Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **505,** 2840–2869 (2021)
Advance Access publication 2021 May 14

https://doi.org/10.1093/mnras/stab1366

Euclid preparation: IX. EuclidEmulator2 – power spectrum emulation with massive neutrinos and self-consistent dark energy perturbations

Euclid Collaboration:, M. Knabenhans,[1★] J. Stadel,[1] D. Potter,[1] J. Dakin,[2] S. Hannestad,[2] T. Tram,[2] S. Marelli,[3] A. Schneider,[1,4] R. Teyssier,[1] P. Fosalba,[5] S. Andreon,[6] N. Auricchio,[7] C. Baccigalupi,[8,9,10] A. Balaguera-Antolínez,[11,12] M. Baldi,[13,14,15] S. Bardelli,[7] P. Battaglia,[16] R. Bender,[17,18] A. Biviano,[10,19] C. Bodendorf,[18] E. Bozzo,[20] E. Branchini,[21,22,23] M. Brescia,[24] C. Burigana,[15,25,26] R. Cabanac,[27] S. Camera,[28,29,30] V. Capobianco,[30] A. Cappi,[7,31] C. Carbone,[32] J. Carretero,[33] C. S. Carvalho,[34] R. Casas,[5,35] S. Casas,[36] M. Castellano,[23] G. Castignani,[37] S. Cavuoti,[24,38,39] R. Cledassou,[40] C. Colodro-Conde,[12] G. Congedo,[41] C. J. Conselice,[42] L. Conversi,[43,44] Y. Copin,[45] L. Corcione,[30] J. Coupon,[20] H. M. Courtois,[46] A. Da Silva,[47,48] S. de la Torre,[49] D. Di Ferdinando,[15] C. A. J. Duncan,[50] X. Dupac,[44] G. Fabbian,[51] S. Farrens,[36] P. G. Ferreira,[50] F. Finelli,[13,52] M. Frailis,[10] E. Franceschi,[7] S. Galeotta,[10] B. Garilli,[32] C. Giocoli,[7,14,15] G. Gozaliasl,[53,54] J. Graciá-Carpio,[18] F. Grupp,[17,18] L. Guzzo,[6,55,56] W. Holmes,[57] F. Hormuth,[58] H. Israel,[17] K. Jahnke,[59] E. Keihanen,[54] S. Kermiche,[60] C. C. Kirkpatrick,[54] B. Kubik,[61] M. Kunz,[62] H. Kurki-Suonio,[54] S. Ligori,[30] P. B. Lilje,[63] I. Lloro,[64] D. Maino,[32,55,56] O. Marggraf,[65] K. Markovic,[57] N. Martinet,[49] F. Marulli,[7,14,15] R. Massey,[66] N. Mauri,[14,15] S. Maurogordato,[31] E. Medinaceli,[16] M. Meneghetti,[7] B. Metcalf,[14,16] G. Meylan,[37] M. Moresco,[7,14] B. Morin,[36] L. Moscardini,[7,14,15] E. Munari,[10] C. Neissner,[33] S. M. Niemi,[67] C. Padilla,[33] S. Paltani,[20] F. Pasian,[10] L. Patrizii,[15] V. Pettorino,[36] S. Pires,[36] G. Polenta,[68] M. Poncet,[40] F. Raison,[18] A. Renzi,[69,70] J. Rhodes,[57] G. Riccio,[24] E. Romelli,[10] M. Roncarelli,[7,14] R. Saglia,[17,18] A. G. Sánchez,[18] D. Sapone,[71] P. Schneider,[65] V. Scottez,[72] A. Secroun,[60] S. Serrano,[5,35] C. Sirignano,[69,70] G. Sirri,[15] L. Stanco,[69] F. Sureau,[36] P. Tallada Crespí,[73] A. N. Taylor,[41] M. Tenti,[15] I. Tereno,[34,47] R. Toledo-Moreo,[74] F. Torradeflot,[73] L. Valenziano,[7,15] J. Valiviita,[53,54] T. Vassallo,[17] M. Viel,[8,9,10,19] Y. Wang,[75] N. Welikala,[41] L. Whittaker,[76,77] A. Zacchei[10] and E. Zucca[7]

Affiliations are listed at the end of the paper

Accepted 2021 May 2. Received 2021 May 2; in original form 2020 October 23

ABSTRACT

We present a new, updated version of the `EuclidEmulator` (called `EuclidEmulator2`), a fast and accurate predictor for the nonlinear correction of the matter power spectrum. 2 per cent level accurate emulation is now supported in the eight-dimensional parameter space of $w_0 w_a$CDM+$\sum m_\nu$ models between redshift $z = 0$ and $z = 3$ for spatial scales within the range $0.01\,h\,\mathrm{Mpc}^{-1} \leq k \leq 10\,h\,\mathrm{Mpc}^{-1}$. In order to achieve this level of accuracy, we have had to improve the quality of the underlying *N*-body simulations used as training data: (i) we use self-consistent linear evolution of non-dark matter species such as massive neutrinos, photons, dark energy, and the metric field, (ii) we perform the simulations in the so-called *N*-body gauge, which allows one to interpret the results in the framework of general relativity, (iii) we run over 250 high-resolution simulations with 3000^3 particles in boxes of $1(h^{-1}\,\mathrm{Gpc})^3$ volumes based on paired-and-fixed initial conditions, and (iv) we provide a resolution correction that can be applied to emulated results as a post-processing step in order to drastically reduce systematic biases on small scales due to residual resolution effects in the simulations. We find that the inclusion of the dynamical dark energy parameter w_a significantly increases the complexity and expense of creating the emulator. The high fidelity of `EuclidEmulator2` is tested in various comparisons against *N*-body simulations as well as alternative fast predictors such as `HALOFIT`, `HMCode`, and `CosmicEmu`. A blind test is successfully performed against the Euclid Flagship v2.0 simulation. Nonlinear correction factors emulated with `EuclidEmulator2` are accurate at the level of 1 per cent or better for

★ E-mail: mischa.knabenhans@uzh.ch

© 2021 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

$0.01\,h\,\mathrm{Mpc}^{-1} \leq k \leq 10\,h\,\mathrm{Mpc}^{-1}$ and $z \leq 3$ compared to high-resolution dark-matter-only simulations. `EuclidEmulator2` is publicly available at https://github.com/miknab/EuclidEmulator2.

Key words: methods: numerical – methods: statistical – cosmological parameters – large-scale structure of Universe.

1 INTRODUCTION

Ongoing and forthcoming cosmological surveys such as DESI[1] (DESI Collaboration 2016), the Vera C. Rubin observatory (formerly LSST)[2] (LSST Science Collaboration 2009), *Euclid*[3] (Laureijs et al. 2011), and the Roman Space Telescope (formerly WFIRST)[4] (Akeson et al. 2019) have the potential to shed light on dark energy (DE), dark matter (DM), and the neutrino masses. It has been confirmed by Solar and atmospheric neutrino experiments that neutrinos have finite mass (e.g. Valle 2005; Schwetz et al. 2008), yet, attempts to pin down the total mass and the mass hierarchy of the neutrino flavour states have so far not been successful. The ongoing neutrino experiment KATRIN (Weinheimer 2002; Fraenkle 2008; Wolf 2010) has been launched in order to tighten the neutrino mass bounds with particle physics. The equation of state (EoS) parameter w describing DE is also poorly constrained. While the current Λ cold dark matter (ΛCDM) concordance cosmological model with a value of $w = -1$ is favoured, the error bars coming from *Planck* data alone are of order 50 per cent. They shrink to $\sim$10 per cent if more DE-sensitive probes such as cluster counts, weak lensing, and supernovae Type Ia are additionally considered (Planck Collaboration XIII 2015). Past and current surveys have a hard time constraining the DE parameter more accurately: Surveys like *Planck* (Planck Collaboration 2006) probe the cosmic microwave background (CMB) only, which is not very sensitive to DE. On the other hand, several current surveys analysed either only spectroscopic probes (e.g. eBOSS, Dawson et al. 2016) or only photometric probes (e.g. KiDS-1000, Kuijken et al. 2019). In contrast, future experiments such as e.g. DESI and the Vera C. Rubin observatory as well as large-scale hybrid photometric and spectroscopic experiments such as e.g. *Euclid* or the Roman Space Telescope will be able to reduce the error bars on the DE parameters, as the combination of weak lensing with galaxy clustering provides a promising handle on DE phenomena.

A common approach for the inference of cosmological parameters is to use Bayesian techniques. One specific possibility is to sample the likelihood function in a Markov chain Monte Carlo (MCMC) approach, compare the predicted observable (e.g. the power spectrum) to the one actually measured in an observation, and extract the maximum likelihood estimator values for the cosmological parameters. However, this requires a large number of accurate theoretical predictions. For studying the nonlinear regime of cosmic structure formation, *N*-body simulations, still the most accurate tool available today, are numerically too expensive to be used for Bayesian inference and hence there is a high demand for more efficient methods. While halo models for massive neutrinos are investigated by several research groups (see e.g. Massara, Villaescusa-Navarro & Viel 2014 or Hannestad, Upadhye & Wong 2020), surrogate models for *N*-body simulations, so-called emulators, have been shown to be very promising candidates. Several emulators are available already. Examples are `FrankenEmu` (Heitmann et al. 2014), `CosmicEmu` (Heitmann et al. 2009, 2010; Lawrence et al. 2010; Heitmann et al. 2016; Lawrence et al. 2017), the emulators of the Aemulus project (DeRose et al. 2019; McClintock et al. 2019; Zhai et al. 2019), `NGenHalofit` (Smith & Angulo 2019), `EuclidEmulator1` (Euclid Collaboration 2019), the `Dark quest` emulator (Nishimichi et al. 2019), and `BE-HaPPY`(Valcin et al. 2019).

In this paper, we will introduce an update of `EuclidEmulator1`. While `EuclidEmulator1` was able to efficiently estimate the nonlinear correction (NLC) to the matter power spectrum for w_0CDM cosmologies (with the time-variable DE EoS parameter w_a set to 0), `EuclidEmulator2` can handle cosmologies with dynamical DE and massive neutrinos. In addition to a bigger parameter space, `EuclidEmulator2` also pushes the upper limit of the k-range to $k_{\max} \sim 10\,h\,\mathrm{Mpc}^{-1}$. The motivation for this is the same as described in Euclid Collaboration (2019): While clearly the dominant source of uncertainties on such small spatial scales is due to hydrodynamics, galaxy formation, and feedback processes, it is important to have best possible control over the gravity-only physics (in both the dark sector and baryons) in this regime in order to avoid additional (and unnecessary) uncertainties. As we describe in Euclid Collaboration (2019), under certain assumptions it is possible to add baryonic and other corrections due to e.g. gas physics and feedback effects as a subsequent step in the pipeline, necessitating that theoretical precision is maintained in the high-k regime of DM clustering. In Schneider et al. (2020a, b) the authors follow this strategy and emulate the effect due to baryons on weak lensing observables on top of the underlying nonlinear DM physics.

This paper is structured as follows. In Section 2, theoretical aspects regarding massive neutrinos, dynamical DE, and the *N*-body gauge are discussed. In Section 3, we introduce our approach to include massive neutrinos in *N*-body simulations using `PKDGRAV3`. In Section 4, we report on an extensive convergence series that we have performed in order to estimate the uncertainties in the input simulations used for the construction of `EuclidEmulator2`, considering volume and resolution effects. We find systematics in this section whose treatment will then be discussed subsequently in Section 5. In Section 6, we investigate a prototype emulator based on `HALOFIT` data (as we have already done in Euclid Collaboration 2019). In Section 7, insights taken from this prototype emulator are used to construct the fully simulation-based `EuclidEmulator2`, whose performance is ultimately assessed in Section 8. We then gain some insight into parameter degeneracies at the level of the matter power spectrum within the $w_0 w_a$ CDM $+ \sum m_\nu$ models. Our conclusions are presented in Section 10.

2 THEORETICAL BACKGROUND

2.1 Massive neutrinos

In oscillation experiments studying solar and terrestrial neutrinos, compelling evidence has been found that the three flavour eigenstates of neutrinos (e, μ, and τ) can be mixed (Becker-Szendy et al. 1992; Fukuda et al. 1998a,b; Ahmed et al. 2004; DUNE Collaboration 2015). This implies that neutrinos must have finite mass eigenstates, a fact lying outside of the current standard model of particle physics. While those experiments were able to measure the differences between the squared neutrinos masses, they cannot measure the

[1] www.desi.lbl.gov/category/announcements/
[2] https://www.vro.org, www.lsst.org/lsst
[3] sci.esa.int/euclid
[4] http://roman.gsfc.nasa.gov

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

absolute mass scale. Currently, we only have bounds on the sum of the neutrino masses, see e.g. Particle Data Group: Patrignani et al. (2016). On cosmological scales, light neutrinos are very abundant and hence are expected to have a significant effect on large-scale structure. As we are considering neutrinos of masses smaller than 1 eV, they have become non-relativistic only after the electron–nucleon recombination and hence they have imprinted only a small signal into the CMB. However, neutrinos constitute a fraction of the DM in our Universe and hence the DM power spectrum is the key quantity to look at when trying to constrain light neutrino masses.

For the rest of this paper we are restricting ourselves to three light mass eigenstates, i.e. we neglect the possibility of heavy sterile neutrinos. Additionally, we consider only the case of three degenerate neutrino masses, i.e. we neglect that the measured squared differences for their masses is given by (Particle Data Group: Tanabashi et al. 2018)

$$\delta m_{21}^2 = m_2^2 - m_1^2 = \left(7.37^{+0.197}_{-0.146}\right) \times 10^{-5}\mathrm{eV}^2, \tag{1}$$

$$\delta m_{31}^2 = \left|m_3^2 - m_1^2\right| = \left(2.56^{+0.043}_{-0.037}\right) \times 10^{-3}\mathrm{eV}^2, \tag{2}$$

or alternatively

$$\delta m_{23}^2 = \left|m_2^2 - m_3^2\right| = (2.54 \pm 0.04) \times 10^{-3}\mathrm{eV}^2, \tag{3}$$

depending on the considered neutrino hierarchy. This simplification is justified as the difference in the resulting power spectra between a degenerate and a realistic (normal or inverted) mass hierarchy is expected to be well sub-per cent. Neutrinos have a non-vanishing mass while at the same time they have a high velocity dispersion, hence they do not belong to the category of CDM. In this paper, we focus on cosmologies where there are non-zero contributions from both cold and non-cold DM particles. Such a CDM plus massive neutrino (CDM+$\sum m_\nu$)[5] cosmology also serves as the new *Euclid* reference cosmology (see Table 2), in contrast to a more standard pure CDM cosmology.

