameters can be compressed into a single variable characterising the global amplitude of density fluctuations.

Sánchez et al. (2022) also outlined the practical application of evolution mapping for designing a new class of emulators of the matter power spectrum or other statistics. In this scheme, the emulator does not directly sample the full space of evolution parameters, nor does it need to explicitly model the redshift dependence of  $P(k)$ . Instead, by focusing on how the non-linear  $P(k)$  responds to changes in the clustering amplitude, this approach can lead to more general predictions valid over a wider range of cosmologies and redshifts. The reduced dimensionality of the problem leads to more accurate predictions from a given number of training simulations.

In this paper, we present **Aletheia**<sup>1</sup>, a new emulator for the non-linear matter power spectrum built upon the evolution mapping framework. **Aletheia** is designed as a proof-of-concept demonstrating the power and flexibility of this emulation strategy. Specifically, we emulate the ratio of the non-linear  $P(k)$  to its de-wiggled linear counterpart (preserving the broadband shape but damping the BAO

signal) as a function of shape parameters and clustering amplitude for a fixed set of reference evolution parameters. The effect of varying evolution parameters is captured by a second emulator for the dependence of the power spectrum on the growth of the structure history. This allows **Aletheia** to efficiently make predictions for general cosmologies while relying on simulations that provide only a single snapshot for each sampled cosmology. We construct **Aletheia** using Gaussian Process regression trained on a suite of N-body simulations specifically designed for this emulator. We validate its performance against independent test simulations, including a dedicated set of dynamic dark energy models, demonstrating excellent accuracy. **Aletheia** shows good agreement with other established emulators such as **EuclidEmulator2** (Euclid Collaboration et al. 2021), while offering predictions over a wider range of cosmological models and redshifts.

The structure of this paper is as follows. In Section 2, we briefly review the general idea of evolution mapping and introduce the parameter  $\tilde{x}$  used to describe deviations from the exact evolution mapping relation. Section 3 details the design of the **Aletheia** emulator, including the parameter space sampled (Section 3.1), the cosmological simulations used for training (Section 3.2), the Gaussian Process emulation methodology (Section 3.3), and the implementation of a high- $k$  resolution correction necessary for the final emulator product (Section 3.4). We present a thorough validation of **Aletheia** in Section 4, assessing its individual components (Section 4.1), its performance for dynamic dark energy cosmologies compared against other publicly available emulators (Section 4.2), and the impact of the high- $k$  resolution correction (Section 4.3). Finally, we summarise our findings and discuss future prospects in Section 5.

## 2 THE EVOLUTION MAPPING FRAMEWORK

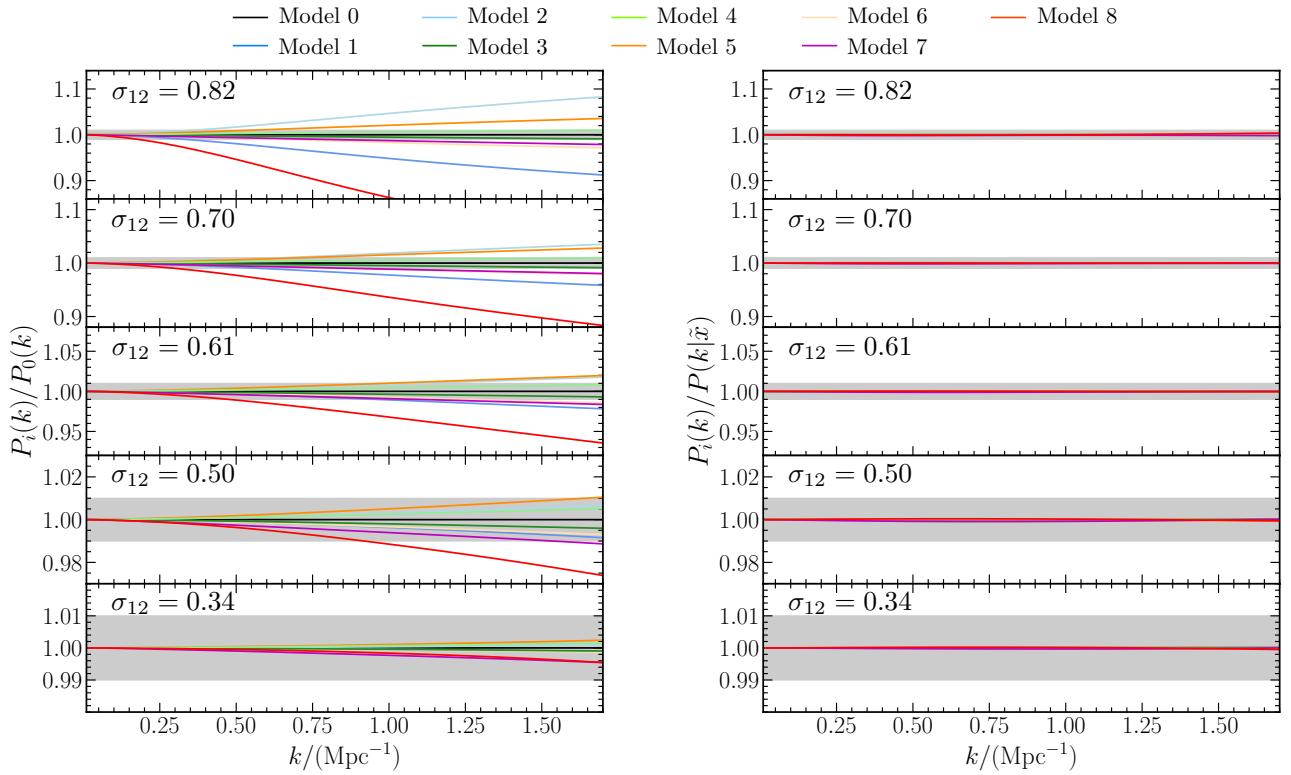
### 2.1 The evolution mapping principle in the linear regime

The framework of evolution mapping introduced by Sánchez et al. (2022) provides a powerful way to understand and model the non-linear evolution of the density field across a wide range of cosmological models. This approach focuses on a set of cosmological parameters free from any explicit dependency on the dimensionless Hubble parameter,  $h$ . It classifies them according to their impact on the linear matter power spectrum,  $P_L(k)$ , into two categories:

- Shape parameters,  $\Theta_s$ , which define the shape of  $P_L(k)$ . Examples of these parameters are the physical baryon density,  $\omega_b$ , the physical cold dark matter density,  $\omega_c$ , and the primordial scalar spectral index,  $n_s$ .
- Evolution parameters,  $\Theta_e$ , which only affect the amplitude of  $P_L(k)$  at any given redshift  $z$  through their impact on the growth rate of cosmic structures. This set includes the physical density parameters of dark energy,  $\omega_{DE}$ , and curvature,  $\omega_K$ , the primordial scalar amplitude,  $A_s$ , as well as the parameters describing the equation of state of dark energy, such as  $w_0$  and  $w_a$  (Chevallier & Polarski 2001; Linder 2003).

This classification is useful as, at the linear level, the impact of all evolution parameters and  $z$  follows a perfect degeneracy and can therefore be fully described by a single quantity characterising the amplitude of  $P_L(k)$ . It is common to describe the amplitude of density fluctuations in terms of the RMS of the linearly evolved mass contrast in spheres of radius  $R = 8 h^{-1} \text{Mpc}$ , which we denote as  $\sigma_{8/h}$  to emphasise its explicit dependence on  $h$ . This dependency makes  $\sigma_{8/h}$  unsuitable for cleanly describing the degeneracy involving evolution parameters. Since  $h$  is given by the sum of all physical energy

<sup>1</sup> From the ancient Greek *alētheia* (pronounced /a le .t̪e .a/), meaning ‘truth’ or ‘disclosure’. In Greek mythology, Aletheia was the personification of truth.



**Figure 1.** The left panel shows the ratio of non-linear power spectra for a set of simulations with identical shape parameters but widely varying evolution parameters, all evaluated at redshifts that correspond to the same values of  $\sigma_{12}$ . The specific cosmological parameters for each model are detailed in Table 2 of Esposito et al. (2024). The right panel shows the ratio of the true  $P(k)$  to the reference prediction (using  $\tilde{x}_0$ ) and the resulting derivative  $\partial R/\partial \tilde{x}$  measured from the simulations, illustrating the accuracy of the first-order Taylor expansion in equation (5).

contributions as

$$h^2 = \sum_i \omega_i, \quad (1)$$

it represents a combination of both shape and evolution parameters. Therefore, its inclusion in  $\sigma_{8/h}$  obscures the fundamental distinction between these parameter classes and hides the degeneracy. Sánchez (2020) showed that this issue can be avoided by using a fixed reference scale in Mpc. A convenient choice is  $\sigma_{12}$ , introduced by Sánchez (2020) as the RMS of linear density fluctuations smoothed over spheres of radius 12 Mpc, as it provides a value comparable to  $\sigma_{8/h}$  for typical cosmological parameters. While other physical characterisations are possible, such as  $\sigma(R)$  for any  $R$  in Mpc units, or the dimensionless power spectrum  $\Delta_L^2(k_p)$  at a reference wavenumber  $k_p$  in  $Mpc^{-1}$  units, we follow the approach of Sánchez et al. (2022) and use  $\sigma_{12}$  for consistency.

The degeneracy between the evolution parameters  $\Theta_e$  and redshift  $z$  can therefore be expressed as

$$P_L(k|z, \Theta_s, \Theta_e) = P_L(k|\Theta_s, \sigma_{12}(z, \Theta_s, \Theta_e)). \quad (2)$$

This equation, referred to as the evolution mapping relation for the linear power spectrum, implies that the time evolution of  $P_L(k)$  in models with the same  $\Theta_s$  but different  $\Theta_e$  can be mapped from one to another simply by relabelling the redshifts that correspond to the same  $\sigma_{12}$ .