While neutrinos are still relativistic, their free-streaming length is of the size of the Hubble scale. Only after the transition to the non-relativistic phase, the comoving free-streaming scale of neutrinos starts to shrink. As a result, neutrino perturbations get washed out on scales below the free-streaming scale. Due to gravitational interaction, this also suppresses the clustering of CDM on small enough scales. The strength of the effect for a specific k-mode depends on both redshift and mass of the neutrinos. But even for small (but finite) neutrino masses we expect a several percent suppression signal in the DM power spectrum due to these effects. The main effect is on the background expansion of the Universe: If only massless neutrinos are considered, $\Omega_{\rm DM}$ is identical to $\Omega_{\rm CDM}$. However, in order to maintain spatial flatness, even when a non-vanishing Ω_ν is present, $\Omega_{\rm DM}$ remains unaltered and thus this requires $\Omega_{\rm CDM}$ to be decreased accordingly:

$$\Omega = \Omega_{\rm DE} + \Omega_{\rm DM} + \Omega_{\rm b} + \Omega_{\rm rad} = \Omega_{\rm DE} + (\Omega_{\rm CDM} + \Omega_\nu) + \Omega_{\rm b} + \Omega_{\rm rad}. \tag{4}$$

The CDM density is thus decreased by Ω_ν (while the DM density parameter does not change). As CDM and neutrinos evolve differently over the history of the Universe, this leads to a suppression of the CDM+baryon power spectrum. This power suppression can help us constrain the sum of the neutrino masses (see e.g. Ichiki, Takada & Takahashi 2009; Coulton et al. 2019; Copeland, Taylor & Hall 2020).

Neutrinos at very high redshifts ($z \gtrsim 1000$) are still a relativistic species, and hence would primarily contribute to $\Omega_{\rm rad}$. However, our simulations focus on the nonlinear growth of structure at $z \lesssim 10$, at which time the contribution of massive neutrinos shifts to $\Omega_{\rm DM}$. It is for this reason that we consider Ω_ν as a contribution to $\Omega_{\rm DM}$ in the above. None the less, in our N-body simulations the transition between fully relativistic and non-relativistic neutrinos, including the full distribution function at any given redshift, is accounted for by the `CLASS` Boltzmann solver. All neutrino effects are self-consistently included in our simulations at the linear level, such that this assignment of Ω_ν to $\Omega_{\rm DM}$ is purely a convenient choice of parametrization. For a more detailed discussion of the neutrino treatment in our N-body simulations please see Section 3.4.

2.2 Dynamical dark energy

In the standard ΛCDM cosmology, DE is assumed to be a cosmological constant with a time-independent EoS parameter given by $w \equiv -1$. This implies that $\rho_\Lambda(t) = \text{const}$ which can be seen from the following conservation equation:

$$T^{\nu}_{0;\nu} = \dot{\rho}_\Lambda + 3H(\rho_\Lambda + p) = 0\,, \tag{5}$$

$$\dot{\rho}_\Lambda + 3H\rho_\Lambda(1 + w) = 0\,, \tag{6}$$

$$\dot{\rho}_\Lambda = 0\,, \tag{7}$$

where $T_{\mu\nu}$ is the energy momentum tensor, p the pressure, ρ the density, and the over-dot denotes a derivative w.r.t. cosmic time. Although this value is close to the best-fitting value from supernova surveys (Betoule et al. 2014), DE with a slightly time-dependent EoS parameter is not ruled out by the data currently available. The effects of DE perturbations become relevant only on the largest scales (usually at $k \lesssim 0.1\,h\,\mathrm{Mpc}^{-1}$) as has recently been studied carefully in Dakin et al. (2019b). Nevertheless, the effects of DE on the matter power spectrum can become quite significant, primarily because changes in the DE component have an impact on the scale factor $a(t)$ which in turn affects the power spectrum on all scales.

Here, we shall just briefly recap the key aspects of the theory of time-dependent DE relevant in the context of `EuclidEmulator2`. We shall follow closely the explanations given in Dakin et al. (2019b) where this topic has been reviewed in greater detail. There are two popular ways of describing DE in the setting of an effective theory: the fluid description and the so-called parametrized post-Friedmann (PPF) description (Hu & Sawicki 2007). In the fluid description, DE is considered a fluid with an equation-of-state parameter $w(a)$ (a being the cosmic scale factor) and a constant rest-frame sound speed $c_{\rm s}$. We will adopt the widely used parametrization $w(a) = w_0 + w_a(1 - a)$. As can be seen in equation (2.9) in Dakin et al. (2019b), the Euler equation for a DE fluid features a factor $(1 + w)^{-1}$, leading to divergences for cosmologies with a DE EoS parameter that evolves across $1 + w = 0$ over time. This is a manifestation of the fact that such a DE is gravitationally unstable for the lack of additional internal degrees of freedom (Fang, Hu & Lewis 2008). The case $w = -1.0$ is sometimes referred to as the 'phantom divide' or 'phantom barrier' and models crossing it are called 'phantom-crossing' cosmologies. It follows, unfortunately, that the fluid description of DE is not well suited to describe phantom-crossing cosmologies and yet there is no reason why these models should not be taken into account in our analysis.

[5]In this paper, we use the notation CDM+$\sum m_\nu$ in order to avoid confusion with mixed DM (MDM) models with significant contributions from more exotic warm or even hot DM species.

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

As described in Fang et al. (2008), common approaches to deal with this problem are either to ignore the DE perturbations in an ad hoc manner or to simply turn them off which violates the energy-momentum conservation for non-Λ models and leads to inconsistencies among the Einstein equations. For dynamical DE models it is hence crucial to thoroughly address the issue of phantom-crossing.

The parametrization of minimally coupled DE is not a complete system of equations but rather it requires two closure conditions. The relation between density and pressure fluctuations of DE, giving rise to its EoS parameter and sound speed, serves as one of these closure relations in the fluid description. The PPF formalism on the other hand takes a direct relation between the momentum density of DE and that of DM on large scales to close the system of equations (Fang et al. 2008), thus circumventing the divergence of sound speed at phantom-crossing. This describes the DE momentum perturbations on large scales. To describe them on small scales, an effective sound speed c_Γ is introduced which is related to the scale below which the DE field becomes sufficiently homogeneous compared to the matter field. From an interpolation between these two limits one can obtain an evolution equation for the potential of DE in its rest frame:

$$\partial_\tau \Gamma = \frac{\partial_\tau a}{a}\left[S\left(1+\frac{c_\Gamma^2 k^2}{\mathcal{H}^2}\right)^{-1} - \Gamma\left(1+\frac{c_\Gamma^2 k^2}{\mathcal{H}^2}\right)\right], \tag{8}$$

with $\mathcal{H}$ the conformal Hubble parameter and ∂_τ denoting the derivative w.r.t. conformal time τ. Γ is related to the rest-frame DE density field $\delta\rho_{\rm DE}^{\rm rest}$ via the Poisson equation

$$k^2\Gamma = -4\pi G a^2 \delta\rho_{\rm DE}^{\rm rest}, \tag{9}$$

c_Γ being the effective sound speed and S being defined as

$$S \equiv \frac{4\pi G a^2}{\mathcal{H}}(\rho_{\rm DE} + p_{\rm DE})\frac{\theta_{\rm t}^{\rm N}}{k^2}. \tag{10}$$

In the last equation, in turn, $\theta_{\rm t}^{\rm N}$ denotes the velocity divergence field of all other species than DE in the conformal Newtonian gauge. For a more complete discussion we refer the reader to Dakin et al. (2019b), section 2.1.2.

2.3 General relativity in *N*-body simulations: the *N*-body gauge

Traditionally, the formation of large-scale structure is simulated with Newtonian *N*-body codes. There are two ways to bring the general theory of relativity (GR) into these simulations, in order to study its effects on structure formation. The first option is to replace the Newtonian equations of motion of structure formation by their general relativistic counterpart. However, this is not an easy task, in particular because scientists have optimized *N*-body codes for Newtonian physics over the last decades. The second option is hence to still use these optimized Newtonian *N*-body simulation codes which is allowed under certain circumstances. As has been shown in Chisari & Zaldarriaga (2011), Newtonian simulations predict the clustering properties of DM and galaxies with negligible errors even on very large scales if non-relativistic components or relativistic but non-clustering components are modelled and if a proper set of coordinates (i.e. gauge) is chosen. According to Fidler et al. (2015) and Brandbyge et al. (2016) it suffices to add a relativistic correction to the Newtonian potential ϕ in the Euler and the Poisson equation (for more practical detail see Section 3.4):

$$(\partial_\tau + \mathcal{H})\mathbf{v}_{\rm CDM+b}^{\rm Nb} = -\nabla\phi + \nabla\gamma^{\rm Nb}, \tag{11}$$

$$\nabla^2\phi = 4\pi G a^2 \delta\rho_{\rm tot}^{\rm Nb}, \tag{12}$$

Figure 1. Ratio between the CDM+baryon transfer functions in the *N*-body gauge (as used for the *N*-body simulations in this work) and the synchronous gauge at three different redshifts. At redshift $z = 99$ where the initial conditions are realized, the two gauges lead to transfer functions differing by almost 4 per cent at very large scales. At the $k \sim 0.006\,h\,{\rm Mpc}^{-1}$ corresponding to the fundamental *k*-mode of the simulations that will be used to construct `EuclidEmulator2` (see Section 4), the effect is just about at the 1 per cent level.

where τ is the conformal time and $\delta\rho_{\rm tot}^{\rm Nb}$ is the total density perturbation from all species,

$$\delta\rho_{\rm tot} = \delta\rho_{\rm CDM} + \delta\rho_{\rm b} + \delta\rho_{\rm photon} + \delta\rho_\nu + \delta\rho_{\rm DE}, \tag{13}$$

where we note that $\delta\rho_{\rm DE} \neq 0$ for $w \neq -1$. The GR correction potential $\gamma^{\rm Nb}$ is built from any other gravitating quantity not already accounted for by $\delta\rho_{\rm tot}^{\rm Nb}$, i.e. the momentum density, pressure and shear of photons, neutrinos, and DE. We further parametrize this as

$$\nabla^2\gamma^{\rm Nb} = -4\pi G a^2 \delta\rho_{\rm metric}, \tag{14}$$

where $\delta\rho_{\rm metric}$ is a fictitious density perturbation, the Newtonian gravity of which implements all general relativistic effects of $\nabla^2\gamma^{\rm Nb}$.

Finally, the full effective potential is split into $\phi - \gamma^{\rm Nb} \equiv \phi_{\rm sim} + \phi_{\rm lin}$, with $\nabla^2\phi_{\rm sim}$ formally equal to $4\pi G a^2(\delta\rho_{\rm CDM}^{\rm Nb} + \delta\rho_{\rm b}^{\rm Nb})$ but in the simulation computed through usual *N*-body (particle) techniques, while

$$\nabla^2\phi_{\rm lin} \equiv 4\pi G a^2\left(\delta\rho_{\rm photon}^{\rm Nb} + \delta\rho_\nu^{\rm Nb} + \delta\rho_{\rm DE}^{\rm Nb} + \delta\rho_{\rm metric}\right) \tag{15}$$

is solved on a grid using Fourier techniques and then applied to the particles as a correction force to the main particle gravity. Notice that $\phi_{\rm lin}$ is the object called $\phi_{\rm GR}$ in other publications such as e.g. Tram et al. (2019), Dakin, Hannestad & Tram (2019a), or Dakin et al. (2019b).

Note that the continuity equation is formally not affected by this additional GR potential $\gamma^{\rm Nb}$. This then can be considered writing the full general relativistic evolution equations of the *N*-body simulations in a very special gauge, the so-called *N-body gauge* (indicated by the superscript 'Nb'). Doing so, it is possible to apply a gauge transformation of the output in a post-processing step in order to reobtain general relativistic results in any observationally relevant gauge up to first order. As is well known and can also be seen from Fig. 1 these general relativistic corrections become important at very large scales and high redshifts. Given that for this work we use simulation boxes of edge length of $1\,h^{-1}$ Gpc, the simulation box volume is too small for those corrections to have a strong effect. Yet, we include those minor corrections into our pipeline as they can be taken into account with moderate effort in a pre-processing step (one merely has to choose a gauge).

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

3 COSMOLOGICAL SIMULATIONS

Cosmic emulators are based on training data (also known as the experimental design) generated by *N*-body simulations. Unsurprisingly, the quality of any emulator hence crucially depends on the quality of the training data. We invested great efforts into optimizing the quality of the training set simulations. The construction of `EuclidEmulator2` training data relies mainly on three codes: `PKDGRAV3`[6] (Potter & Stadel 2016; Potter, Stadel & Teyssier 2017), which is the main simulation code, and `CONCEPT`[7] (Dakin 2015) which in turn depends on the Einstein–Boltzmann solver `CLASS`[8] (Blas, Lesgourgues & Tram 2011; Lesgourgues 2011). At this point it shall be emphasized that `CONCEPT` is a full-fledged *N*-body code in its own right with the added trait of very tight integration with `CLASS`. However, for this work we merely use it to compute the transfer functions in the *N*-body gauge seeding the subsequent *N*-body simulations which are performed with `PKDGRAV3`. All three codes had to be optimized in order to be able to fully self-consistently (at the linear level) treat cosmologies with massive neutrinos and dynamical DE. In this section, we shall describe how these codes were optimized to generate a suite of simulations used as the experimental design for `EuclidEmulator2` and thereby put emphasis on the differences with respect to the generation of the experimental design of `EuclidEmulator1`.

3.1 Pipeline overview

We shall give a quick overview over the implemented pipeline in order to facilitate the understanding of the steps involved in generating the training data for the emulator. More detail is given for each step in dedicated sections below.

3.1.1 Pre-computation of cosmological quantities

We have designed a pipeline in which all cosmological background and linear quantities, such as the time dependence of the Hubble parameter $H(z)$, are computed before running any simulation. This is done through `CONCEPT` with `CLASS` in order to take all relevant physics into account (e.g. this approach allows for a fully relativistic treatment). In this step, transfer functions are computed to linear level at many different redshifts for all species contributing to $\phi_{\rm lin}$, which are later used to provide the small corrective force contributions in the *N*-body simulations. In the case of `EuclidEmulator1`, however, we only computed the transfer function at $z = 0$, which we later used for initial condition (IC) generation via the usual back-scaling approach, and used an analytical form for $H(z)$ to compute the background evolution.

3.1.2 Gauge transformation

This step was entirely missing in `EuclidEmulator1` where we used the transfer function computed in synchronous gauge to set up the ICs. `CLASS`, as well as `CAMB`, compute quantities either in the synchronous or in the conformal Newtonian gauge. However, in order to make Newtonian *N*-body codes consistent with GR, their input transfer functions have to be mapped to the *N*-body gauge described in Section 2.3. This transformation is performed by the `CONCEPT` code. The results are then stored into look-up tables inside HDF5 files that can be queried by the *N*-body code. Notice that `CONCEPT` allows to combine this step with the pre-computation of the cosmological quantities described in the previous paragraph.

3.1.3 N-body simulation

Both the initial condition generation and the actual *N*-body simulations are performed with `PKDGRAV3`. Paired-and-fixed (P+F) (Angulo & Pontzen 2016) first-order Lagrangian perturbation theory (1LPT) initial conditions are set up at redshift $z = 99$ (see also Section 4 for a more in-depth discussion). The nonlinear evolution of DM particles is computed with a binary tree-based fast multipole method (FMM). To these tree forces we add a particle-mesh (PM) field for the fluctuations due to massive neutrinos, photons, DE, and the metric field in order to study their effect on cosmological structure growth at linear level (for simulations treating neutrinos fully nonlinearly see e.g. Banerjee et al. 2018; Bird et al. 2018). This aspect is one of the key differences between `EuclidEmulator1` and `EuclidEmulator2`.

3.1.4 Post-processing

From the simulations we obtain the power spectra for each realization of the fixed-IC simulation pair. In order to get the final P+F power spectrum we compute the average of those individual realizations.[9] As ultimately we are primarily interested in the NLC, we compute it w.r.t. the linear theory power spectrum (in our case computed by `CLASS`) at each redshift. Explicitly, the NLC $B(k, z)$ is the dimensionless quantity defined through the relation

$$P_{\rm nl}(k, z) = P_{\rm LinTh}(k, z)B(k, z)\,, \tag{16}$$

where the subscript nl stands for 'nonlinear' and the subscript LinTh for 'linear theory'. Notice that the defining equation for the NLC, equation (16), looks slightly differently compared to its counterpart, equation (4), in Euclid Collaboration (2019).[10] This is a consequence of our new time-stepping procedure (explained in Section 3.4), which guarantees that $B(k, z) = P_{\rm nl}(k, z)/P_{\rm LinTh}(k, z)$ approaches 1 for small k at all redshifts.

We also compute an additional factor that corrects for power suppression at small scales caused by resolution effects in the simulations. This factor we shall refer to as the resolution correction factor (RCF) and it will be discussed in Section 5.

3.1.5 Emulator construction

The obtained data matrix is then principal component analysed. The emulator then predicts the vector of principal component weights.

[6] http://pkdgrav.org
[7] https://github.com/jmd-dk/concept/
[8] http://class-code.net

[9] The P+F approach (Angulo & Pontzen 2016) and its application in the context of cosmic emulation have already been described in Euclid Collaboration (2019).
[10] A major shortcoming of the IC generation and time-stepping methods employed in Euclid Collaboration (2019) is that the ratio $P_{\rm nl}(k, z)/P_{\rm LinTh}(k, z)$ would not approach unity at large scales (except for the one single redshift at which the transfer functions were computed, usually $z = 0$). To deal with this challenge, for `EuclidEmulator1` we computed the NLC via division by the IC power spectrum of the simulation.

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

3.2 Pre-processing with CO*N*CEPT: transfer functions

For generation of particle initial conditions inside PKDGRAV3, matter density and velocity transfer functions $\delta_{\rm CDM+b}(a_{\rm ini}, k)$ and $\theta_{\rm CDM+b}(a_{\rm ini}, k)$ are required, where $a_{\rm ini}$ is the scale factor at the start of the simulation. As we seek to carry out the simulation in *N*-body gauge, this is also the gauge of these transfer functions. Note that, due to the contributions from species other than CDM, we cannot obtain accurate particle velocities from only $\delta_{\rm CDM+b}$.

To keep the simulation in *N*-body gauge, we need to repeatedly apply the linear GR correction force $-\nabla\phi_{\rm lin}$. From (15) we see that we are thus in need of $\delta\rho_{\rm photon}(a, k)$, $\delta\rho_{\nu}(a, k)$, $\delta\rho_{\rm DE}(a, k)$, and $\delta\rho_{\rm metric}(a, k)$. Obtaining these *N*-body gauge transfer functions, the CO*N*CEPT code has been run in advance. Internally, CO*N*CEPT takes the synchronous gauge output from CLASS, converts it to *N*-body gauge and uses it for both particle initial conditions and GR corrections. The relevant gauge transformations are

$$\delta\rho_{\alpha}^{\rm Nb} = \delta\rho_{\alpha}^{\rm s} + 3\mathcal{H}(1+w_{\alpha})\bar{\rho}_{\alpha}\frac{\theta_{\rm tot}^{\rm s}}{k^2}, \tag{17}$$

$$\theta_{\alpha}^{\rm Nb} = \theta_{\alpha}^{\rm s} + \partial_{\tau}\left(\frac{h_{\rm mp}}{2} - 3\mathcal{H}\frac{\theta_{\rm tot}^{\rm s}}{k^2}\right), \tag{18}$$

where superscripts 'Nb' and 's' indicate the *N*-body and synchronous gauge, respectively, α labels the species, $h_{\rm mp}$ is the trace of the metric perturbation in synchronous gauge, and $\theta_{\rm tot}$ is the total velocity divergence of all species.

Though more complicated, the $\delta\rho_{\rm metric}(a, k)$ transfer function is similarly constructed from various CLASS outputs, some of which are only available in the CLASS version that ships with CO*N*CEPT (see Dakin et al. 2019a for more details). We also note that the computation of the DE pressure perturbation within CLASS, needed for $\delta\rho_{\rm metric}$, is wrong in the standard version of CLASS, but fixed in the version shipping with CO*N*CEPT (see Dakin et al. 2019b for details).

For use with PKDGRAV3, the so-called CLASS-utility was added to CO*N*CEPT, which saves the *N*-body gauge transfer functions to an external HDF5 file, which is then read in by PKDGRAV3. In the HDF5 all requested transfer functions (i.e. on top of the CDM+baryon transfer function also that of the photons, the massive neutrinos, the DE, and the metric perturbations) are stored on a global (a, k) grid, where special attention must be given to the interpolations performed within this grid in order to achieve the required precision. Various background quantities, for example $H(a)$, that are computed by CLASS are also stored in the HDF5 and used by PKDGRAV3, as these generally are non-trivial to compute in the presence of massive neutrinos.

3.3 Initial condition generation

In order to get the initial conditions for the particle positions, we generate a regular grid of particles which we displace from their grid points using first-order Lagrangian perturbation theory. We imprint an initial power spectrum $P_{\rm ini}$ with the pairing-and-fixing technique (Angulo & Pontzen 2016), i.e.

$$P(|\delta_{i,\rm lin}|, \theta_i) = \frac{1}{2\pi}\delta_{\rm D}(|\delta_{i,\rm lin}| - \sqrt{P_{\rm ini}}), \tag{19}$$

with $\delta_{\rm D}$ being the Dirac delta function and the index i labelling the Fourier modes (for more info about how we use the pairing-and-fixing technique to generate initial conditions, please refer to Euclid Collaboration 2019). The initial power spectrum is computed based on the CDM+baryon overdensity field $\delta_{\rm m}$ pre-computed by CLASS+CO*N*CEPT at $z_{\rm ini} = 99$. While PKDGRAV3 has no further use for $\delta_{\rm m}(z < z_{\rm ini})$ we still tabulate these, as they are used when computing the NLC factors. We then use

$$P_{\rm m,\,ini}(k, z_{\rm ini}) = \zeta^2(k)\delta_{\rm m}^2(k, z_{\rm ini}), \tag{20}$$

where the ζ-function is defined as

$$\zeta(k) = \pi\sqrt{2A_{\rm s}}k^{-3/2}\left(\frac{k}{k_*}\right)^{\frac{n_{\rm s}-1}{2}}\exp\left[\frac{\alpha_{\rm s}}{4}\ln\left(\frac{k}{k_*}\right)^2\right]. \tag{21}$$

Here, $A_{\rm s}$ is the spectral amplitude, $n_{\rm s}$ is the spectral index, and the running $\alpha_{\rm s} = 0$ in all cases. The pivot scale is set to its standard *Planck* value, $k_* = 0.05\,{\rm Mpc}^{-1}$. The resulting initial power spectrum is used to displace the particles from their regular grid points.

3.4 Linear and nonlinear evolution

Once the initial condition is generated, the DM particles are evolved by the tree code PKDGRAV3 using FMM and a multi-time-stepping approach. For further technical details about the gravity evolution of DM in PKDGRAV3, we refer to Potter et al. (2017).

What is newly introduced in the version of PKDGRAV3 that is used for this work (which is also new compared to the version used to construct EuclidEmulator1) is the interaction with massive neutrinos as well as other linearly evolved species, namely photons, DE, and the metric field. For this, PKDGRAV3 fetches the linearly evolved perturbations of these species from the pre-computed CLASS transfer functions (stored into lookup tables), then computes and applies the corresponding gravitational kick from these linear species to the nonlinear matter particles. This is done at every base time-step (and not on the adaptive sub time-steps) by solving the Poisson equation (equation 15) in Fourier space for the linear potential,

$$\phi_{\rm lin}(a, \boldsymbol{k}) = -\frac{4\pi G a^2}{k^2}\delta\rho_{\rm lin}^{\rm Nb}(a, \boldsymbol{k}), \tag{22}$$

with $\delta\rho_{\rm lin}^{\rm Nb} \equiv \delta\rho_{\rm photon}^{\rm Nb} + \delta\rho_{\nu}^{\rm Nb} + \delta\rho_{\rm DE}^{\rm Nb} + \delta\rho_{\rm metric}$. This linear potential is averaged over the time-step i and realized on a grid, as

$$\langle\phi_{\rm lin}^{\boldsymbol{k}}\rangle_i = -\frac{4\pi G}{k^2}\frac{1}{t_{i+1}-t_i}\int_{t_i}^{t_{i+1}}{\rm d}t\, a^2\delta\rho_{\rm lin}^{\rm Nb}(a, \boldsymbol{k}), \tag{23}$$

where t is cosmic time. Cubic interpolation in $\log k$ and $\log a$ is used to interpolate the tabulated values on to the grid and to evaluate the above integral. The forces from the realized potential $\langle\phi_{\rm lin}^{\boldsymbol{k}}\rangle_i$ are then applied to the particles using standard particle-mesh methods, specifically through differentiation in Fourier space along all three dimensions (requiring two further scalar grids) and separate inverse Fourier transforms, followed by PCS interpolation from the grid points to the particle locations. This added particle-mesh interaction then provides an additional gravity source term for the particles, complementing the standard particle–particle interaction, such that the DM field is evolved taking the linear species into account.

During the early Universe, it is important to capture the effects of the high-frequency oscillations in the linear fields, particularly in the metric field, otherwise we would see a slight offset of the power spectrum at linear scales. The linear evolution may only be done when the simulation is 'time synchronized' which happens at the start of a base time-step. One approach would be to take sufficient base time-steps to capture this effect at high redshift, but it would be computationally wasteful at lower redshifts. Instead we adopted the following scheme. Each *N*-body simulation was started at redshift $z_{\rm ini} = 99$ and evolved in 60 base time-steps to $z_{\rm intrm} = 10$. From $z_{\rm intrm}$ we continued each simulation down to $z_{\rm fin} = 0$ in another 100 base time-steps. Note that particles drop to much smaller sub time-steps when

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

calculating the particle–particle forces; this being done adaptively based on the acceleration of the particle ($\Delta t_i < \eta\sqrt{\epsilon/|\boldsymbol{a}_i|}$, with $\eta = 0.2$ and softening $\epsilon = 0.02\,\ell$).[11] For a more detailed description of the adaptive time-stepping algorithm see Potter et al. (2017).

By employing this approach we achieve agreement between the linear evolution of all particle species and CLASS over the entire simulation starting from $z_{\rm ini}$ all the way to $z = 0$. A comparison of the neutrino implementation in PKDGRAV3 to CONCEPT is presented in Tram et al. (2019). Further, activities to compare a larger number of cosmological simulation codes with implemented neutrinos (including gevolution Adamek et al. 2016; Adamek, Durrer & Kunz 2017, Gadget Springel et al. 2020, RAMSES Teyssier 2002, and PKDGRAV3 Potter et al. 2017 as well as approximate methods such as PINOCCHIO Rizzo et al. 2017) are currently ongoing in *Euclid*.

4 CONVERGENCE TESTS

We have performed extensive convergence tests for the power spectrum and the NLC in different dimensions. These convergence test results serve to update those presented in Euclid Collaboration (2019).

4.1 Volume

We start by determining the minimal volume required to achieve convergence at the 1 per cent level. We compare a series of box size of edge lengths $L \in \{512, 1024, 2048, 4096\}\,h^{-1}$ Mpc to a reference volume $V = L^3 = (8192\,h^{-1}\,\mathrm{Mpc})^3$. In this process we fix the resolution parameter to $\ell^{-1} \equiv N/L = 1/3\,h\,\mathrm{Mpc}^{-1}$, where in this paper we use N for the number of particles per dimension.