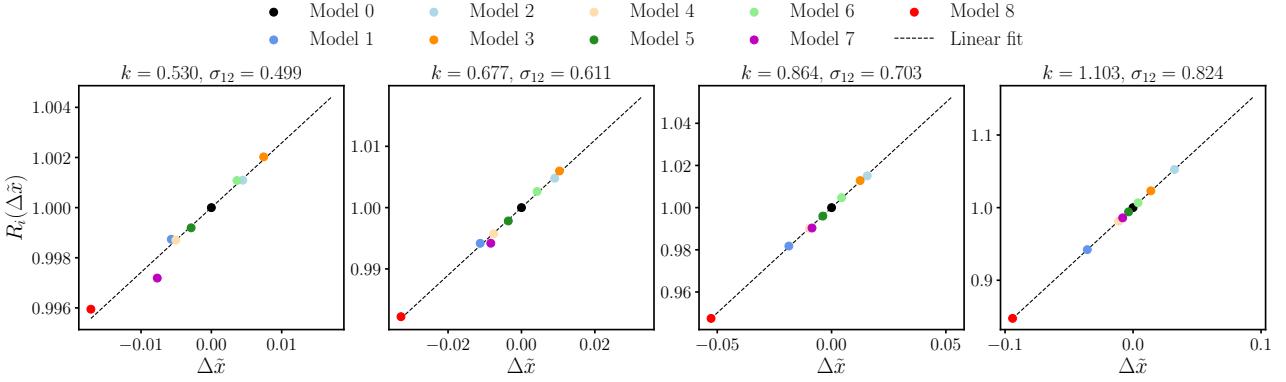
The left panel of Fig. 2 in Sánchez et al. (2022) provides a clear illustration of this principle, showing identical linear power spectra for nine cosmological models with identical shape parameters but vastly different evolution parameters when evaluated at redshifts where their respective  $\sigma_{12}(z)$  values match.

This framework has been extended to describe statistics of the velocity field (Esposito et al. 2024) and to incorporate the impact of massive neutrinos (Pezzotta et al. 2025), highlighting the robustness of the evolution mapping principle.

## 2.2 Deviations in the non-linear regime

The non-linear evolution of the matter power spectrum,  $P(k)$ , is predominantly determined by its linear-theory counterpart,  $P_L(k)$  (e.g., Hamilton et al. 1991; Peacock & Dodds 1996; Smith et al. 2003). Standard perturbation theory (SPT) and related approaches suggest that, if the perturbation theory kernels are independent of cosmology (a good approximation, see e.g., Takahashi 2008; Taruya 2016; Garny & Taule 2021), then the non-linear  $P(k)$  should be a function of  $P_L(k)$  that is independent of the specific cosmological parameters beyond those encapsulated in the power spectrum itself (see, e.g., Scoccimarro et al. 1998). This is exploited by the emulator COMET, which uses evolution mapping to emulate the perturbation theory models commonly employed in present-day LSS analyses (Eggemeier et al. 2023; Pezzotta et al. 2025). Under this assumption, the evolution mapping relation of equation (2) would be expected to hold even for the fully non-linear  $P(k)$ .

Sánchez et al. (2022) used  $P(k)$  measurements from N-body simulations to show that the evolution mapping relation provides a remarkably good description even in the non-linear regime. Indeed, cosmologies with identical shape parameters and widely different evolution parameters yield very similar non-linear structure when viewed at equal  $\sigma_{12}$ . However, these models exhibit small deviations in their non-linear power spectra, particularly in the deeply non-linear



**Figure 2.** The ratio of power spectra,  $R(k)$ , as a function of the integrated growth history parameter  $\tilde{x}$  for four choices of  $k$  and  $\sigma_{12}$ , as indicated in the legend. The points show the measurements from the Aletheia simulations, where each colour represents a different cosmology (with parameters detailed in Table 2 of Esposito et al. 2024), all evaluated at a fixed  $\sigma_{12}$ . The dashed lines show the best-fitting linear relation for each case. The remarkable linearity of the response validates the first-order approximation in equation (5) and allows for a direct measurement of the derivative  $\partial R/\partial \tilde{x}$  from the simulations.

regime. This is illustrated in the left panel of Fig. 1 (which is analogous to Fig. 3 of Sánchez et al. 2022). These differences increase with wavenumber  $k$  and with  $\sigma_{12}$ , although they remain at the sub-percent to few-percent level for a wide range of scales and cosmologies.

To account for these deviations, we must identify a parameter that captures the relevant differences in the growth history. In SPT, the primary cosmology dependence of the solutions, beyond the linear power spectrum itself, is encoded in the parameter combination

$$x(z) = \frac{\Omega_m(z)}{f^2(z)}, \quad (3)$$

where  $\Omega_m(z)$  is the fractional matter density parameter and  $f(z) = d \ln D(z)/d \ln a$  is the linear growth rate parameter (Scoccimarro et al. 1998). While the instantaneous value of  $x(z)$  is important, we find that the small deviations from the evolution mapping relation are sensitive to the recent growth history. We therefore define a new parameter,  $\tilde{x}$ , which is an average of  $x$  over the past history, using  $\tau = \ln(\sigma_{12})$  as the time variable. This parameter is defined as

$$\tilde{x}(\tau) = \int_{-\infty}^{\tau} x(\tau') K(\tau - \tau' | \eta) d\tau', \quad (4)$$

where  $K(\tau | \eta)$  is a Gaussian kernel with a width  $\eta$  that represents a characteristic memory of the non-linear evolution, indicating the  $\Delta \ln(\sigma_{12})$  interval over which the past growth history significantly impacts the power spectrum at any time. We find that a fixed value of  $\eta = 0.12$  provides an excellent description of the deviations for all relevant cosmologies and scales considered here, so we adopt it throughout.

This new parameter,  $\tilde{x}$ , effectively describes the differences in the power spectrum due to varying structure growth histories. For cosmologies sharing the same  $\Theta_s$  and  $\sigma_{12}$ , but having different values of  $\tilde{x}$ , we can approximate the ratio of their non-linear power spectra with a first-order Taylor expansion around a reference value  $\tilde{x}_0$  as

$$R(k, \tilde{x}) \equiv \frac{P(k|\Theta_s, \sigma_{12}, \tilde{x})}{P(k|\Theta_s, \sigma_{12}, \tilde{x}_0)} \approx 1 + \left. \frac{\partial R(k)}{\partial \tilde{x}} \right|_{\tilde{x}_0} (\tilde{x} - \tilde{x}_0). \quad (5)$$

This linear correction is sufficient to capture the bulk of the differences in  $P(k)$  due to varying growth histories, forming the basis of our emulator design.

Figure 2 provides strong empirical support for this approach. It shows the ratios  $R(k, \tilde{x})$  measured from the Aletheia simulations of Esposito et al. (2024) for four different choices of  $k$  and  $\sigma_{12}$ . The

specific cosmological parameters for each model shown are detailed in Table 2 of Esposito et al. (2024). The results demonstrate that the dependence of this ratio on  $\tilde{x}$  is remarkably linear. The dashed lines show the best-fitting linear relation, and their excellent agreement justifies the use of the Taylor expansion in equation (5). The slope of this linear fit can be interpreted as a direct measurement of the derivative  $\partial R(k)/\partial \tilde{x}$ , which is the second key quantity our emulators are designed to predict. The remarkable accuracy of this first-order approximation is empirically validated in the right panel of Fig. 1, where the ratio of the true  $P(k)$  to the prediction using the  $\tilde{x}_0$  reference model agrees with unity to well below the 1% level.

### 3 EMULATING THE MATTER POWER SPECTRUM

Building upon the evolution mapping framework described in Section 2, we introduce **Aletheia**, a new emulator for the non-linear matter power spectrum,  $P(k)$ . The design of **Aletheia** aims to provide accurate and general predictions for  $P(k)$  by separating the emulation task into two distinct components. This approach uses the fact that once the values the shape parameters  $\Theta_s$  and  $\sigma_{12}$  are fixed, the non-linear  $P(k)$  is largely determined, with only minor corrections needed to account for differing structure growth histories.

#### 3.1 Emulator design

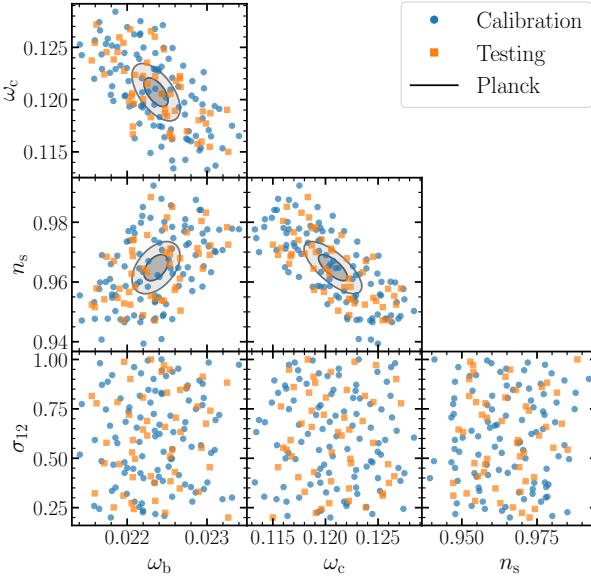
The **Aletheia** emulator predicts the non-linear matter power spectrum,  $P(k|\Theta_s, \Theta_e, z)$ , for a target cosmology defined by its shape parameters,  $\Theta_s = (\omega_b, \omega_c, n_s)$ , and evolution parameters,  $\Theta_e$ , using a two-stage process.