As our new pipeline (described in Section 3) allows us to recover the linear scales very accurately at all redshifts, we could try to actually perform the convergence comparison directly against linear theory itself. However, as for simulation box sizes $L < 2048\,h^{-1}$ Mpc even the smallest k modes are already slightly nonlinear, the finite volume effects overlap with the pre-virialization dip such that particularly at low redshifts no clear conclusion can be drawn.

For this reason, we performed additionally the volume convergence test based on the NLC factor (i.e. we compare to the NLC factor of the simulation in the $L = 8192\,h^{-1}$ Mpc box). The result is shown in Fig. 2. In this test the pre-virialization dip is cancelled out such that we are only left with the finite volume effects. We find that a simulation box of $L = 1024\,h^{-1}$ Mpc is converged at the level of $\sim$1 per cent (the cosmic variance of only a few individual mildly nonlinear k modes exceeds the 1 per cent limit). Notice that this result is unchanged if one would perform the comparison to the largest box at the power spectrum level. As the NLC is computed with respect to linear theory, dividing two NLCs by each other leads to cancellation of the linear theory out of the expression such that one is left with a ratio of power spectra only:

$$\frac{B_L}{B_{L=8192}} = \frac{P_L}{P_{\rm LinTh}}\frac{P_{\rm LinTh}}{P_{L=8192}} = \frac{P_L}{P_{L=8192}}\,. \tag{24}$$

[11]This is the default adaptive time-stepping algorithm with default simulation accuracy parameter values. In Schneider et al. (2016), Garrison, Eisenstein & Pinto (2019), and Springel et al. (2020) (mainly in subsection 8.3) it is shown that with these parameters very good agreement with other cosmological simulation codes such as Gadget-4, ABACUS (Garrison et al. 2016; Garrison et al. 2018) and RAMSES (Teyssier 2002) is achieved.

Based on this result we choose the simulation box side length for our simulations to be $L = 1\,h^{-1}$ Gpc. This finding agrees exactly with Heitmann et al. (2010) and also broadly with the Euclid Collaboration (2019). A similar study was also conducted in Schneider et al. (2016) but with a reference simulation computed in a much smaller volume of only $V = (1024\,h^{-1}\,\mathrm{Mpc})^3$.

4.2 Resolution

For the investigation of the mass resolution required to get convergence of the power spectrum at a desired level, we are interested in the minimal value of the resolution parameter ℓ^{-1} of N-body simulations (for a similar analysis see e.g. Heitmann et al. 2010). In Euclid Collaboration (2019) we claimed that the power spectra are converged at the level of 1 per cent up to $k \sim 5\,h\,\mathrm{Mpc}^{-1}$ for $\ell^{-1} \geq 1.6\,h\,\mathrm{Mpc}^{-1}$. We state clearly that this statement was overly optimistic. We underestimated the minimal ℓ^{-1} because we compared to a simulation with $\ell^{-1} = 4\,h\,\mathrm{Mpc}^{-1}$ which at that time was the best we could do. For this work we were able to double the resolution parameter of our reference simulation to $\ell^{-1} = 8\,h\,\mathrm{Mpc}^{-1}$. In turn, this increases our current estimate of the minimal ℓ^{-1} value required to achieve convergence at the 1 per cent level.

From Fig. 3 one clearly observes that simulations with $\ell^{-1} \geq 4\,h\,\mathrm{Mpc}^{-1}$ are required to be converged at the 1 per cent level at $k = 10\,h\,\mathrm{Mpc}^{-1}$ with respect to the $\ell^{-1} = 8\,h\,\mathrm{Mpc}^{-1}$ simulation at $z = 0$ and the resolution needs to be even higher if 1 per cent-convergence at higher redshifts is required. Given the minimal volume found in Section 4.1, such high ℓ^{-1} values imply prohibitively large particle numbers for our simulations. Taking our computational budget into account, we decided to run simulations with $\ell^{-1} = 3\,h\,\mathrm{Mpc}^{-1}$ thereby doubling the resolution parameter compared to EuclidEmulator1. This means that our simulations, according to the currently best estimate available, are converged at redshift $z = 0$ up to $k_{2\ \mathrm{per\ cent}} = 9.42\,h\,\mathrm{Mpc}^{-1}$ at the level of 2 per cent and up to $k_{1\ \mathrm{per\ cent}} \sim 5\,h\,\mathrm{Mpc}^{-1}$ at the level of 1 per cent. Further results showing how the values for $k_{1\ \mathrm{per\ cent}}$ and $k_{2\ \mathrm{per\ cent}}$ evolve with redshift are shown in Table 1.

Notice that the suppression of power due to low mass resolutions is very systematic. We investigate this further in Section 5.

Further we summarize that a compromise between the requirements estimated from the convergence tests and our computational budget leads to the following specifications for our N-body simulations. We evolve 3000^3 particles in boxes with a volume of $1\,h^{-3}\,\mathrm{Gpc}^3$.

4.3 Paired-and-fixed versus Gaussian random field initial conditions

We also re-evaluate the comparison between P+F simulations and simulations based on Gaussian random field (GRF) initial conditions. For this comparison we ran one pair of fixed amplitude simulations and an ensemble of 50 different GRF simulations. Over all redshifts of interest ($z \leq 3$), P+F simulations approximate the GRF ensemble average very well on all but the largest scales. However, on all scales the deviation of the P+F power spectrum w.r.t. the GRF ensemble average lies well within the bound set by $\max(\sigma_{\rm GRF}, 1\ \mathrm{per\ cent})$. This confirms the finding of Angulo & Pontzen (2016). It needs to be taken into account, though, that in that publication an ensemble of 300 GRF simulations (i.e. six times larger than our ensemble) has been used for comparison. Based on this exploration and the results found already in Euclid Collaboration (2019) we again use the pairing-and-fixing approach to efficiently reduce cosmic variance.

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Figure 2. Convergence test results for volume of the NLC factor B(k,z) of simulations with $L \in \{512, 1024, 2048, 4096, 8192\}\,h^{-1}$ Mpc. We show results for two different redshifts. Apart from individual k modes on mildly nonlinear scales, simulation boxes with $L = 1024\,h^{-1}$ Mpc are converged at the 1 per cent level.

4.4 Quantities not investigated in this convergence series

Quantities we have not tested in this convergence series are the starting redshift, the main time-stepping parameter, and the resolution of the PM-grid for the linear species. As we describe below, the softening is related to the mass resolution but its convergence was not tested independently either. We set the softening parameter to default values of `PKDGRAV3`: $\epsilon = 1/(50N)$ and the time-stepping parameter $\eta = 0.20$. The number of PM-grid cells for the linear species has been conservatively chosen to be a quarter of the CDM+b particle number N per dimension, i.e. we use $N_{\rm lin} = 750$ PM-grid cells per dimension. The ratio $N/N_{\rm lin} = 4$ has already been used in the generation of the Euclid Flagship v2.0 Simulation (EFS2) simulation, where it has been proved to be a more than adequate choice. We also follow the EFS2 simulation for the choice of the initial redshift, $z_{\rm ini} = 99$. Convergence of these quantities has been studied in Schneider et al. (2016). Further, we have not investigated 2LPT second-order Lagrangian perturbation theory (2LPT) ICs for this work but use 1LPT ICs to set up our simulations. While 2LPT is expected to improve the resolution convergence results presented above, `PKDGRAV3` does not yet support 2LPT for multiple fluids.

4.5 Results of this convergence analysis

We have identified that one requires simulations with 3000^3 particles and a resolution parameter of $\ell^{-1} = 3\,h\,{\rm Mpc}^{-1}$ (corresponding to a minimally resolved mass of $\sim 3.3 \times 10^9\ {\rm M}_\odot/h$ and a Nyquist frequency of $k_{\rm max} \sim 9.4\,h\,{\rm Mpc}^{-1}$) in order to achieve satisfactory accuracy on both large and small scales. A simulation of this size and resolution takes a bit more than 2000 node hours (on GPU accelerated nodes[12]).

5 RESOLUTION CORRECTION

In Fig. 3 in Section 4.2 we show that a too low mass resolution leads to a suppression of power on small scales where the amplitude of this suppression grows both with a growing ratio $\ell^{-1}_{\rm lowRes}/\ell^{-1}_{\rm highRes}$ and as the redshift z increases. However, as this effect is very systematic it is possible to correct for it in a post-processing step by compensating the suppression with an RCF.

5.1 Cosmology dependence of the resolution-induced power suppression

In general it has to be assumed that the precise shape of the resolution-induced power suppression (and equivalently that of the RCF) depends on cosmology. To test this statement we compute a series of simulations for 20 different cosmologies. We choose these cosmologies to be a subset of the experimental design (ED) of `EuclidEmulator2`. Of course, running multiple high-resolution reference runs in large boxes is too expensive, so we perform our test in a smaller box. We run these simulations in boxes of $L = 128\,h^{-1}$ Mpc side length with $N^3 = 384^3$ and $N^3 = 1024^3$ particles, corresponding to resolution parameters of $\ell^{-1} = 3\,h\,{\rm Mpc}^{-1}$ and

[12]The simulations were run on the Piz Daint supercomputer at the Swiss National Scientific Supercomputing Centre (CSCS).

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Figure 3. Convergence test results for mass resolution of full nonlinear power spectra showing results for simulations with $\ell^{-1} \in \{1, 2, 3, 4, 8\}\, h\,\mathrm{Mpc}^{-1}$. We show results for two different redshifts. The faint, dark green vertical lines correspond to the k values for which the dark green, solid curves (corresponding to the $N1536L512$-case) deviates more than 1 per cent (dashed lines) and 2 per cent (dotted lines), respectively, from the reference $N4096L512$ ($\ell^{-1} = 8\, h\,\mathrm{Mpc}^{-1}$) simulation. At $z = 2.76$, we find that the $\ell^{-1} = 3\, h\,\mathrm{Mpc}^{-1}$ simulation is converged within 1 per cent up to $k \sim 2\, h\,\mathrm{Mpc}^{-1}$ and within 2 per cent up to $k \sim 3.5\, h\,\mathrm{Mpc}^{-1}$. At $z = 0$, however, the same simulation is converged within 2 per cent all the way up to $k = k_{\mathrm{max}} \sim 10\, h\,\mathrm{Mpc}^{-1}$ and even stays within 1 per cent from the reference up to $k \sim 5\, h\,\mathrm{Mpc}^{-1}$.

Table 1. Dependency on the redshift of the k modes at which the $\ell^{-1} = 3\, h\,\mathrm{Mpc}^{-1}$ simulation are converged at 1 and 2 per cent, respectively. For redshifts 0, 0.5, and 1 the $k_{2\,\mathrm{per\,cent}}$ values correspond to the highest k mode obtained from the simulations and hence should be understood as lower bounds.

	$k_{1\,\mathrm{per\,cent}}$ [$h\,\mathrm{Mpc}^{-1}$]	$k_{2\,\mathrm{per\,cent}}$ [$h\,\mathrm{Mpc}^{-1}$]
$z = 0$	4.87	≥9.42
$z = 0.5$	4.36	≥9.42
$z = 1$	3.99	≥9.42
$z = 2.76$	1.97	3.57
$z = 10$	1.34	1.99

$\ell^{-1} = 8\, h\,\mathrm{Mpc}^{-1}$, respectively. The RCF is then simply defined as:

$$f_{\mathrm{res}}^{3\to 8}(k, z; \boldsymbol{c}) = \frac{P^{\ell^{-1}=3\,h\,\mathrm{Mpc}^{-1}}(k, z; \boldsymbol{c})}{P^{\ell^{-1}=8\,h\,\mathrm{Mpc}^{-1}}(k, z; \boldsymbol{c})}\,, \tag{25}$$

where $\boldsymbol{c}$ denotes a specific cosmology. We thus compare the same resolutions as we do in the case of the dark green and yellow curves in Fig. 3. We show the results of this analysis in Fig. 4. In this figure, the two upper panels correspond to the lower panels in Fig. 3. The variability in the ratio of power spectra due to variability in the cosmology is then shown in the lower panels of Fig. 4. As is clearly visible, the cosmology dependence of this ratio is weak, particularly at very low redshifts. Also at high redshifts, the 1σ-standard deviation is considerably smaller than the biases at the same k modes for all 20 tested cosmologies. Accordingly, we may simplify

$$f_{\mathrm{res}}^{3\to 8}(k, z; \boldsymbol{c}) \approx f_{\mathrm{res}}^{3\to 8}(\boldsymbol{c}^*; k, z) := f_{\mathrm{res}}^{3\to 8}(k, z)\,, \tag{26}$$

where $\boldsymbol{c}^*$ denotes any reasonable cosmology not too different from the cosmologies for which low resolution simulations are run. In our case the Euclid Reference Cosmology, the *Planck* 2015 best-fitting cosmology, or the central cosmology of the parameter box defined below in Table 2 would all be viable choices for $\boldsymbol{c}^*$.

From this we conclude that a cosmology-independent correction factor, although introducing a new source of uncertainty, improves the power spectrum measurement of the N-body simulation by a few per cent at high k. Of course, in future work this should be improved even further by emulating the cosmology dependence of the RCF. This would greatly reduce the newly introduced uncertainty while still mostly removing the bias.

We do compute $f_{\mathrm{res}}^{3\to 8}(k, z)$ at the Euclid Reference Cosmology as defined in Table 2. The RCF curve is shown in Fig. 4 for different redshifts covering the entire redshift range of interest. Notice that values $f_{\mathrm{res}}^{3\to 8}$ for $k < 2\pi/(128\, h^{-1}\,\mathrm{Mpc})$ are set to unity (see discussion below in Section 5.2).

5.2 Dependence on simulation box size

For this approach to be a practical strategy, the RCF must not depend strongly on simulation box sizes. Otherwise, the RCF would itself rely on high-resolution runs in large boxes which are exactly

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Figure 4. The resolution-induced power suppression is systematic and can be corrected for. In this plot we show the resolution corrections $f_{\rm res}^{3\to8}(k, z; \boldsymbol{c})$ computed from a sample of 20 cosmologies represented by the yellow curves. The newly introduced uncertainty due to variance in cosmology, which we neglect, is merely 0.5 per cent (standard deviation, green shaded region) or lower at $z = 0$ and at the level of 1 per cent at $z \sim 2$. Given the fact that the bias itself is at the 5 per cent level at higher redshifts, correcting for the resolution has a positive net effect. The curves corresponding to the Euclid Reference Cosmology are overplotted. We shall use this cosmology to compute the RCF which we apply to all other cosmologies.

Table 2. Parameter box for `EuclidEmulator2` defined through its lower bounds ('min') and its upper bounds ('max'). The central cosmology of the parameter box ('EE2centre'), which is *almost* identical to the Euclid Reference Cosmology, is also given here. $\Omega_{\rm rad}$ is the same for all cosmologies, corresponding to $T_{\rm CMB} = 2.7255$ K. Notice that the maximum value for w_a mentioned reflects the actual upper bound allowed by the emulator, while values $0.5 < w_a \leq 0.7$ were considered as well during the construction of the `HALOFIT`-based mock emulator. See Section 6.4 for a discussion.

	min	max	EE2centre	Euclid Reference
$\Omega_{\rm b}$	0.04	0.06	0.05	0.049
$\Omega_{\rm m}$	0.24	0.40	0.32	0.319
$\sum m_\nu$	0.0 eV	0.15 eV	0.075 eV	0.058 eV
$n_{\rm s}$	0.92	1.00	0.96	0.96
h	0.61	0.73	0.67	0.67
w_0	-1.3	-0.7	-1.0	-1.0
w_a	-0.7	0.5	0.0	0.0
$A_{\rm s}$	1.7×10^{-9}	2.5×10^{-9}	2.1×10^{-9}	2.1×10^{-9}

the simulations that are not affordable. Whether this is the case was tested by comparing the RCF from a $L = 128\,h^{-1}$ Mpc box to the corresponding RCF computed from a simulation run in a $L = 512\,h^{-1}$ Mpc box. As the latter curve (requiring a simulation with $N^3 = 4096^3$ particles) is already very expensive to produce, we did this test for only one single cosmology. Further, we tested if the RCF $f_{\rm res}^{3\to8}(z, k)$ depends on the box size and we found that it does not in any other way than the fact that $k_{\rm min} = 2\pi/L$ is of course changed. There is a limit to the minimally allowed box size, though. While computing an RCF one must make sure that a box size is chosen such that $f_{\rm res}^{3\to8}(z, k_{\rm min}) = 1$ for all z of interest such that the resulting RCF can be safely extrapolated to larger scales by setting $f_{\rm res}^{3\to8}(z, k) \equiv 1$ for all $k < k_{\rm min}$.

5.3 Correction strategy

Our suggested strategy to fight this resolution effect is hence as follows. Starting from a power spectrum with a resolution parameter of $\ell^{-1} = 3\,h\,\mathrm{Mpc}^{-1}$, we can resolution-correct it by multiplying it with a k- and redshift-dependent (but cosmology-independent) RCF $f_{\rm res}^{3\to8}(k, z)$, i.e.

$$P^{\ell^{-1}=8\,h\,\mathrm{Mpc}^{-1}}(k, z; \boldsymbol{c}) \approx f_{\rm res}^{3\to8}(k, z) P^{\ell^{-1}=3\,h\,\mathrm{Mpc}^{-1}}(k, z; \boldsymbol{c})\,. \tag{27}$$

This corresponds to lowering the minimally resolved mass by roughly an order of magnitude from $\sim 3.3\times10^9$ to $\sim 1.7\times10^8$ $\mathrm{M}_\odot/h$ (corresponding to a Nyquist frequency of $k_{\rm max} \sim 25\,h\,\mathrm{Mpc}^{-1}$).

As we define the NLC for `EuclidEmulator2` with respect to linear theory which is not affected by this resolution effect, the very same correction can be applied to those quantities:

$$B^{\ell^{-1}=8\,h\,\mathrm{Mpc}^{-1}}(k, z; \boldsymbol{c}) \approx f_{\rm res}^{3\to8}(k, z) B^{\ell^{-1}=3\,h\,\mathrm{Mpc}^{-1}}(k, z; \boldsymbol{c})\,. \tag{28}$$

6 PROJECTION STUDIES USING CLASS-BASED MOCK EMULATORS

N-body simulations of the matter field in $w_0 w_a$CDM+$\sum m_\nu$ cosmologies are expensive, even when the mass resolution is low. It is hence not affordable to run thousands of simulations that would allow for an in-depth analysis of a given emulator that involves (potentially several) training, test, and validation sets. In order to develop an understanding of the final emulation error and its dependence on the dimensionality of the parameter space as well as on the size

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

of the experimental design, we construct mock emulators based on HALOFIT (Bird, Viel & Haehnelt 2012) data. We have followed this strategy already in Euclid Collaboration (2019) where it has proved to yield a reliable estimate for the performance of the real, simulation-based emulator.

In this section we first define the parameter space inside which the emulator is constructed. This parameter space is the same for the HALOFIT-based emulator as for the actual simulation-based EuclidEmulator2. In a next step we (approximately) optimize the hyperparameters of the model in order to use the resulting 'architecture' for the computation of learning curves, etc. We shall then apply the findings to the training of the actual, simulation-based EuclidEmulator2.

6.1 Definition of the parameter space

For EuclidEmulator2 we consider CDM models with dynamical DE and including massive neutrinos, often abbreviated as w_0w_aCDM+$\sum m_\nu$ models. More precisely, this means that we parametrize the considered cosmologies via the following eight parameters:

(i) Ω_b, the total baryon density parameter in the Universe ,
(ii) Ω_m, the total matter density parameter in the Universe ,
(iii) Σm_ν, the sum of masses of all neutrino families ,
(iv) h, the dimensionless Hubble parameter ,
(v) n_s, the spectral index ,
(vi) w_0, the time-independent part of the DE EoS parameter ,
(vii) w_a, the linearly scale factor-dependent part of the DE EoS parameter ,
(viii) A_s, the spectral amplitude .

The radiation density Ω_rad is given by the CMB temperature which we fix at 2.7255 K. The DE density Ω_DE is then inferred from the flatness condition given by equation (4) with $\Omega = 1$ on the left-hand side.

We do not want to include any prior knowledge about a most likely cosmology into the construction of the emulator other than the assumption that a flat w_0w_aCDM+$\sum m_\nu$ model is sufficiently accurate in order to describe our Universe. In the context of EuclidEmulator2, we explicitly ignore alternative gravity and other more exotic cosmological models. Emulators for such models have been published by other research groups as e.g. Winther et al. (2019) and Giblin et al. (2019). We thus apply flat (uniform) priors to each of the eight input parameters. In order to have a well-defined, normalized prior probability distribution function we thus need to define compact intervals along each dimension over which the final emulator will be defined. Mathematically this means that we have to define intervals $[a_1, b_1], \ldots, [a_8, b_8]$ such that the final parameter box Π is the Cartesian product of all intervals:

$$\Pi := \bigtimes_{i=1}^{8} [a_i, b_i] \tag{29}$$

The choice of the interval boundaries is mostly arbitrary and depends mainly on the tasks that will be tackled by the emulator. Without imposing any restrictions, we assume that EuclidEmulator2 will be mostly applied to MCMC searches of the cosmological parameter space to solve the inverse problem of finding the parameter values best describing our Universe. It is reasonable to assume that these values are not too far away from the current best-fitting values as published by modern cosmological experiments. We centre the parameter box for EuclidEmulator2 around the cosmology

Figure 5. Evolution of the first-order principal component weight w_{PC_1} as a function of w_0 and w_a. The exponentially increasing value is evident for $w_0 + w_a \to 0$. The white, dashed line indicates where the cut of the parameter space is made along the w_a axis in order to avoid the problematic region. We have investigated also less aggressive cuts all of which lead to worse performance of the emulator, suggesting that the problematic region is even more extended in higher dimensions.

'EE2centre' which is defined in Table 2. Notice that this central cosmology is identical to the Euclid Reference Cosmology (up to two decimal places) for all dimensions but the sum of the neutrino masses. We are thus left with the definition of the width of the intervals along each dimension in such a way that the resulting parameter box remains small enough that a sample of size ∼200 (corresponding to our computational budget) contains enough information to achieve a generalization error of $\lesssim$1 per cent. To determine these in a systematic way we run a number of full N-body simulations along each parameter axis, both below and above the central value. This allows us to determine by how much each individual parameter has to be varied in order to cause roughly a ±10 per cent variation in the output NLC factor (the emulation target). Following this prescription, the output variation is an order of magnitude larger than the uncertainty in the output ensuring significant discrimination power while keeping the parameter box, Table 2, reasonably small. Yet, for certain tasks (particularly in the field of weak gravitational lensing), the resulting parameter box may be too small. However, further away from the central cosmology, 1 per cent accuracy is no longer necessary and EuclidEmulator2 could be extended via a multifidelity procedure. In addition, it is expected that cosmologies outside this parameter box can be ruled out by the linear theory power spectrum alone.

The range for neutrino masses allows for less than 10 per cent of output variability. We accept this compromise in order to improve emulation accuracy (due to a reduced volume of the parameter box). At the same time we do not expect this cut to have a large impact because the neutrino signal is expected to mostly affect the linear scales. Notice further that, for reasons discussed in Section 6.4, we trained EuclidEmulator2 only on cosmologies with $w_a < 0.5$. While this does in principle not change the parameter box over which the emulator is defined, in the actual implementation of EuclidEmulator2 the range is restricted to $w_a \in [-0.7; 0.5]$ in order to not allow the emulator to extrapolate. The cut at $w_a = 0.5$ is illustrated in Fig. 5.

6.2 Halofit mock data sets

For the HALOFIT-based analyses we create multiple data sets for training and validation. The training data sets were sampled using Latin hypercube sampling (LHS) first published in McKay,

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Beckmann & Conover (1979), just in the same way as reported on in Euclid Collaboration (2019). We create a series of Latin hypercube (LH) samples with different sizes, $n_{\rm ED} \in \{25, 50, 100, 200, 300, 400\}$. For each size we generate $5n_{\rm ED}$ sets and choose the realization that maximizes the minimum distance between all sampling points.

For validation of the HALOFIT-based emulator we create much larger sets than for training. Notice that sampling large sets with LHS is computationally demanding as it scales polynomially with the number of points to sample. Additionally, we are primarily interested in the performance of the emulator inside an axis-aligned ellipsoid inscribed in the parameter box (Table 2). LHS, however, is designed to be space-filling and thus the high computational cost comes with a low efficiency as due to the high dimensionality of the parameter space most sample points lie outside that ellipsoid. We hence decided to sample the validation sets purely randomly and filter the sample with an ellipsoidal mask. Following this procedure we generate in total 30 validation sets with roughly 1500 sampling points each (resulting in 45 000 validation points).

In addition to these data sets, we have also created data sets for cosmologies organized in a grid of 50×50 points in each parameter plane, resulting in another 70 000 HALOFIT evaluations. These sets are on the one hand used for analysis of the principal component analysis (PCA) eigenvectors (see Section 6.4) and, on the other hand, also for validation purposes (see Sections 6.7 and 6.9).

For each sampled cosmology in all of the sets mentioned above, we run HALOFIT and evaluate the matter power spectrum at redshift $z = 0$. This result is then divided by the linear matter power spectrum in order to get the NLC factors which ultimately form the data sets of interest.

6.3 Emulation strategy: PCE

We use a supervised regression technique called polynomial chaos expansion (PCE) to emulate the NLC factor. We use the implementation of this method in the MATLAB package UQLab[13] (Marelli & Sudret 2014, 2017; Marelli, Lamas & Sudret 2017). PCE in its generality is well documented in several publications such as Xiu & Karniadakis (2002), Blatman (2009), Blatman & Sudret (2009), Blatman & Sudret (2010), Blatman & Sudret (2011), Marelli & Sudret (2017), Marelli & Sudret (2018), Torre et al. (2019), and its application to cosmological emulation is discussed in Euclid Collaboration (2019). As a reminder we repeat that we express NLC factors using PCA and PCE with the following emulation equation

$$B(k, z; \boldsymbol{c}) \approx \mu_{\rm PCA}(k, z) + \sum_{j=1}^{n_{\rm PCA}} \sum_{\alpha \in \mathcal{A}_j^{p,q,r}} \hat{\eta}_{j,\alpha} \Psi_j^{\alpha}[f(\boldsymbol{c})] {\rm PC}_j(k, z)\,, \quad (30)$$

where $\boldsymbol{c}$ stands for a vector of the eight cosmological parameters discussed here which is transformed through f into the standard unit hypercube $[-1, 1]^8$. The PCA quantities are the mean $\mu_{\rm PCA}(k, z)$ and the eigenvectors ${\rm PC}_j(k, z)$. The actual PCE is given by the inner sum, with Ψ_j^{α} being the PCE basis functions $\hat{\eta}_{j,\alpha}$ the coefficients and α being an element from a multi-index set $\mathcal{A}_j^{p,q,r}$.

We shall also stress again that this procedure of combining PCE and PCA is the standard approach for emulating vector-valued quantities with PCE (Blatman & Sudret 2013). This implies, however, that EuclidEmulator2, having a target space of $n_{\rm PCA}$ dimensions, is actually a conglomerate of $n_{\rm PCA}$ single, scalar-valued emulators.

[13] www.uqlab.com

The actual learning algorithm we use is a regularized (i.e. LASSO-type) version of a least-squares minimization called Least angle regression (LAR), discussed in detail in Efron et al. (2004) and Blatman & Sudret (2011). This regression algorithm minimizes the loss function

$$\hat{\eta}_{j,\alpha} = \underset{\eta_j \in \mathbb{R}^{|\mathcal{A}_j|}}{\rm argmin}\ \mathbb{E}\left[\left(\eta_j^{\top} \cdot \Psi_j(\boldsymbol{c}) - w_{j,{\rm true}}(\boldsymbol{c})\right)^2 + \lambda \sum_{\alpha \in \mathcal{A}_j} |\eta_{j,\alpha}|\right]\,. \quad (31)$$

Here, $w_{j,\,{\rm true}}$ denotes the true value of the j-th principal component weight and we use $\mathcal{A}_j$ as a shorthand notation for $\mathcal{A}_j^{p,q,r}$ for the sake of readability.