The first stage focuses on modelling the overall change in the broad-band shape of the non-linear power spectrum for a fixed reference set of evolution parameters,  $\Theta_{e0}$ . We define the boost factor,  $B(k)$ , as the ratio of the non-linear power spectrum to its de-wiggle linear counterpart,  $P_{DW}(k)$ , as

$$B(k|\Theta_s, \sigma_{12}) = \frac{P(k|\Theta_s, \Theta_{e0}, \sigma_{12})}{P_{DW}(k|\Theta_s, \Theta_{e0}, \sigma_{12})}. \quad (6)$$

Here,  $P_{DW}(k)$  is a version of the linear power spectrum where the BAO signal is damped following the predictions of perturbation theory. It is constructed as

$$P_{DW}(k) = P_L(k)G(k) + P_{NW}(k)[1 - G(k)]. \quad (7)$$



**Figure 3.** Distribution of cosmological parameters for the training (blue points) and testing (orange points) sets of simulations used for  $\mathcal{E}_B(k)$ . The panels show 2D projections of the parameter space ( $\omega_b$ ,  $\omega_c$ ,  $n_s$ ,  $\sigma_{12}$ ). The grey ellipses represent the  $1\sigma$  and  $2\sigma$  confidence regions derived from *Planck* 2018 data for the shape parameters ( $\omega_b$ ,  $\omega_c$ ,  $n_s$ ). Our sampling strategy uses the eigenvector directions of the *Planck* covariance matrix to broadly cover the relevant parameter space. The sampling for  $\sigma_{12}$  covers the range from 0.2 to 1.0.

Here,  $P_{\text{NW}}(k)$  is a smoothed version of the linear power spectrum that retains its broadband shape but has the BAO signal removed, computed using the discrete sine transform algorithm proposed in Hamann et al. (2010). The function  $G(k)$  is a Gaussian damping kernel,

$$G(k) = \exp\left(-\frac{1}{2}k^2\sigma_v^2\right), \quad (8)$$

which depends on the linear velocity dispersion,  $\sigma_v^2$ , calculated from the linear power spectrum as

$$\sigma_v^2 = \frac{1}{6\pi^2} \int_0^\infty P_L(k) dk. \quad (9)$$

This formulation smoothly transitions from the full linear power spectrum at large scales to the non-wiggled component at small scales.

We build an emulator of the boost factor,  $\mathcal{E}_B(k)$ , solely as a function of the shape parameters  $\Theta_s$  and the clustering amplitude  $\sigma_{12}$ . Specifically, for each training cosmology defined by  $(\Theta_s, \sigma_{12})$ , we run an N-body simulation with evolution parameters  $\Theta_{e0}$  until the redshift  $z$  at which  $\sigma_{12}(z, \Theta_s, \Theta_{e0})$  matches the target  $\sigma_{12}$  of the node. In practice, rather than fixing the evolution parameter  $\omega_{DE}$  for  $\Theta_{e0}$ , we fix the value of  $h$ . Since  $\omega_b$  and  $\omega_c$  vary across the training set, this implies that  $\omega_{DE}$  subtly varies to maintain the fixed  $h$  for a flat universe (see Section 3.2).

The second stage in our prediction of  $P(k)$  accounts for the small deviations that arise when the target cosmology has different evolution parameters  $\Theta_e$  than the reference set  $\Theta_{e0}$ . We define the ratio

$$R(k|\Theta_s, \Theta_e, \sigma_{12}) = \frac{P(k|\Theta_s, \Theta_e, \sigma_{12})}{P(k|\Theta_s, \Theta_{e0}, \sigma_{12})}, \quad (10)$$

where both power spectra are evaluated at the same  $\Theta_s$  and  $\sigma_{12}$ , but

with differing evolution parameters, leading to different integrated growth histories,  $\tilde{x}$  and  $\tilde{x}_{e0}$ . Following the linear approximation of equation (5), we estimate this ratio as

$$R(k|\Theta_s, \Theta_e, \sigma_{12}) \approx 1 + \frac{\partial R(k)}{\partial \tilde{x}} \Big|_{\tilde{x}_{e0}} (\tilde{x} - \tilde{x}_{e0}). \quad (11)$$

We build a second emulator,  $\mathcal{E}_{\partial R/\partial \tilde{x}}$ , to predict the derivative term  $\partial R(k)/\partial \tilde{x}$  as a function of  $(\Theta_s, \sigma_{12})$ .

The final prediction for the non-linear power spectrum for an arbitrary cosmology  $(\Theta_s, \Theta_e)$  at a redshift  $z$ , characterised by the clustering amplitude  $\sigma_{12}$  and the integrated growth history parameter  $\tilde{x}$ , is given by

$$\begin{aligned} \mathcal{E}_P(k|\Theta_s, \Theta_e, z) &= P_{\text{DW}}(k|\Theta_s, \Theta_e, z) \times \mathcal{E}_B(k|\Theta_s, \sigma_{12}) \\ &\times [1 + \mathcal{E}_{\partial R/\partial \tilde{x}}(k|\Theta_s, \sigma_{12}) (\tilde{x} - \tilde{x}_{e0})]. \end{aligned} \quad (12)$$

Here,  $P_{\text{DW}}(k|\Theta_s, \Theta_e, z)$  is the de-wiggled linear power spectrum of the target cosmology and  $\tilde{x}_{e0}$  is the value of the integrated growth history parameter for the reference evolution parameters  $\Theta_{e0}$  evaluated at the redshift  $z_0$  that corresponds to the same clustering amplitude, i.e.,  $\sigma_{12}(z_0, \Theta_s, \Theta_{e0}) = \sigma_{12}(z, \Theta_s, \Theta_e)$ .

By construction, this design ensures that the emulator is broadly applicable to various evolution histories, including dynamic dark energy models, without explicitly sampling their parameters or redshift during the primary training of the emulators  $\mathcal{E}_B$  and  $\mathcal{E}_{dR/dx}$ , which only depend on  $(\omega_b, \omega_c, n_s, \sigma_{12}, \ln k)$ . The value of  $\ln k$  is treated as an input parameter for both emulators, allowing their evaluation at any desired  $k$  value. The ranges for these parameters will be detailed in Section 3.2.

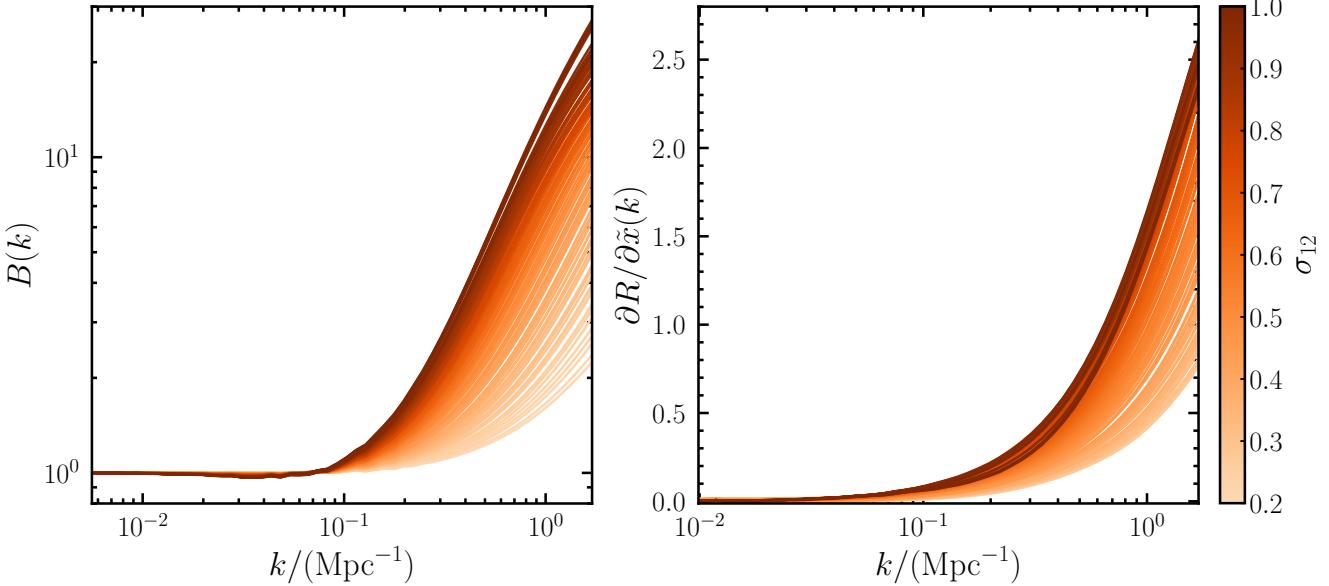
### 3.2 Cosmological simulations

Two different sets of N-body simulations, which we collectively refer to as the AletheiaEmu suite, were performed to train and test the Aletheia emulators. All simulations use initial conditions generated at  $z_{\text{IC}} = 99$  with 2LPTic (Crocce et al. 2006, 2012) and evolved to the desired redshift using the publicly available version of the code GADGET-4 (Springel et al. 2021). Both codes were modified to incorporate various dark energy models at the level of the background expansion. The Plummer-equivalent softening length was set appropriately for the resolution of each simulation set.

#### 3.2.1 Simulations for $\mathcal{E}_B(k)$

To train the emulator  $\mathcal{E}_B(k)$ , we generated a set of 100 cosmological models using a four-dimensional maximin Latin hypercube (LH) design (McKay et al. 1979; Stein 1987). The sampling of the shape parameters ( $\omega_b$ ,  $\omega_c$ ,  $n_s$ ) for these models was designed to efficiently cover the physically relevant parameter space, based on current cosmological constraints. Rather than directly sampling these parameters, we used the covariance matrix of constraints derived from *Planck* data (Planck Collaboration et al. 2020). These constraints are largely robust to assumptions about dark energy evolution, making them suitable for defining the shape parameter space.