The regularization term enforces low-rank (i.e. sparse) solutions. This is motivated by the so-called *sparsity-of-effects principle* according to which most of the variance of the underlying model is encoded in interaction terms among only a small number of parameters. Further, enforcing sparse basis representations serves the purpose of reducing the memory requirements of the emulator code.

6.4 Principal component analysis

Vanilla PCE can only predict scalars. In order to create a PCE emulator for a non-scalar quantity like NLC, it is thus mandatory to decompose the full signal into a series such that only scalar coefficients have to be emulated. This is done in the standard way using PCA (also used in other emulators based on different techniques than PCE, such as e.g. Heitmann et al. 2010; Nishimichi et al. 2019 and others). The coefficients of the principal components (also called weights or eigenvalues) are thus the quantities that are actually emulated. To our knowledge, the dependence of those coefficients on the cosmological parameters has not been investigated in any of the papers about cosmic emulators employing PCA published over the last decade.

As we use a polynomial regression method of finite order for emulation, we implicitly assume that the dependence of the PC weights on the cosmological parameters is sufficiently polynomial. To investigate if this assumption is justified we perform a PCE on each set of the above-mentioned 2500 data points sampled in the coordinate planes and plot the resulting first-order coefficients $w_{{\rm PC}_1}$ as heatmaps. The (w_0, w_a) plane stands out as for $w_0 + w_a \to 0$ the corresponding first-order coefficient grows exponentially (see Fig. 5). From a physical point of view this does not come as a surprise as a cosmology with such a DE EoS is highly exotic as it implies a DE with an almost matter-like nature shortly after the big bang. Such a cosmology is not of interest to us as it is highly unrealistic. We have thus identified a clear non-polynomial dependence in the functions we try to emulate. We mitigate this problem by masking out the critical region. In practice, we train our emulator based on all cosmologies within our training set but exclude the very 19 cosmologies that do not meet the condition:

$$w_a < 0.5\,. \quad (32)$$

As we will show in Section 6.7, this modification of the training data set is crucial for the performance of the emulator (although its size is decreased from 127 to only 108 training cosmologies!). This does, however, not restrict the allowed input parameters of the resulting emulator in any way. Clearly, the generalization performance of the emulator is considerably worse in the masked region than for cosmologies with $w_a < 0.5$. We have tried other, less aggressive cuts, too (e.g. cutting along $w_0 + w_a < 0.5$) but have found that

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

only the cut along $w_a < 0.5$ leads to a satisfactory generalization performance.

6.5 About error measurements

As in all machine learning (ML) and uncertainty quantification (UQ) tasks, error quantifications play a central role in this work. Ultimately, we are primarily interested in creating an emulator that generalizes well in an L_1 sense in the cosmological parameter space. To be more precise, we try to minimize the generalization error of the emulator given by the maximum of the relative mean absolute error:

$$\varepsilon_{\mathrm{maxrMAE}} = \max_{k,z} \left\langle \left| \frac{B_{\mathrm{emu}}(k,z;\boldsymbol{c}) - B_{\mathrm{true}}(k,z;\boldsymbol{c})}{B_{\mathrm{true}}(k,z;\boldsymbol{c})} \right| \right\rangle_{\boldsymbol{c}\in C} . \tag{33}$$

Notice that the mean indicated by the angle brackets is taken over the cosmologies $\boldsymbol{c}$ defined in the parameter space C given in Table 2. The maximization, on the other hand, is performed over the non-regressed parameters k and z. This generalization error will be approximated by a validation error of the form

$$\hat{\varepsilon}_{\mathrm{maxrMAE}} = \max_{k,z} \frac{1}{N_{\mathrm{val}}} \sum_{\boldsymbol{c}\in C_{\mathrm{val}}} \left| \frac{B_{\mathrm{emu}}(k,z;\boldsymbol{c}) - B_{\mathrm{true}}(k,z;\boldsymbol{c})}{B_{\mathrm{true}}(k,z;\boldsymbol{c})} \right| \tag{34}$$

where C_{val} designates the set of validation cosmologies.

However, since an estimate of the above error requires a validation set which we do not have readily available in all situations, we also often use a cross-validation error metric as an alternative. This is given by the leave-one-out (LOO) error defined as

$$\varepsilon_{\mathrm{LOO},j} = \frac{\sum_{i=1}^{n_{\mathrm{ED}}} \left[w_{\mathrm{true},j}\left(\boldsymbol{c}^{(i)}\right) - w_{\mathrm{emu},j}^{\mathrm{PCE}\backslash i}\left(\boldsymbol{c}^{(i)}\right) \right]^2}{\sum_{i=1}^{n_{\mathrm{ED}}} \left[w_{\mathrm{true},j}\left(\boldsymbol{c}^{(i)}\right) - \hat{\mu}_{w_j} \right]^2} . \tag{35}$$

Here, w stands for the eigenvalues of the PCA which are the quantities that are actually emulated in this work (for more details on this refer to Euclid Collaboration 2019). Each w_j corresponds to the inner sum in equation (30) running over the multi-index α. To compute this error one trains a PCE emulator on all training example but the i-th one (indicated by the superscript PCE$\backslash i$). In this very example the emulator is then evaluated (second term in the numerator). The quantity is finally rescaled by the overall variance of the quantity w_j. More details on this quantity and how to compute it efficiently can be found in Marelli & Sudret (2017). It shall be emphasized that the subscript i runs over cosmologies in the *training* set and no reference to any validation examples is made. Further, this metric measures the emulator performance not in the NLC space but rather in the more abstract, associated principal component space (hence the subscript j which refers to the order of the principal component).

6.6 Hyperparameter optimization using the `Halofit` mock data

The hyperparameters of PCE for emulation of non-scalar quantities are given by

(i) the minimum percentage a_{PCA} of explained variance retained in the PCA (strongly related to the number n_{PCA} of principal components taken into account),
(ii) the polynomial order p at which the PCE is truncated,
(iii) the maximum interaction r (number of factors per monomial in PCE),
(iv) the q-norm.

Remember that actually there is a separate PCE for each principal component (as described in Section 6.3). As a result, the hyperparameters p, q, and r can be chosen differently for each $i = 1, \ldots, n_{\mathrm{PCA}}$ and the accuracy parameter a is the only hyperparameter that has to be chosen globally for obvious reasons. As training a PCE is relatively cheap, there is no need for a sophisticated optimization algorithm. Rather, we perform a (partially greedy) grid search over the grid given by

$$\begin{aligned} a_{\mathrm{PCA}} &\in \{0.9, 0.99, \ldots, 0.99999999\}, \\ r &\in \{1, 2, \ldots, 5\}, \\ p &\in \{2, 3, \ldots, 20\}, \\ q &\in \{0.3, 0.35, 0.4, \ldots, 0.9\}. \end{aligned} \tag{36}$$

The small set of low numbers looked at for r are motivated by the 'sparsity-of-effects' principle (Marelli & Sudret 2017). Strictly speaking, there are two steps here. In the first step a vector of a_{PCA}-values is created and, in the second step, for each value of a_{PCA} a multitude of (p, q, r)-grids is searched (one for each principal component). This is important because the optimal point in the (p, q, r)-grid is chosen based on a different criterion than the optimal a_{PCA} value. The former is chosen based on a minimization of the LOO cross-validation error without ever seeing a validation point. This happens entirely on the level of principal component weights (i.e. eigenvalues of the covariance matrix) and thus this step is performed independently for each principal component. For identification of the (near-)optimal a_{PCA} value, in contrast, an emulator is trained for each value of a_{PCA} and evaluated on a validation set. We then aim to minimize $\hat{\varepsilon}_{\mathrm{maxrMAE}}$ as defined in equation (34). Notice that, while the latter error is the correct quantity to look at when judging the overall performance of the emulator, this can only be investigated once a separate emulator is trained (and, as well, a separate set of hyperparameters is optimized) for each principal component. It is of paramount importance to understand that we go through this procedure really only to fix the values for a_{PCA}. As p, q, and r can be optimized without the need of a separate validation set, they can be optimized once the final emulator is being trained based on actual simulations without the risk of overfitting to a validation set. After evaluating the $\hat{\varepsilon}_{\mathrm{maxrMAE}}$ for all candidate values of a_{PCA}, we conclude that

$$a_{\mathrm{PCA}} = 0.99999 \tag{37}$$

is the optimal (i.e. loss minimizing) value for this global hyperparameter. If we continue to increase a_{PCA}, we find that we are attempting to capture the numerical noise in the simulations, leading to an increase again of $\hat{\varepsilon}_{\mathrm{maxrMAE}}$.

6.7 Learning curves

Now that we have defined the bounds for all parameters, have created the necessary data sets, and optimized the global hyperparameter a_{PCA}, it is natural to ask the following two questions:

(i) How does the validation error of the emulator decrease as the number of training examples increases?
(ii) Given a fixed number of training examples, how does the validation error increase as a consequence of adding the two additional dimensions compared to version 1 of `EuclidEmulator`?

To answer these questions we train emulators for different cosmological models with different dimensionalities, namely ΛCDM (5D), w_0CDM (6D, corresponding to `EuclidEmulator1`), w_0w_aCDM (7D), and w_0w_aCDM+$\sum m_\nu$ (8D, corresponding to `EuclidEmulator2`). Each model is trained on the series of training sets mentioned in Section 6.2. The emulators are then validated on the 30

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Figure 6. Learning curve for HALOFIT-based emulators for four different cosmological models. On the *y*-axis the estimated mean of the validation error distribution over the 30 different validation sets is plotted. The error bars show the standard error of this estimated mean. Notice that we explicitly plot the point at $n_{\mathrm{ED}} = 127$ for the w_0w_aCDM+$\sum m_\nu$ model as this point corresponds to our actual training set. The learning curves have been computed (i) taking all training examples into account (dashed lines, labelled 'uncut') and (ii) ignoring those training examples with $w_a > 0.5$ (solid lines, labelled 'cut'). For the ΛCDM and w_0CDM models this distinction makes no difference because for those models w_a was fixed to 0.

Validation examples were sampled randomly inside axis-aligned hyperellipsoid inscribed in the parameter box (see Section 6.2 for a description of the data sets).

ellipsoidal validation sets. In this context it becomes evident why we have produced so many validation sets: It allows one to get statistics on the validation error (namely the standard error of the estimated mean error). Notice that such a representative test is by far beyond what is achievable with simulation data as several tens of thousands of simulations with at least moderate mass resolution would have to be run. The resulting learning curves are plotted in Fig. 6.

While it is easy to achieve validation errors < 1 per cent for the 5D and 6D models with only 50 training examples, the complexity of the 7D and 8D models is considerably higher, particularly if we train based on examples in the entire original parameter space without masking out the problematic region in the (w_0, w_a) plane (see the discussion in Section 6.4). In this scenario, we would require $\gtrsim$400 training examples to reach accuracies of 1 per cent or better. Masking out the region where $w_0 + w_a \sim 0$ in the training set reduces the training set size to only 100 to 200 examples.

Notice that in Fig. 6 we show the error estimate for $n_{\mathrm{ED}} = 127$ for the case of w_0w_aCDM+$\sum m_\nu$ cosmologies. This measurement was added in hindsight because it turned out that the actual simulation-based emulator would achieve the target accuracy already with a training set of only 127 simulations selected from an LHS of size 200 (see Section 7.2 for a short discussion).

6.8 Application to the training of EuclidEmulator2

We shall now anticipate some training aspects of the actual, simulation-based EuclidEmulator2. We use the same surrogate model in order to train the emulator as for the HALOFIT-based mock emulator discussed above and for EuclidEmulator1, i.e. sparse PCE combined with PCA. The number of principal components is defined through the threshold value for the minimally explained variance in the data set (which is independent of the size of the training set). As this was investigated in Section 6.2 and found that $a_{\mathrm{PCA}} = 0.99999$ is the optimal value, we can now use the same value for the training of the actual emulator. Having found in the learning curves analysis that a training set of 127 cosmologies is enough to achieve the targeted accuracy, this value for a_{PCA} corresponds to retaining $n_{\mathrm{PCE}} = 14$ principal components (i.e. the target space of the full vector-valued emulator is 14-dimensional).

Table 3. Table with optimal hyperparameter values for all 14 scalar-valued PCEs of EuclidEmulator2. The resulting PCE contains 574 non-trivial terms.

PC order	p	r	q
1	3	2	0.45
2	3	2	0.45
3	4	2	0.5
4	4	2	0.45
5	3	2	0.45
6	4	2	0.4
7	4	2	0.5
8	4	2	0.45
9	4	2	0.5
10	16	3	0.4
11	4	2	0.5
12	12	4	0.5
13	4	3	0.5
14	15	2	0.4

Also, we found for the HALOFIT-based mock emulator that setting the maximal interaction $r_{\mathrm{max}} = 4$, the maximal polynomial order $p_{\mathrm{max}} = 20$, and varying the q-norm between $q_{\mathrm{min}} = 0.3$ and $q_{\mathrm{max}} = 0.9$ leads to convergence in the selection of terms in the expansion series. So we recycle this here, too. The parameters p and q, as well as the actual interaction number r, are optimized individually for each principal component and listed in Table 3. The full training process (i.e. optimizing the values for p, r, and q for each principal component and fitting the coefficients) takes only 9 s on a usual MacBook pro with a 2.8 GHz Intel Core i7 CPU. As a result we get a PCE with a total of 574 terms (all other coefficients vanish). Notice that this corresponds to a very small number of terms, i.e. an extremely sparse PCE, as in our case there were 3108 105 terms without sparsification.