We computed the eigenvalues and eigenvectors of this covariance matrix, which define the principal axes and their corresponding variances in the parameter space. The first three components of our LH sampling were identified with these eigenvector directions. For each eigenvector direction, the sampling range was set to  $\pm 5$  times the square root of its corresponding eigenvalue, ensuring a broad coverage that far exceeds the constraints from *Planck*. The fourth parameter in our LH design was associated with  $\sigma_{12}$  and sampled from



**Figure 4.** Raw training data for the Aletheia emulators. The left panel shows the measured boost factor,  $B(k)$ , for the 100 simulations to train  $\mathcal{E}_B(k)$ , colour-coded by the clustering amplitude  $\sigma_{12}$  at which they were evaluated. This panel illustrates the strong, smooth dependence of the boost factor on  $\sigma_{12}$ . The right panel displays the measured derivative term  $\partial R / \partial \tilde{x}$  for the same cosmologies, determined from the derivative simulations described in Section 3.2.2. The smooth dependency of both quantities on the cosmological parameters and  $k$  makes them well-suited for Gaussian Process emulation.

0.2 to 1.0, covering a wide range of relevant clustering amplitudes. The distribution of these 100 training cosmologies, along with their relationship to the *Planck* 2018 constraints, is illustrated in Fig. 3.

For each of these 100 nodes, the primordial scalar amplitude was kept fixed to  $A_s = 2.1 \times 10^{-9}$ , and the dimensionless Hubble parameter was fixed to  $h = 0.673$ . As  $\omega_b$  and  $\omega_c$  vary across the training set, fixing  $h$  implies that the dark energy density parameter,  $\omega_{\text{DE}}$ , varies slightly across these simulations to maintain a flat universe. This approach defines our reference evolution parameters  $\Theta_{e0}$  for each node. Each simulation was run until the specific redshift  $z_i$  at which  $\sigma_{12}(z_i)$  for that cosmology matched the target  $\sigma_{12,i}$  of the corresponding LH node. Thus, each of these primary simulations yields a single snapshot used for training  $\mathcal{E}_B(k)$ .

For each LH node, we ran two N-body simulations employing the “fixed-paired” technique to suppress cosmic variance (Angulo & Pontzen 2016). These simulations follow the evolution of  $2048^3$  dark matter particles in a periodic box of side length  $L_{\text{box}} = 1500$  Mpc. All simulation pairs use the same initial random phases to minimise stochastic differences among the nodes.

The non-linear matter power spectra of these simulations were measured in Mpc units using a Cloud-In-Cell (CIC, Hockney & Eastwood 1988) mass assignment scheme on a grid of 2048 points per side. We corrected for the effect of aliasing with the interlacing method of Sefusatti et al. (2016). This provides reliable measurements of  $P(k)$  up to the Nyquist frequency,  $k_{\text{Ny}} = 4.3$  Mpc $^{-1}$ . The power spectra from the paired simulations were averaged to obtain the final estimate of the non-linear  $P(k)$  for each node. Finally, these estimates were used to compute the boost factor,  $B(k)$ , for each node by dividing them by the corresponding de-wiggled power spectrum  $P_{\text{DW}}(k)$ , as defined in equation (6). We focus our emulation on the range  $0.006 \text{ Mpc}^{-1} < k < 2 \text{ Mpc}^{-1}$ . The upper bound is chosen to mitigate the impact of resolution effects, for which we apply a correction to the final emulator results (see Section 3.4). The left panel of Fig. 4 shows the measured  $B(k)$  for all 100 training cosmologies,

colour-coded by their corresponding values of  $\sigma_{12}$ , illustrating the primary dependence captured by  $\mathcal{E}_B(k)$ .

An additional set of 50 cosmological models was generated for testing the performance of the emulator. These testing cosmologies were sampled using a different maximin LH, but followed the same eigenvector-based parameter space definition, fixed values of  $A_s$  and  $h$ , target  $\sigma_{12}$  selection, and paired simulation specifications with  $2048^3$  particles. The distribution of these 50 testing cosmologies is also shown in Fig. 3.

### 3.2.2 Simulations for $\mathcal{E}_{\partial R / \partial \tilde{x}}(k)$

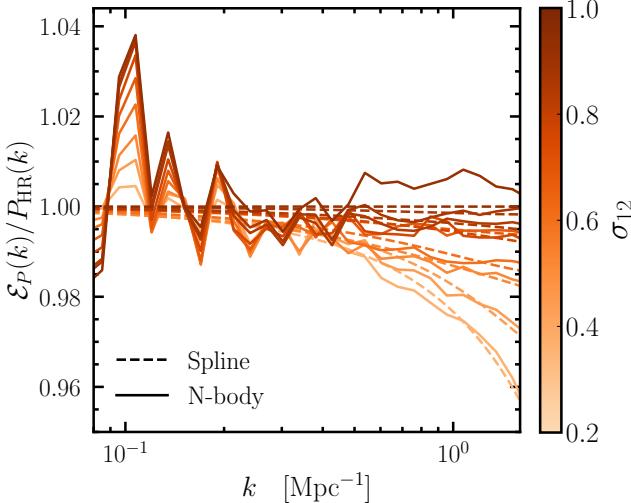
To train our second emulator  $\mathcal{E}_{\partial R / \partial \tilde{x}}(k)$ , we used the same set of 100 base cosmological models described in Section 3.2.1. For each node, we ran five additional simulations for cosmological models characterised by the same shape parameters of the main node, but exploring variations in the evolution parameters. Their purpose is to measure the response of the non-linear power spectrum to variations in the integrated growth history parameter  $\tilde{x}$ . These additional cosmologies include:

- (i) One simulation with the same base evolution parameters ( $\Theta_{e0}$ ), serving as the reference  $P_0(k)$ .
- (ii) Two simulations where the dark energy density is varied relative to the base value as  $\omega_{\text{DE}} = 0.925 \omega_{\text{DE},0}$  and  $\omega_{\text{DE}} = 1.125 \omega_{\text{DE},0}$ .
- (iii) Two simulations where the physical curvature density is set to  $\omega_K = -0.04$  and  $\omega_K = +0.02$ .

By inducing these targeted variations in the evolution parameters, we generate models with different integrated growth histories ( $\tilde{x} \neq \tilde{x}_0$ ) while keeping the clustering amplitude  $\sigma_{12}$  fixed.

All five simulations for each node were run with  $1500^3$  particles in a box of size  $L_{\text{box}} = 1500$  Mpc and used the same initial random phases. The power spectra of each simulation was measured using the same method described in Section 3.2.1.

These measurements were used to define the ratios  $R_j(k) =$



**Figure 5.** Correction of the Aletheia emulator for resolution effects. The solid lines show the measured ratio of the uncorrected emulator prediction,  $\mathcal{E}_P(k)$ , to the high-resolution power spectrum,  $P_{\text{HR}}(k)$ , derived from the AletheiaMass simulations. The lines are colour-coded by the value of  $\sigma_{12}$ . The fluctuations are caused by cosmic variance due to the smaller box size of the high-resolution runs. The dashed lines represent the corresponding two-dimensional smoothing spline interpolation,  $C(k, \sigma_{12})$ , which is used to apply the correction to the final emulator prediction.

$P_j(k)/P_0(k)$ , where  $P_0(k)$  is the power spectrum of the reference simulation ( $\Theta_{e0}$ ). Assuming the linear approximation established in equation (5), the derivative  $\partial R/\partial \tilde{x}(k)$  is then obtained by performing a linear fit to the measured  $R_j(k)$  versus  $\Delta \tilde{x} = (\tilde{x}_j - \tilde{x}_0)$ , constrained to  $R(k) = 1$  when  $\Delta \tilde{x} = 0$ .

Since the derivative term  $\partial R/\partial \tilde{x}$  is a small relative correction factor, it does not require the same demanding fidelity as the absolute  $P(k)$  measurements for  $\mathcal{E}_B(k)$ . Furthermore, we verified that the measured power spectrum ratios,  $R_j(k)$ , are insensitive to the resolution of the simulations. This allowed us to use a lower resolution for this second set of simulations than the ones used for the training of the primary emulator  $\mathcal{E}_B(k)$ , significantly reducing their computational cost.

The set of 100 derivative measurements was used to train the emulator  $\mathcal{E}_{\partial R/\partial \tilde{x}}(k)$ . As illustrated in the right panel of Fig. 4, this derivative term is a smooth function of  $k$  and the input parameters. The same procedure is applied to the 50 test cosmologies described in Section 3.2.1 by running the corresponding derivative simulations, providing an independent set of measurements for validation.

### 3.3 Gaussian process emulation

Both  $\mathcal{E}_B(k)$  and  $\mathcal{E}_{\partial R/\partial \tilde{x}}(k)$  are constructed using Gaussian Process (GP) regression, a powerful non-parametric Bayesian method for interpolating complex functions from a finite set of training data (Rasmussen & Williams 2006). A GP defines a probability distribution over functions. Given a set of training inputs  $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$  and corresponding outputs  $\mathbf{y} = \{y_1, \dots, y_N\}$ , a GP model can predict the output  $y$  at a new input point  $\mathbf{x}$  by providing a mean prediction and an estimate of the prediction uncertainty. The key component of a GP is the covariance function, or kernel,  $K(\mathbf{x}_i, \mathbf{x}_j)$ , which defines the similarity between output values based on their input points.

For both our emulators,  $\mathcal{E}_B(k)$  and  $\mathcal{E}_{\partial R/\partial \tilde{x}}(k)$ , we employ a Matérn kernel with a smoothness parameter  $\nu = 3/2$ . This kernel

takes the form

$$K_{\nu=3/2}(r) = \sigma_f^2 \left( 1 + \sqrt{3}r \right) \exp \left( -\sqrt{3}r \right), \quad (13)$$

where  $\sigma_f^2$  is the signal variance,  $r$  is the anisotropic Mahalanobis distance between input points  $\mathbf{x}$  and  $\mathbf{x}'$  in  $D$  dimensions, defined as

$$r = \sqrt{\sum_{i=1}^D \left( \frac{x_i - x'_i}{\ell_i} \right)^2}, \quad (14)$$

with  $\ell_i$  giving the characteristic length-scales for each input dimension. The hyperparameters  $\sigma_f^2$  and  $\ell_i$  were optimised by maximising the marginal likelihood of the training data.

The input parameters for both emulators are  $(\omega_b, \omega_c, n_s, \sigma_8, \ln(k))$ . As mentioned before,  $\ln(k)$  is treated as an explicit input parameter, allowing the emulators to be queried at any arbitrary  $k$  value within the trained range, rather than being trained independently at a fixed grid of  $k$ -bins, offering more flexibility.

The training of the GPs for Aletheia was performed using the `scikit-learn` Python library (Pedregosa et al. 2011). The training process involves finding the optimal kernel hyperparameters, and the testing process (on the 50 separate cosmologies) evaluates the predictive accuracy of the emulators and quantifies their uncertainties.

### 3.4 Correction for resolution effects

The prediction of the non-linear power spectrum from our combined emulator,  $\mathcal{E}_P(k)$ , is inherently limited by the resolution of the AletheiaEmu simulations of the primary training, which used  $2048^3$  particles in a  $1500$  Mpc box. At high  $k$ , these predictions exhibit a suppression of power due to the finite particle resolution. To extend the accuracy of Aletheia more deeply into the non-linear regime, we implemented a correction factor based on higher-resolution simulations.

This correction is defined using the AletheiaMass suite of N-body simulations, described in detail in Fiorilli et al. (in prep.). These are a set of pair-fixed runs for a cosmology close to the best-fitting  $\Lambda$ CDM model of Planck Collaboration et al. (2020) that sample the growth of structure over a range of  $\sigma_{12}$  matching that covered by our emulator (0.2 to 1.0). We obtained high-resolution power spectra,  $P_{\text{HR}}(k)$ , by extracting data from the AletheiaMass runs with  $2048^3$  particles in a  $700$  Mpc box. This configuration results in significantly higher mass resolution than the AletheiaEmu training suite.

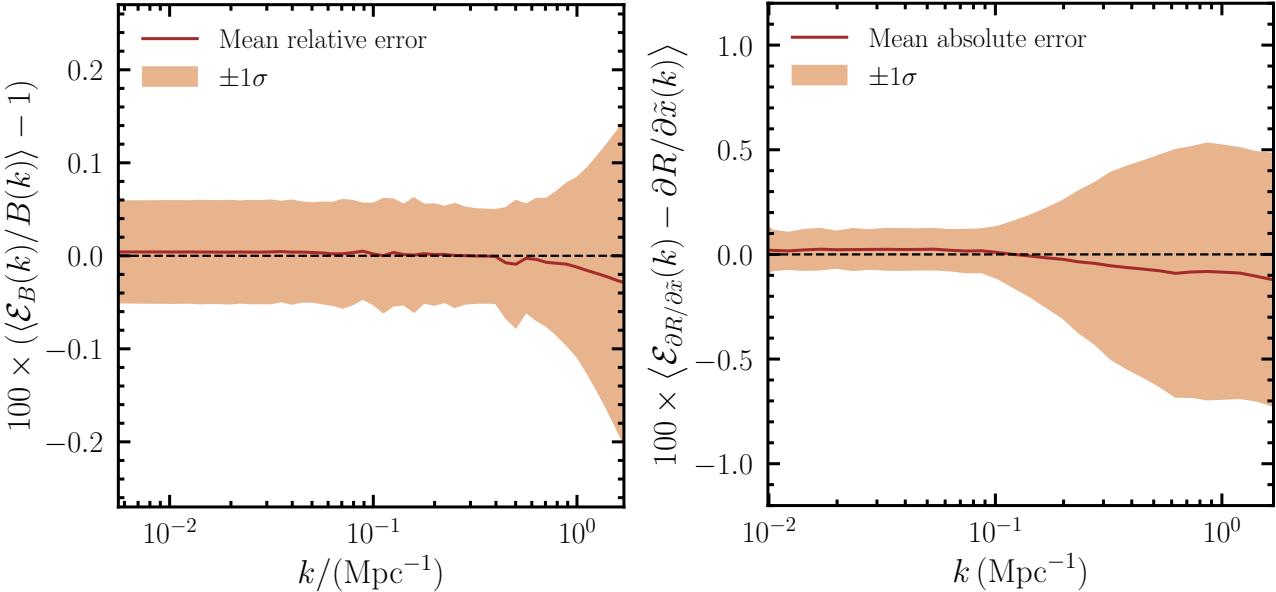
We define the resolution correction factor,  $C(k, \sigma_{12})$ , as the ratio of the uncorrected emulator prediction to the higher-resolution power spectrum,

$$C(k, \sigma_{12}) = \frac{\mathcal{E}_P(k|\Theta_{\text{fid}}, \sigma_{12})}{P_{\text{HR}}(k|\Theta_{\text{fid}}, \sigma_{12})}. \quad (15)$$

We assume that this correction is independent of the shape parameters and depends only on the wavenumber  $k$  and the clustering amplitude  $\sigma_{12}$ . Fig. 5 shows the measured correction ratios  $C(k, \sigma_{12})$  as solid lines, colour-coded by  $\sigma_{12}$ . We model this ratio using a two-dimensional smoothing spline over the  $(k, \sigma_{12})$  plane. The correction is assumed to be unity for high clustering amplitudes, specifically  $\sigma_{12} > 0.83$ . The dashed lines in Fig. 5 illustrate the resulting smoothed spline interpolation,  $C_s(k, \sigma_{12})$ . The final corrected prediction of the emulator,  $\mathcal{E}_{\text{final}}(k)$ , is then given by:

$$\mathcal{E}_{\text{final}}(k) = \frac{\mathcal{E}_P(k)}{C_s(k, \sigma_{12})}. \quad (16)$$

This is the final recipe implemented in the public Aletheia Python package (see Appendix A).



**Figure 6.** Performance of the individual emulator components of **Aletheia**. The left panel shows the relative error of the  $\mathcal{E}_B(k)$  emulator,  $(P_{\text{emu}}/P_{\text{sim}} - 1)$ , with the shaded region indicating the  $1\sigma$  variance across the test set. The right panel displays the absolute error of the  $\mathcal{E}_{dR/d\tilde{x}}(k)$  emulator. Both emulators demonstrate high accuracy across the relevant wavenumber range, with the error of  $\mathcal{E}_B(k)$  being consistently below 0.2% and the absolute error of  $\mathcal{E}_{dR/d\tilde{x}}(k)$  being very small.

#### 4 VALIDATION OF THE EMULATORS

An essential step in the development of any emulator is a rigorous validation of its performance against independent data not used during the training phase. In this section, we present a comprehensive assessment of **Aletheia**, starting with the individual components of our two-stage emulation strategy and then moving on to the validation of the full, combined emulator.

##### 4.1 Validation of individual components

We trained the emulators  $\mathcal{E}_B(k)$  and  $\mathcal{E}_{dR/d\tilde{x}}(k)$  on the sets of simulations described in Section 3.2. To obtain an unbiased assessment of their accuracy, we validated their performance using a dedicated set of 50 testing cosmologies, which were entirely excluded from the training process.

Figure 6 illustrates the performance of both emulator components. The left panel shows the relative error of the emulator  $\mathcal{E}_B(k)$  with respect to the boost factors  $B(k)$  measured in the test simulations. The solid line indicates the mean relative error across all testing cosmologies, and the shaded region shows the corresponding  $1\sigma$  variance. The mean relative error is remarkably small, remaining consistently below 0.03% across the full range of scales considered. The variance of the predictions is stable at the level of 0.05% for  $k \leq 0.8 Mpc^{-1}$ , and gradually increases for larger wavemodes while staying below 0.2%, demonstrating excellent accuracy.

The right panel of Fig. 6 presents the absolute error of the  $\mathcal{E}_{dR/d\tilde{x}}(k)$  emulator. We chose to show the absolute error for this component because the derivative  $\partial R/\partial \tilde{x}(k)$  approaches zero at low  $k$ , making relative errors ill-suited for a clear visual representation. The solid line shows the mean absolute error across the testing cosmologies, and the shaded region indicates the  $1\sigma$  variance. The mean absolute error for  $\mathcal{E}_{dR/d\tilde{x}}(k)$  is generally below 0.007 at high wavenumbers. This absolute error typically corresponds to less than 0.3% of the maximum value of the derivative on those scales.

Both emulators demonstrate good performance, well within the requirements for precision cosmology. It is worth noting that the emulator for the derivative term,  $\mathcal{E}_{dR/d\tilde{x}}(k)$ , does not require the same level of stringent accuracy as  $\mathcal{E}_B(k)$ . This is because the predicted derivative is multiplied by the term  $(\tilde{x} - \tilde{x}_0)$  in equation (12), which is typically of the order of  $10^{-2}$ . Therefore, a small percentage error in  $\partial R/\partial \tilde{x}$  translates into an even smaller fractional error in the final power spectrum prediction. The high performance of these individual components provides strong confidence in the overall accuracy of the **Aletheia** emulator.