6.9 Performance estimation of the mock emulator

From the learning curves presented above we can expect EuclidEmulator2 to be sub-1 per cent accurate at $z = 0$ over the entire k range of interest for the available training set of 108 simulations (being a subset of the originally planned set of 200 training examples). While this is the final goal, it is yet interesting to see how the error evolves as a function of cosmology. For this we evaluated the HALOFIT-based mock emulator on all 70 000 validation cosmologies sampled in the 28 parameter planes of the feature space. The results are 28 error maps that we show in Appendix B. It is clearly visible that the validation error is below 2 per cent (and hence at the same level as the mass resolution-related uncertainty in the simulations at small scales) for the vast majority of cosmologies lying inside the axis-aligned hyperellipsoid inscribed in the parameter box (indicated by a grey ellipse in the error maps). Only for cosmologies with a large value of $w_0 + w_a$ the error grows to ∼5 per cent. We reiterate, however, that on average over the entire 8D-hyperellipsoid the error drops below 1 per cent. Outside that ellipsoidal region, the errors sometimes exceed the 2 per cent limit. We also designate the $w_a = 0.5$ boundary by a grey dashed line.

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Figure 7. Distribution of LH-sampled cosmologies represented in the coordinate planes of the parameter space. The blue circles and histograms show the full ED of 200 data points, while the green dots indicate the subset of 127 training examples actually used to create `EuclidEmulator2`. The grey ellipses correspond to the slices of the hyperellipse within which the performance of `EuclidEmulator2` is optimized and the dashed, grey lines correspond to the cut at $w_a = 0.5$ explained in Section 6.4. The ranges of each subplot (mentioned along the x-axis for convenience) exactly match those specified in Table 2.

7 THE TRAINING SET OF EUCLIDEMULATOR2

7.1 Experimental design: sampling

We sample the points in the parameter space defined in Table 2 using LHS, as we have done already in Euclid Collaboration (2019). LHS is a very straightforward sampling technique that is widely used and accepted in the cosmological emulator community (Heitmann et al. 2009, 2010, 2014; DeRose et al. 2019; Gration & Wilkinson 2019; Nishimichi et al. 2019; Rogers et al. 2019) and extensively presented in the statistical sampling literature (McKay et al. 1979; Tang 1993; Liefvendahl & Stocki 2006; Swiler, Slepoy & Giunta 2006; Crombecq, Laermans & Dhaene 2011; Damblin, Couplet & Iooss 2013; Garg & Stogner 2017; Sheikholeslami & Razavi 2017; Yang, Liu & Li 2017). Endowed with an additional optimization step (we use a distance-based criterion), its main advantage is that it combines good space-filling properties with a high degree of randomness. For an in-depth explanation of the exact steps we go through to generate the sample we refer to Euclid Collaboration (2019).

We chose to generate a sample with 200 points based on the argument that in Fig. 6 we show that a training set of this size should be large enough to achieve a validation error below 1 per cent, while a set of only 100 examples is expected to just miss this requirement in the 8D parameter box. In fact, we use exactly the same LHS of size $n_{\rm ED} = 200$ to run the simulations as we used in the investigation of the mock emulators in Section 6.

The resulting sample of cosmologies is shown in blue in Fig. 7.

7.2 Experimental design: simulations

The generation of each training example (i.e. each pair of simulations per cosmology) corresponds to an investment of about 4000 node hours of computation. Unfortunately, when we sampled the ED of 200 cosmologies, we underestimated the required computational resources. As a result, we ran out of computing time after having completed merely 127 pairs of simulations. Luckily, we noticed that the emulator achieved the targeted sub-percent accuracy when trained on only 127 examples. In this case the sample of course is no longer an LHS but rather resembles a random sampling. Fortunately, we realized that there was no need to invest more time and effort in running the remaining 73 pairs of simulations. The completed set of 127 simulation pairs used for the training of `EuclidEmulator2` is plotted in Fig. 7. Notice that it was for this very reason why we also looked at the sample of size 127 when investigating the `HALOFIT`-based mock emulator in Section 6.7.

From this fact we can learn two important conclusions for future projects:

(i) At least as long as the marginal distributions of sample points along all parameter dimensions do not have regions where there are no sampling points at all, LHS is not a necessity for good performance of PCE as random sampling works fine too.

(ii) When generating an ED is related to large computational costs (and hence to a non-negligible risk of failing to generate a large sample in one go), it is advisable to choose an enrichable sampling technique such as e.g. the LHS-based adaptive response surface method (LHS-ARSM, Wang 2003) or active learning (as e.g. done in Rogers et al. 2019).

We thus use an experimental design of 127 (respectively 108 after applying the cut in w_a) P+F simulations randomly sampling the parameter box. Each simulation samples the power spectrum at 613 k-modes and 100 time-steps between $z_{\rm intrm} = 10$ and $z_{\rm fin} = 0$. While all this data are used to compute the emulator, we only allow the user to emulate up to $z_{\rm max} = 3.0$ because the overall accuracy

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Figure 8. Sobol' analysis plots for the first three principal components. Clearly, Ω_m is by far the most impactful cosmological parameter of all as its Sobol' index S_1 is large for the first two principal components. While the neutrino mass, the spectral index, and both DE EoS parameters have only a weak influence on the first-order principal component, the spectral index is the most important parameter for PC_2 and w_0 and w_a are dominant at third PC level. Neutrino mass becomes only relevant in the seventh principal component (not shown) highlighting the fact that its impact is very small (see discussion in Section 8.2.2, right-hand panel of Fig. 11 in particular, and Section 9).

decreases considerably for higher redshifts (primarily because the underlying simulations have not converged for higher redshifts as can be extrapolated from Fig. 3).

7.3 Post processing: computation of the NLC

We compute the NLC for each power spectrum by dividing the nonlinear power spectrum resulting from the simulation by the linear theory power spectrum computed by `CLASS`. Notice that this is different from what is done for `EuclidEmulator1` where the NLC was computed via a division by the re-scaled power spectrum measured from the simulation particle realization at the initial condition. For training the emulator, the NLC is converted into log space because we have shown in Euclid Collaboration (2019) that this improves the generalization of the emulator. We stress that the RCF introduced in Section 5 is *not* applied to the training data set such that users of `EuclidEmulator2` can decide individually whether they want to apply this correction or not to the emulated result. We compile the NLC data into a data matrix $\mathcal{D}^{\mathrm{CDM+b}} \in \mathbb{R}^{n_{\mathrm{ED}} \times n_z n_k}$. This data matrix is then decomposed into its principal component basis $\{\mathrm{PC}_i | i \leq n_{\mathrm{PCA}}\}$ where n_{PCA} denotes the number of principal components taken into account. As a result, the m-th row in $\mathcal{D}$ can be represented as follows:

$$\mathcal{D}_m = \sum_{i=1}^{n_{\mathrm{PCA}}} w_i(c_m)\mathrm{PC}_i(k, z)\,, \tag{38}$$

where the argument c_m of the PC weight stands for the vector of parameters defining the m-th cosmology and the arguments of the principal components are the k mode and the redshift. We hence build n_{PCA} individual training sets defined by

$$\mathcal{T}_i = \{w_i(c_m) \mid m \leq n_{\mathrm{ED}}\}\,,\ \forall\, i \in \{i, \ldots, n_{\mathrm{PCA}}\} \tag{39}$$

that are used to train the n_{PCA} individual, scalar-valued emulators.

8 EMULATOR PERFORMANCE, ERRORS AND SENSITIVITY TO PARAMETERS

8.1 Sensitivity analysis

As for `EuclidEmulator1`, we have again performed a Sobol' analysis to investigate the relative importance of each cosmological parameter on the final NLC. Notice that since the parametrization of the emulator changed significantly from `EuclidEmulator1`, it cannot be expected that the Sobol' indices remain unaltered. Clearly, for `EuclidEmulator2`, the matter density parameter Ω_m dominates the behaviour of the resulting NLC. At the same time, to first principal component order, $\sum m_\nu$ and n_s are almost entirely negligible (therefore, post-processing approaches to add a neutrino-contribution as suggested e.g. in Agarwal & Feldman 2011 work fine as long as only the power spectrum is considered). While in the case of neutrino masses this does not come as a surprise (the effect of massive neutrinos is mostly captured by the linear signal already), one might not have guessed that for the spectral index. This is resolved when looking at the Sobol' indices of the second principal component where the neutrino mass still has almost no impact at all, while the spectral impact becomes actually the dominant parameter.

The fact that the sum of the neutrino masses is almost entirely negligible when computing the NLC supports our suggestion mentioned in Euclid Collaboration (2019) that to good approximation one can emulate nonlinear power spectra with massive neutrino cosmology by simply computing the corresponding linear power spectrum and multiplying that by an NLC as produced by `EuclidEmulator1`, i.e. an NLC that does not know anything about massive neutrinos. The test of this hypothesis is deferred to Section 8.2.2 (see Fig. 11 in particular).

8.2 Generalization performance of `EuclidEmulator`

In this section we shall compare `EuclidEmulator2` to other fast prediction techniques such as `HALOFIT` (Bird et al. 2012), `HMCode` (Mead et al. 2016), `CosmicEmu` (Lawrence et al. 2010, 2017), the very recent emulator based on the BACCO simulation project (Angulo et al. 2020), hereafter referred to as the 'BACCO-emulator', and the predecessor `EuclidEmulator1` (Euclid Collaboration 2019) as well as with `PKDGRAV3` (Stadel 2001; Potter & Stadel 2016; Potter et al. 2017) simulations. While the comparisons of `EuclidEmulator2` with `EuclidEmulator1` and `PKDGRAV3`, respectively, can be conducted at the NLC-level, all comparisons

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

with HALOFIT, HMCode, and CosmicEmu are performed at the level of the fully nonlinear power spectrum. To this end we multiply the NLC computed by EuclidEmulator2 with a linear power spectrum computed by CLASS for the same cosmological parameters.

We compare each pair of predictors in two ways. On the one hand, we compare them for a set of different cosmologies at redshift $z = 0$, while on the other hand we chose a single cosmology equal to the Euclid Reference Cosmology but with a higher total neutrino mass for comparison at different redshifts $z \leq 2$.

For the comparison between EuclidEmulator2 and PKDGRAV3 we have used a small validation data set containing three validation cosmologies.

In order to compare EuclidEmulator2 to EuclidEmulator1, we primarily focus on two extreme cases: the Euclid Reference Cosmology as defined in Euclid Collaboration (2019), once with massless and once with massive neutrinos.

For all comparisons with HALOFIT, HMCode, and CosmicEmu in this section, we choose the cosmologies from a set of 291 cosmologies, put together by an LHS of size 200, cosmologies along the coordinate axis and along one of the diagonals. From this set, we filter out all cosmologies that are not accepted by any of the emulators. This results in a set of 84 comparison cosmologies (47 in the case of the comparison to the BACCO-emulator), which are listed and plotted in appendix C. We emphasize that some of these validation cosmologies lie outside the hyperellipsoid inside which the hyperparameters were optimized (see discussion in Section 6.6). Some validation cosmologies even have w_a values larger than 0.5 and thus lie outside the training region of the emulator, indicating that EuclidEmulator2 may to some degree be able to extrapolate relatively accurately.

8.2.1 Comparison of EuclidEmulator2 and PKDGRAV3 simulations

We start our series of comparisons by checking how well EuclidEmulator2 is able to approximate simulation data. To this end, we generate a validation set of P+F simulations with the same resolution as the training data. The validation set contains only three cosmologies (all unseen by the training process) because the generation of a significantly bigger training set is too expensive. These three cosmologies are all sampled from the ellipsoid inscribed the parameter box with axes given by the limits of each parameter range.

We observe in Fig. 9 that the validation error (given by the relative mean absolute error, rMAE) between emulated and simulated NLC factors is well below 1 per cent for k-modes and redshifts of interest to the *Euclid* mission, i.e. $k \leq 10\,h\,\mathrm{Mpc}^{-1}$ and $z \leq 3$. Of course, as the validation set is very small, there is a substantial uncertainty on this estimate and the rMAE is likely to exceed the 1 per cent limit as one exits the hyperellipsoid inscribed by the parameter box. Yet, the overall error is expected to be dominated by uncertainties in the underlying simulations, especially at very small scales, $k \gtrsim 5\,h\,\mathrm{Mpc}^{-1}$.

In the context of comparing EuclidEmulator2 to PKDGRAV3, it is natural to compare our emulator to the EFS2 simulation. To this end, we evaluate both EuclidEmulator2 and CLASS at the Euclid Reference Cosmology defined in Table 2. We then produce a nonlinear power spectrum by multiplying the linear power spectrum, the NLC, and the RCF. This product is then compared to the EFS2 power spectrum in Fig. 10. Because EFS2 is not a P+F simulation, the cosmic variance is clearly visible as oscillations at the level of a few per cent at linear scales. Generally, the agreement between EuclidEmulator2 and EFS2 is at the 1 per cent level or better for nonlinear k modes.

Figure 9. Comparison of NLC factors predicted by EuclidEmulator2 and ones computed directly from PKDGRAV3 simulations (averaged over three different cosmologies). The agreement is at the sub-percent level and thus respects the target accuracy.

Figure 10. Comparison of full nonlinear power spectra between EFS2 and resolution-corrected product of a CLASS linear power spectrum times an EuclidEmulator2-emulated NLC at the Euclid Reference Cosmology as defined in Table 2. The agreement is generally very good and even at the 1 per cent level or better at small scales. The oscillations at linear scales are expected as EFS2 is not a P+F simulation. We also show the comparison to EuclidEmulator2 when we do not apply the RCF, which shows a clear 2 per cent bias at nonlinear scales.