##### 4.2 Validation of the full **Aletheia** emulator

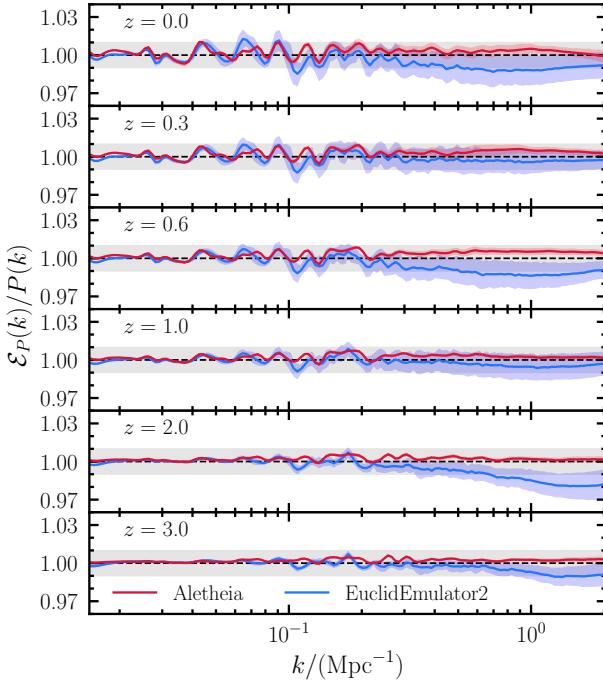
Having validated the individual components of **Aletheia**, we now present an assessment of its performance in predicting the full non-linear matter power spectrum  $P(k)$  for diverse cosmological scenarios. We compare **Aletheia**'s predictions against a set of independent N-body simulations, which we collectively refer to as the **AletheiaDE** test suite. We benchmark its accuracy against the results of **EuclidEmulator2** (Euclid Collaboration et al. 2021). We choose **EuclidEmulator2** as a representative of the state of the art in the field due to its wide adoption, well-documented performance, extended parameter space, and broad redshift range.

For this validation, we selected ten distinct dynamic dark energy cosmological models, including one  $\Lambda$ CDM model ( $w_0 = -1, w_a = 0$ ). These models were chosen such that their parameters fall within the parameter space covered by **EuclidEmulator2**, and their shape parameters ( $\omega_b, \omega_c, n_s$ ) are within the prior ranges of **Aletheia** described in Section 3.2. The full list of cosmological parameters defining these cosmologies is summarised in Table 1.

For each of these cosmologies, we ran two paired and fixed N-body simulations, using initial phases distinct from those employed for the training simulations. These simulations used  $2048^3$  particles in a periodic box of size  $L_{\text{box}} = 1500 \text{ Mpc}$ . We obtained snapshots at six different redshifts:  $z = 3, 2, 1, 0.6, 0.3$ , and 0, spanning the full

**Table 1.** Cosmological parameters for the AletheiaDE simulation set. This table lists the physical baryon ( $\omega_b$ ), and cold dark matter ( $\omega_c$ ) densities, the primordial scalar spectral index ( $n_s$ ), the physical dark energy density ( $\omega_{\text{DE}}$ ) and its equation of state parameters ( $w_0, w_a$ ), the dimensionless Hubble parameter ( $h$ ), and the values of  $\sigma_{12}$  and  $\sigma_{8/h}$  at redshift  $z = 0$ . The first ten rows correspond to cosmologies within the parameter ranges of EuclidEmulator2. The last row represents the best-fit dynamic dark energy cosmology to DESI DR1 BAO data, combined with CMB and SN measurements (DESI Collaboration et al. 2025b).

ID	$\omega_b$	$\omega_c$	$n_s$	$\omega_{\text{DE}}$	$w_0$	$w_a$	$h$	$\sigma_{12}(z = 0)$	$\sigma_{8/h}(z = 0)$
1	0.02213	0.11671	0.9837	0.2865	-1.1118	0.4722	0.6522	0.8144	0.8021
2	0.02219	0.12459	0.9488	0.2308	-0.8340	-0.2718	0.6145	0.9291	0.8774
3	0.02136	0.12821	0.9474	0.3136	-0.9505	0.3855	0.6806	0.8296	0.8416
4	0.02207	0.11665	0.9806	0.3433	-0.7545	-0.3737	0.6943	0.7741	0.7961
5	0.02277	0.12048	0.9628	0.2854	-1.0852	0.3930	0.6547	0.7513	0.7419
6	0.02236	0.12269	0.9461	0.3461	-0.9064	0.2483	0.7008	0.7804	0.8076
7	0.02249	0.11512	0.9699	0.2357	-1.0547	0.1569	0.6110	0.7496	0.7609
8	0.02309	0.11680	0.9645	0.2571	-1.2528	-0.6151	0.6301	0.8087	0.7778
9	0.02167	0.12578	0.9570	0.3833	-0.9794	0.3158	0.7285	0.7464	0.7937
10	0.02229	0.11652	0.9762	0.3100	-1.0000	0.0000	0.6700	0.8078	0.8106
DESI Y1 Best-fit	0.02236	0.12021	0.9648	0.3000	-0.6500	-1.2700	0.6653	0.8230	0.8218



**Figure 7.** Comparison of the performance of the full Aletheia emulator against N-body simulations and EuclidEmulator2. Each panel shows the mean and variance of the ratio  $\mathcal{E}_P(k)/P(k)$  across ten test cosmologies for various redshifts computed using Aletheia (red) and EuclidEmulator2 (blue). Both emulators show excellent agreement with the simulations, with deviations largely driven by cosmic variance. Aletheia exhibits significantly lower variance than EuclidEmulator2, illustrating the higher accuracy enabled by the evolution mapping framework.

redshift range covered by EuclidEmulator2. For each cosmology and redshift, we estimated the matter power spectrum  $P(k)$  following the same methodology as described in Section 3.2.1 and then obtained the corresponding predictions from both Aletheia and EuclidEmulator2.

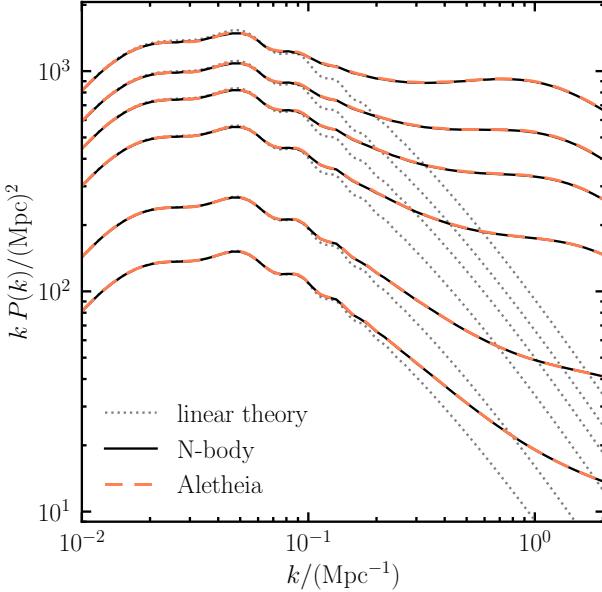
Figure 7 presents the comparison of the performance of both emulators. Each panel corresponds to a different redshift and shows the mean and variance of the ratio  $\mathcal{E}_P(k)/P(k)$  for both Aletheia (red lines) and EuclidEmulator2 (blue lines) across the ten test cos-

mologies. Note that the AletheiaDE simulations used for this test were run at the same resolution as the primary emulator training suite ( $2048^3$  particles in  $1500 \text{ Mpc}$  boxes), therefore, no resolution correction was applied to the Aletheia predictions shown here. The results demonstrate good agreement between both emulators and the N-body simulation results, with overall percent-level differences. Both emulators exhibit similar deviations from the simulation measurements, which are primarily due to cosmic variance associated with the particular initial phases of the simulations. This indicates that both emulators are capturing the underlying physics robustly.

Although the mean ratio curves for both emulators are close to unity, a notable difference can be appreciated in the variance of the predictions. While the variance of the predictions of EuclidEmulator2 is of the order of 1%, Aletheia shows a variance of approximately 0.2%. Although this result is based on a relatively small set of test simulations, this significantly lower variance confirms the power of the evolution mapping approach: the compressed parameter space of the evolution mapping framework results in more accurate and stable predictions compared to those obtained with the standard approach of directly emulating the dependency on all evolution parameters and redshift.

A further advantage of the evolution mapping emulation strategy is its potential for broad generalisation to much wider ranges of redshifts and cosmological parameters, while still maintaining high accuracy. To illustrate this, we performed an additional test using a cosmology explicitly outside the typical parameter space of state-of-the-art emulators. We ran an additional pair of fixed-paired simulations corresponding to the best-fitting dynamical dark energy cosmology obtained from a joint fit to DESI DR1 BAO data, CMB measurements, and SN data (DESI Collaboration et al. 2025b). The parameters for this specific cosmology are listed in the last row of Table 1. This cosmology is outside the limits of most existing emulators (for example, its  $w_a$  value is almost twice the prior range allowed by EuclidEmulator2). Our Aletheia emulator, however, can reproduce accurately the results of this simulation, illustrating the greater generality of the evolution mapping approach.

Figure 8 shows the comparison of the measured  $P(k)$  from these DESI-motivated simulations (black solid lines) with Aletheia's predictions (dashed red lines), showing excellent agreement. The linear theory predictions (dotted grey lines) indicate the amplitude of the non-linear boost  $B(k)$  at each redshift. This agreement can be seen more clearly in Fig. 9, which presents the ratio  $\mathcal{E}_P(k)/P(k)$



**Figure 8.** Non-linear matter power spectrum for the best-fitting dynamic dark energy cosmology of DESI Collaboration et al. (2025b). N-body simulation results (black solid lines) at six redshifts are compared against the predictions of Aletheia (red dashed lines). The linear-theory predictions (grey dotted lines) illustrate the impact of non-linear evolution. The evolution mapping design allows Aletheia to provide reliable predictions for cosmologies outside the parameter space of conventional emulators.

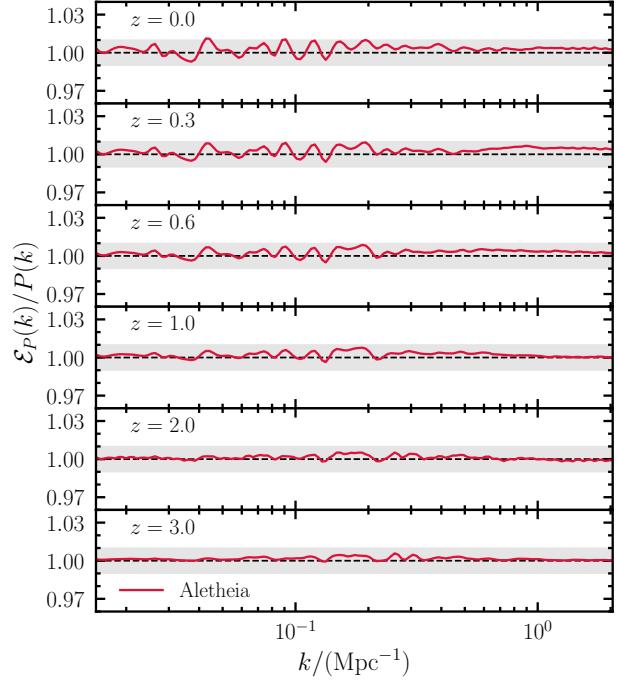
between the predictions of Aletheia and the N-body simulation results. The grey shaded regions mark a  $\pm 1\%$  agreement band at each redshift. The sub-percent agreement on small scales (where cosmic variance is minimal) shows how the evolution mapping framework provides a powerful method for modelling the non-linear matter power spectrum even in cosmological scenarios far from the fiducial  $\Lambda$ CDM model.

The evolution mapping approach provides a significant advantage in handling dark energy models, as the framework does not assume a specific parametrisation for the evolution of  $w_{\text{DE}}(a)$ , such as the standard CPL form (Chevallier & Polarski 2001; Linder 2003). This is possible because the effects of the dark energy only enter the emulator through the background expansion history, which influences  $\sigma_{12}$  and the growth history parameter  $\bar{x}$ . However, this flexibility is limited to models in which dark energy perturbations are negligible and the growth of structure is scale-independent at the linear level. More general models that include dark energy perturbations would require an extension of the current framework.

### 4.3 Validation of the resolution correction

We validated the effectiveness and general applicability of the resolution correction,  $C_s(k, \sigma_{12})$ , using an independent high-resolution simulation of the same challenging DESI-motivated cosmology described in Section 4.2. We ran a paired-fixed simulation of this specific cosmology using  $2048^3$  particles in a  $700 \text{ Mpc}$  box, and obtained high-resolution power spectrum estimates  $P_{\text{HR}}(k)$  at six target redshifts ( $z = 3$  to  $z = 0$ ).

Figure 10 presents the ratio of the predictions of Aletheia to this high-resolution simulation. Each of the six panels shows the ratio obtained without applying the resolution correction ( $\mathcal{E}_P(k)/P_{\text{HR}}(k)$ ,



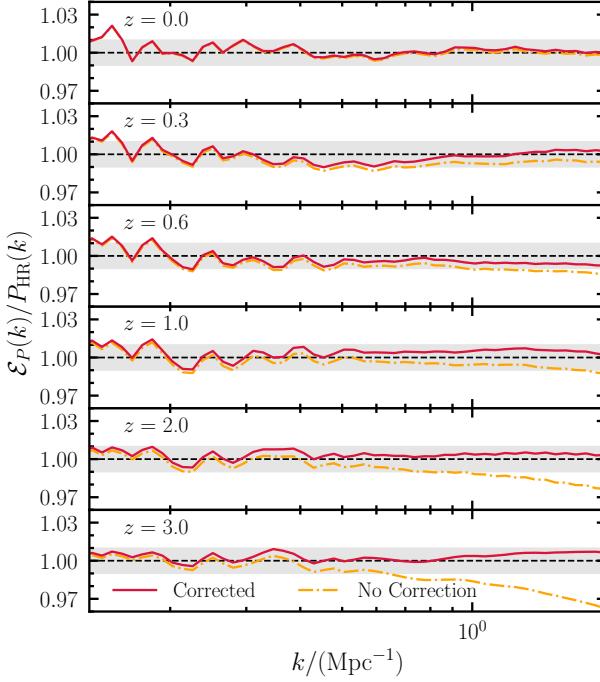
**Figure 9.** Ratio of Aletheia’s prediction to N-body simulation results for the best-fit dynamic dark energy cosmology of DESI Collaboration et al. (2025b). The red solid lines in each panel show  $\mathcal{E}_P(k)/P(k)$  for six different redshifts. The grey shaded regions indicates a  $\pm 1\%$  agreement band. The accuracy achieved demonstrates the validity of the evolution mapping design over a significantly wider cosmological parameter space than conventional emulators.

dot-dashed lines) and the ratio obtained with the correction ( $\mathcal{E}_{\text{final}}(k)/P_{\text{HR}}(k)$ , solid lines).

As these high-resolution simulations have a smaller volume, the impact of cosmic variance on the measured  $P(k)$  is larger, leading to increased scatter and fluctuations in the ratios compared to the tests presented in Section 4.2. Despite this increased uncertainty, these results demonstrate a clear benefit from the correction. At high redshift, corresponding to low clustering amplitudes, the uncorrected results show a systematic deficit of power at high  $k$  with respect to the high-resolution simulation. However, the corrected prediction is consistent with the high-resolution simulation results across the full  $k$ -range, confirming that  $C_s(k, \sigma_{12})$  successfully accounts for the resolution-induced power suppression, even for a cosmology that deviates significantly from the fiducial one used to define the correction.

## 5 CONCLUSIONS

In this paper, we have presented Aletheia, a new emulator for the non-linear matter power spectrum,  $P(k)$  based on the evolution mapping framework proposed by Sánchez et al. (2022). This approach addresses the limitations of traditional emulation techniques, which struggle with wide parameter space coverage and redshift generalisation, by decoupling cosmological dependencies to offer a highly flexible and accurate analysis tool. The core of this framework relies on focusing on  $h$ -independent cosmological parameters, which can be classified into those governing the shape of the linear power spectrum ( $\Theta_s$ ) and those controlling its subsequent growth ( $\Theta_e$ ). The influence of the evolution parameters and redshift can be largely compressed



**Figure 10.** The ratio of the **Aletheia** predictions to the power spectra inferred from high-resolution simulations at varying redshifts. The dot-dashed lines show the results of the uncorrected emulator prediction  $\mathcal{E}_P(k)$ , which underestimates the power at high  $k$  for low clustering amplitudes. The solid lines show the results of the final emulator prediction  $\mathcal{E}_{\text{final}}(k)$ , including the resolution correction described in Section 3.4. The corrected predictions are consistent with  $P_{\text{HR}}(k)$  across all scales.

into a single amplitude parameter,  $\sigma_{12}$ , with small deviations from a universal trend described in terms of the integrated growth history parameter  $\tilde{x}$ . Building upon this principle, **Aletheia** employs a two-stage emulation strategy: a first emulator reproduces the non-linear boost factor  $B(k)$  as a function of  $\Theta_s$  and  $\sigma_{12}$ , and a second emulator accounts for deviations from this ideal mapping with a linear correction dependent on  $\tilde{x}$ . This design allows **Aletheia** to efficiently generate predictions across a broad spectrum of cosmologies and redshifts without the need for dense, explicit sampling of the full parameter space. Our final emulator predictions include a correction to account for resolution effects at high wavenumbers, further extending the accuracy of the predictions more deeply into the non-linear regime.

Our validation against N-body simulations demonstrates the high accuracy and robustness of **Aletheia**. The individual emulator components,  $\mathcal{E}_B(k)$  and  $\mathcal{E}_{\partial R/\partial \tilde{x}}(k)$ , were shown to predict their respective quantities with sub-percent accuracy. When combined, the full **Aletheia** emulator provides excellent agreement with N-body simulations across a range of redshifts and for various dynamical dark energy cosmological models. A direct comparison with the state-of-the-art emulator **EuclidEmulator2** showed that, while both give consistent results, **Aletheia** exhibits a significantly lower variance in its predictions (approximately 0.2% compared to approximately 1% for **EuclidEmulator2**). This smaller variance is a direct consequence of the evolution mapping strategy, which drastically reduces the dimensionality of the problem by leveraging the physical degeneracies in structure growth, leading to more stable and precise predictions from a given training set. **Aletheia**'s design is also more general, allowing it to accurately reproduce the power spectrum for the

best-fit dynamic dark energy cosmology derived from recent DESI, CMB, and SN data (DESI Collaboration et al. 2025b), whose parameters lie well outside the typical training ranges of conventional emulators. Furthermore, our approach can also provide predictions of  $P(k)$  for an arbitrary evolution of  $w_{\text{DE}}(a)$ , as long as these models remain in a regime where dark energy perturbations are negligible and the growth of structure remains scale-independent at the linear level.

The success of **Aletheia** as a proof-of-concept demonstrates the potential of the evolution mapping framework. The emulator is publicly available as a Python package, and a description of its usage is provided in Appendix A. In future work, we plan to extend this setup to emulate other non-linear statistics of the large-scale structure, such as the galaxy-matter cross-power spectrum, the galaxy auto-power spectrum, and higher-order statistics like the bispectrum, further enhancing the capabilities of the **Aletheia** package for cosmological analyses.

Beyond its direct application in building emulators, the evolution mapping framework offers a more physically intuitive understanding of how cosmic structures evolve across different cosmological scenarios. It provides a powerful way to simplify the description of the non-linear evolution of the cosmic density field. The approach fundamentally highlights the importance of expressing theoretical predictions in a manner that is independent of the dimensionless Hubble parameter,  $h$ , simplifying the apparently complex interplay of cosmological parameters in the formation of the LSS of the Universe.

## ACKNOWLEDGEMENTS

We would like to thank Jiamin Hou, Soumadeep Maiti, Andrea Pezzotta, Agne Semenaitė, Mathias Garny, Alex Eggemeier, Martín Crocce, Román Scoccimarro, and Lukas Schwörer for their help and useful discussions. The **AletheiaEmu** simulations were carried out and post-processed on the HPC system Raven of the Max Planck Computing and Data Facility (MPCDF, <https://www.mpcdf.mpg.de>) in Garching, Germany. This research was supported by the Excellence Cluster ORIGINS, which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC-2094 - 390783311. JGJ acknowledges funding by ANID (Beca Doctorado Nacional, Folio 21210846). FR acknowledges support from the ICTP through the Junior Associates Programme 2023-2028. This project has received funding from the European Union's HORIZON-MSCA-2021-SE-01 Research and Innovation programme under the Marie Skłodowska-Curie grant agreement number 101086388 - Project acronym: LACEGAL.

## DATA AVAILABILITY

The **Aletheia** emulator described in this work is publicly available and registered on the Python Package Index. The simulation data underlying this article will be shared on reasonable request to the corresponding authors.

## REFERENCES

- Abbott T. M. C., et al., 2022, *Phys. Rev. D*, **105**, 023520
- Alam S., et al., 2017, *MNRAS*, **470**, 2617–2652
- Alam S., et al., 2021, *Phys. Rev. D*, **103**, 083533
- Anderson L., et al., 2014, *MNRAS*, **441**, 24

- Angulo R. E., Pontzen A., 2016, *MNRAS*, **462**, L1
- Angulo R. E., Zennaro M., Contreras S., Aricó G., Pellejero-Ibañez M., Stücker J., 2021, *MNRAS*,
- Bocquet S., Heitmann K., Habib S., Lawrence E., Uram T., Frontiere N., Pope A., Finkel H., 2020, *ApJ*, **901**, 5
- Brout D., et al., 2022, *ApJ*, **938**, 110
- Chevallier M., Polarski D., 2001, *International Journal of Modern Physics D*, **10**, 213
- Cole S., et al., 2005, *MNRAS*, **362**, 505–534
- Crocce M., Pueblas S., Scoccimarro R., 2006, *MNRAS*, **373**, 369
- Crocce M., Pueblas S., Scoccimarro R., 2012, 2LPTIC: 2nd-order Lagrangian Perturbation Theory Initial Conditions, Astrophysics Source Code Library, record ascl:1201.005
- DESI Collaboration et al., 2025a, *arXiv e-prints*, p. arXiv:2503.14738
- DESI Collaboration et al., 2025b, *J. Cosmology Astropart. Phys.*, **2025**, 021
- DESI Collaboration et al., 2025c, *J. Cosmology Astropart. Phys.*, **2025**, 028
- Dalal R., et al., 2023, *Phys. Rev. D*, **108**, 123519
- DeRose J., et al., 2019, *ApJ*, **875**, 69
- Eggemeier A., Camacho-Quevedo B., Pezzotta A., Crocce M., Scoccimarro R., Sánchez A. G., 2023, *MNRAS*, **519**, 2962
- Eisenstein D. J., et al., 2005, *ApJ*, **633**, 560
- Esposito M., Sánchez A. G., Bel J., Ruiz A. N., 2024, *MNRAS*, **534**, 3906
- Euclid Collaboration et al., 2019, *MNRAS*, **484**, 5509
- Euclid Collaboration et al., 2021, *MNRAS*, **505**, 2840
- Euclid Collaboration et al., 2025, *A&A*, **697**, A1
- Garny M., Taule P., 2021, *J. Cosmology Astropart. Phys.*, **2021**, 020
- Garrison L. H., Eisenstein D. J., Ferrer D., Tinker J. L., Pinto P. A., Weinberg D. H., 2018, *The Astrophysical Journal Supplement Series*, **236**, 43
- Grieb J. N., et al., 2017, *MNRAS*, **467**, 2085
- Hamann J., Hannestad S., Lesgourgues J., Rampf C., Wong Y. Y. Y., 2010, *J. Cosmology Astropart. Phys.*, **2010**, 022
- Hamilton A. J. S., Kumar P., Lu E., Matthews A., 1991, *ApJ*, **374**, L1
- Heitmann K., White M., Wagner C., Habib S., Higdon D., 2010, *ApJ*, **715**, 104
- Heitmann K., et al., 2016, *ApJ*, **820**, 108
- Hockney R. W., Eastwood J. W., 1988, Computer Simulation Using Particles, 1 edn. CRC Press, Boca Raton, FL, doi:10.1201/9780367806934, <https://doi.org/10.1201/9780367806934>
- Ivanov M. M., Simonović M., Zaldarriaga M., 2020, *J. Cosmology Astropart. Phys.*, **2020**, 042
- Li X., et al., 2023, *Phys. Rev. D*, **108**, 123518
- Linder E. V., 2003, *Physical Review Letters*, **90**, 091301
- Louis T., et al., 2025, *arXiv e-prints*, p. arXiv:2503.14452
- McKay M. D., Beckman R. J., Conover W. J., 1979, *Technometrics*, **21**, 239
- Peacock J. A., Dodds S. J., 1996, *MNRAS*, **280**, L19
- Pedregosa F., et al., 2011, *Journal of Machine Learning Research*, **12**, 2825
- Pezzotta A., et al., 2025, *Phys. Rev. D*, **112**, 023520
- Planck Collaboration et al., 2020, *A&A*, **641**, A6
- Rasmussen C. E., Williams C. K. I., 2006, Gaussian Processes for Machine Learning. The MIT Press
- Sánchez A. G., 2020, *Phys. Rev. D*, **102**, 123511
- Sánchez A. G., Baugh C. M., Percival W. J., Peacock J. A., Padilla N. D., Cole S., Frenk C. S., Norberg P., 2006, *MNRAS*, **366**, 189
- Sánchez A. G., et al., 2017, *MNRAS*, **464**, 1493
- Sánchez A. G., Ruiz A. N., Jara J. G., Padilla N. D., 2022, *MNRAS*, **514**, 5673
- Scoccimarro R., Colombi S., Fry J. N., Frieman J. A., Hivon E., Melott A., 1998, *ApJ*, **496**, 586
- Scolnic D. M., et al., 2018, *ApJ*, **859**, 101
- Scolnic D., et al., 2022, *ApJ*, **938**, 113
- Sefusatti E., Crocce M., Scoccimarro R., Couchman H. M. P., 2016, *MNRAS*, **460**, 3624
- Smith R. E., et al., 2003, *MNRAS*, **341**, 1311
- Springel V., Pakmor R., Zier O., Reinecke M., 2021, *MNRAS*, **506**, 2871
- Stein M., 1987, *Technometrics*, **29**, 143151
- Takahashi R., 2008, *Progress of Theoretical Physics*, **120**, 549
- Taruya A., 2016, *Phys. Rev. D*, **94**, 023504
- Tristram M., et al., 2024, *A&A*, **682**, A37
- Tröster T., et al., 2020, *A&A*, **633**, L10
- Wright A. H., et al., 2025, *arXiv e-prints*, p. arXiv:2503.19441
- Yang Y., Bird S., Ho M.-F., 2025, *Phys. Rev. D*, **111**, 083529
- d'Amico G., Gleyzes J., Kokron N., Markovic K., Senatore L., Zhang P., Beutler F., Gil-Marín H., 2020, *J. Cosmology Astropart. Phys.*, **2020**, 005

## APPENDIX A: FUNCTIONALITY OF THE ALETHEIA PACKAGE

The non-linear power spectrum emulator described in this work is publicly available as a Python package named `Aletheia`. This appendix provides a brief guide to its installation and basic usage. For more detailed information, tutorials, and the source code, we refer the reader to the public repository<sup>2</sup> and the full online documentation<sup>3</sup>.

The package is registered on the Python Package Index (PyPI) and can be installed in a Python environment using `pip`, as

```
pip install AletheiaCosmo
```

This command will automatically handle the installation of the required dependencies, which include `numpy`, `scipy`, and `camb`.

The core of the package is the `AletheiaEmu` class. The typical workflow involves creating a cosmology dictionary, initialising the emulator, and then calling its main prediction method. A minimal working example is shown below.

```
import numpy as np
from aletheiacosmo import AletheiaEmu

# 1. Define the cosmology using the helper function.
# This function handles the conversion of parameters
# and ensures the dictionary is correctly formatted
# for the emulator.
cosmo_params = AletheiaEmu.create_cosmo_dict(
    h=0.67,
    omega_b=0.0224,
    omega_c=0.120,
    n_s=0.96,
    A_s=2.1e-9,
    model='LCDM' # Automatically sets w0=-1, wa=0
)

# 2. Initialise the emulator instance.
# This loads the internal GP models
# and correction functions.
emu = AletheiaEmu()

# 3. Define the wavenumbers and redshift for the
# prediction.
k_values = np.logspace(-2, 0.3, 200) # k in 1/Mpc
redshift = 1.0

# 4. Get the non-linear power spectrum.
# The get_pnl method returns a NumPy array
# of P_NL(k).
P_NL = emu.get_pnl(k_values, cosmo_params, redshift)

# The result can now be used for analysis.
```

<sup>2</sup> <https://gitlab.mpcdf.mpg.de/arielsan/aletheia>

<sup>3</sup> <https://aletheia-46606f.pages.mpcdf.de/index.html>

The primary method of the class is `get_pnl(k, cospar, z)`, which takes an array of wavenumbers in units of  $\text{Mpc}^{-1}$ , a cosmology dictionary, and a single redshift, and returns the non-linear power spectrum in units of  $\text{Mpc}^3$ . The package also includes internal validation to ensure that the input cosmology lies within the parameter space where the emulator was trained, raising an error if the boundaries are exceeded.

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