8.2.2 Comparison of EuclidEmulator1 and EuclidEmulator2

It is natural to compare the performance of EuclidEmulator2 with its predecessor EuclidEmulator1 (cf. Fig. 11). As both emulators predict the NLC, we can perform the comparison on this level. In a first step we perform this comparison using the version of

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Figure 11. In this figure we compare the NLC prediction of `EuclidEmulator2` to that of `EuclidEmulator1` for two cosmologies: one for which $\sum m_\nu = 0.0$ eV (left-hand panel) and $\sum m_\nu = 0.15$ eV for the other (right-hand panel). A_s is the same in both cosmologies, however the value of σ_8 differs due to the mass difference in the neutrino sector. It is not surprising that the agreement on the largest scales is nearly perfect for both cosmologies as both emulators return values close to unity irrespective of the cosmology by construction. The oscillations around BAO scales are due to cosmic variance. The vertical, black, dashed lines indicate the two k-modes where the curves for `EuclidEmulator1` and `EuclidEmulator2` deviate the most from each other in Fig. 12. The time-dependent mismatch on small scales is due to the different mass resolutions of the training of the two emulators. A key point is that this plot shows that the agreement between `EuclidEmulator1` and `EuclidEmulator2` is not only very good for massless neutrino cosmologies where `EuclidEmulator1` is supposed to work well by construction, but also for w_0CDM+$\sum m_\nu$ models which `EuclidEmulator1` was not designed for. This suggests that `EuclidEmulator1` is able to predict the nonlinear power spectrum within 1 per cent of accuracy for $z < 0.5$ even if the cosmology contains massive neutrinos (of course, the linear power spectrum must be computed taking the neutrino masses into account).

the Euclid Reference Cosmology as defined in Euclid Collaboration (2019), i.e. a cosmology without massive neutrinos.

We observe very good agreement on large scales which is achieved by construction as the variability of the NLC is negligible at these scales. The sub-percent differences at these scales are due to the fact that the simulations volumes of the training simulations underlying both emulators are different.

On intermediate scales around the baryon acoustic oscillations (BAO) one observes a peaky pattern at the level of $\lesssim$2.5 per cent. We show in Fig. 12 that this can be explained by cosmic variance. We reiterate a point already reported in Euclid Collaboration (2019): For `EuclidEmulator1`, cosmic variance is strongest not on large but on intermediate scales. This is because on large scales the cosmic variance is not significantly amplified by nonlinear structure formation. As a consequence the residual (after pairing and fixing) cosmic variance drops out because we compute the NLC for `EuclidEmulator1` by dividing the nonlinear power spectrum at z by the properly rescaled initial condition. On the other hand, we show in Fig. 2 that for `EuclidEmulator2` we choose the volume large enough to render cosmic variance irrelevant. However, on intermediate scales, the residual cosmic variance did not get cancelled out even for `EuclidEmulator1` as on these scales it is already non-negligibly amplified by nonlinear evolution. As a result, when comparing `EuclidEmulator1` to `EuclidEmulator2`, one actually divides two signals with oscillatory behaviour on intermediate scales, manifesting itself as oscillations on intermediate scales observed in Fig. 11.

Figure 12. Comparison of `EuclidEmulator1` and `EuclidEmulator2` to the NLC computed from an ensemble average of 50 Gaussian random field simulations. It can be seen that the cosmic variance is responsible for the oscillatory peaky pattern in Fig. 11, as the vertical, black, dashed lines correspond to the two highlighted peaks in that figure.

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Table 4. Mutually corresponding cosmologies in order to approximate the EuclidEmulator2-based NLC with an EuclidEmulator1-based NLC neglecting massive neutrinos. The relative difference of the NLC factors produced with the respective version of EuclidEmulator are shown in Fig. 11, right-hand panel. Notice that $\Omega_{\rm rad}$ is the same for both cosmologies corresponding to $T_{\rm CMB} = 2.7255$ K.

	EE1	EE2
$\Omega_{\rm b}$	0.049	0.049
$\Omega_{\rm m}$	0.3154	0.319
$\sum m_\nu$	0.0 eV	0.15 eV
$n_{\rm s}$	0.96	0.96
h	0.67	0.67
w_0	− 1.0	− 1.0
w_a	0.0	0.0
σ_8	0.799	–
$A_{\rm s}$	–	2.1×10^{-9}

Unsurprisingly, EuclidEmulator1 underestimates power at small scales compared to EuclidEmulator2. This is simply due to a too low mass resolution of the training simulations of EuclidEmulator1. Here, we can confidently report that the baseline in Fig. 11 given by EuclidEmulator2 is the (more) correct answer.

Now, as we claimed above that massive neutrinos do not have a significant impact on the NLC we should test this hypothesis. To actually do so, we compared predictions of EuclidEmulator1 and EuclidEmulator2 to each other in Fig. 11: We evaluated each emulator at the respective cosmology in Table 4. The ratio between the two NLC factors is clearly dominated by cosmic sample variance and resolution effects. This suggests that EuclidEmulator1 can indeed be used to estimate the NLC to good approximation for $z < 0.5$ for w_0CDM+$\sum m_\nu$ models. The key point to get the correct answer is to account for the difference in $\Omega_{\rm m}$. Given a particular $\Omega_{\rm m}^{\rm EE2} = \Omega_{\rm b} + \Omega_{\rm CDM} + \Omega_\nu$ for EuclidEmulator2, one needs to choose $\Omega_{\rm m}^{\rm EE1} = \Omega_{\rm b} + \Omega_{\rm CDM}$ for EuclidEmulator1, such that $\Omega_{\rm m}^{\rm EE2} = \Omega_{\rm m}^{\rm EE1} + \Omega_\nu$. As a result, the value for σ_8 has to be adjusted accordingly. The value $\sigma_8 = 0.799$, corresponding to $A_{\rm s} = 2.1 \times 10^{-9}$ at the 'EE1' cosmology listed in Table 4, was computed by CLASS.

8.2.3 *Comparison to HALOFIT*

EuclidEmulator2 is compared to the extension of HALOFIT by Bird et al. (2012) in Fig. 13. The comparison across multiple cosmologies shows almost perfect agreement for all cosmologies on large scales. This is expected as HALOFIT builds on linear theory as does EuclidEmulator2. On intermediate scales around BAOs we find systematic oscillations which are in agreement with what we have found in the corresponding comparison between HALOFIT by Takahashi et al. (2012) and EuclidEmulator1 (Euclid Collaboration 2019). While there we attributed those oscillations to HALOFIT's inability to capture the BAOs correctly, this may play a less relevant role for this version of HALOFIT. Rather, the oscillations may be mostly explained by the higher mass resolutions and smaller simulation box sizes used in Bird et al. (2012) compared to the those used in this work. The fact that on average less power is found by EuclidEmulator2 compared to HALOFIT (at the level of roughly 3 per cent) is consistent with the findings presented in Fig. 2 of Bird et al. (2012) where it is reported that PM-based neutrino simulations tend to find less power on intermediate to small scales compared to simulations treating neutrinos as particles. This is also why on small scales we then find an overestimation of power in

Figure 13. The comparison between HALOFIT (Bird et al. 2012) and EuclidEmulator2 is consistent with what is found in Euclid Collaboration (2019). While there are disagreements on intermediate and nonlinear scales, the measured errors stay within the bounds of 5 to 10 per cent as reported in Bird et al. (2012).

HALOFIT relative to EuclidEmulator2. The mean including the 1σ-region stays within the 5 to 10 per cent error margin, respecting the bounds published in Takahashi et al. (2012) and Bird et al. (2012). The error evolution with redshift looks again very similar to what we have already found for EuclidEmulator1. The systematic oscillations on intermediate scales grow with time while on small scales the disagreement is largest for higher redshifts.

8.2.4 *Comparison to HMcode*

The comparison of EuclidEmulator2 and HMCode is shown in Fig. 14. In Mead et al. (2016) it is reported that HMCode achieves an accuracy of a few per cent for cosmologies with massive neutrinos and dynamical DE. We find an agreement at the few per cent level both over all tested cosmologies as well as over all redshifts (see Fig. 14). Independent of redshift and cosmology the agreement on large scales is virtually perfect. Around BAO scales we find a systematic overprediction of power in HMCode relative to EuclidEmulator2 (degrading as z increases) which relaxes again at $k \sim 0.6\,h\,{\rm Mpc}^{-1}$. On small scales, however, the variance in the relative difference is quite large (though always within the few per cent limit) both as the cosmology varies as well as over the probed redshift range.

8.2.5 *Comparison to CosmicEmu*

The comparison of EuclidEmulator2 and CosmicEmu is shown in Fig. 15. In Lawrence et al. (2017), they report that for predictions of the eight-parameter model they find an approximation accuracy of 5 per cent or better. On average over all probed cosmologies, the comparison error is far below that and it is even

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Figure 14. Comparison between `EuclidEmulator2` and `HMCode` over a set of 84 cosmologies at $z = 0$. The agreement on large scales is nearly perfect while the errors stay within the few per cent level over all k as reported in Mead et al. (2016).

Figure 15. Comparison between `EuclidEmulator2` and `CosmicEmu` over a set of 84 comparison cosmologies at $z = 0$. The mean is clearly and the ± 1 standard deviation range is broadly within the 5 per cent level reported in Lawrence et al. (2017).

Figure 16. Comparison between `EuclidEmulator2` and the BACCO emulator over several redshifts. The relative error is mostly smaller than 3 per cent corresponding to the expected accuracy of the BACCO emulator reported in Angulo et al. (2020).

relatively constant over the entire k-range of interest. Even the standard deviation of the entire set of comparisons is only at the level of 5 per cent over all k (for $z = 0$). There are, however, a few cosmologies for which the comparison is significantly poorer. The fact that there is no k-region where the comparison is nearly perfect is explained by the fact that `CosmicEmu` emulates the full nonlinear power spectrum directly while `EuclidEmulator2` emulates the NLC only. It is thus not surprising that there is some generalization error also on large scales for `CosmicEmu`, while `EuclidEmulator2` is accurate in this regime by construction.

8.2.6 Comparison to the BACCO-emulator

The quantity emulated by the BACCO-emulator (Angulo et al. 2020) is also the NLC. For this very reason the comparison between `EuclidEmulator2` and the BACCO-emulator (version 1.1.1) is conducted at the level of the NLC rather than at the fully nonlinear power spectrum level. The result of this comparison is shown in Fig. 16. Clearly, the agreement between these two state-of-the-art emulators is extremely good over wide ranges of spatial scales and redshifts. First, we discuss the comparison between the two emulators at the Euclid Redshift Cosmology for different redshifts. Notice that the BACCO-emulator allows prediction of the NLC only up to $z = 1.5$. For this reason, the comparison at $z = 2$ included in the previous comparisons to `HALOFIT`, `HMCode`, and `CosmicEmu` is omitted here. It is found that the agreement at the tested redshifts is mostly at the per cent level, where a suppression of power in the BACCO-emulator is observed relative to `EuclidEmulator2` at small scales. This is explained by the fact that the BACCO-emulator is based on simulation with a resolution parameter of $\ell^{-1} = 3\,h\,{\rm Mpc}^{-1}$ while the `EuclidEmulator2` NLC were resolution corrected as explained in Section 5.

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Figure 17. Linear and nonlinear power spectra variation due to varying the cosmological parameters (neutrino mass in the left-hand panel, w_a in the right-hand panel). The power spectra are normalized to the Euclid Reference Cosmology power spectrum. The NLCs for the nonlinear power spectra have been predicted with `EuclidEmulator2` in all cases. These plots show clearly that the power spectra differ the most at scales around $k \sim 1\,h\,\mathrm{Mpc}^{-1}$.

The two emulators were also compared at 47 different cosmologies at redshift $z = 0$. The overall agreement is also mostly at the 3 per cent level over the entire k-range, where it is reported in Angulo et al. (2020) that the BACCO-emulator is expected to predict the NLC with an accuracy of 3 per cent. The high-frequency oscillatory pattern at intermediate k-scales may be explained by a somewhat poor sampling of the BAOs in the BACCO simulations.

9 EXPLORATION OF DEGENERACIES IN THE NONLINEAR MATTER POWER SPECTRUM

`EuclidEmulator2` is expected to be applied to parameter forecasts because it is able to very efficiently produce highly accurate predictions of the NLC and hence of the fully nonlinear power spectrum. The *Euclid* mission aims at measuring the absolute neutrino mass scale (Laureijs et al. 2011) by analysing the effects of neutrinos on cosmic structure formation. Massive neutrinos suppress power particularly at small scales (see e.g. Viel, Haehnelt & Springel 2010; Bird et al. 2012; Hannestad, Haugbølle & C. 2012 and others). While this is true also for the linear power spectrum, the effect is largest in the nonlinear power at scales around $k \sim 1\,h\,\mathrm{Mpc}^{-1}$. The reaction of the linear and nonlinear power spectra to varying the total neutrino mass is shown in Fig. 17. In this figure, the base line is given by the Euclid Reference Cosmology with $\sum m_\nu = 0.058$ eV.

While of course a proper Bayesian inference is required to forecast the neutrino mass (as is done e.g. in Audren et al. 2013), we shall use `EuclidEmulator2` in order to investigate the uniqueness of the neutrino signal in the nonlinear matter power spectrum. To this end, we use a reference cosmology which has all parameters set identically to the Euclid Reference Cosmology except the sum of neutrino masses, which is set to 0.15 eV. We then try to fit the corresponding nonlinear power spectrum with a $w_0 w_a$CDM cosmology that has only massless neutrinos. Our goal is to fit the reference with an accuracy ≤ 1 per cent on all scales $0.01\,h\,\mathrm{Mpc}^{-1} \leq k \leq 10\,h\,\mathrm{Mpc}^{-1}$. We emphasize that we *do not* perform a proper forecasting by any means; we simply manually adjust all other cosmological parameters but $\sum m_\nu$ until we find a fit. It is worthwhile to note that such a procedure would not be practical without an emulator. The result of this procedure is shown in Fig. 18. We find that the nonlinear power spectra of the two cosmologies defined in Table 5 agree at a level of better than 1 per cent over the entire k range of interest at $z = 0$. We thus managed to find a cosmology (we call it 'fit') which is highly degenerate with the reference. The relative difference between the resulting power spectra is below the expected measurement accuracy of the *Euclid* mission and hence, *based on this information alone*, *Euclid* would not be able to tell these two cosmologies apart. However, taking the information from higher redshifts into account, the degeneracies are broken. This emphasizes the importance of weak lensing tomography for the *Euclid* survey in particular and of tomographic surveys in general.

We have further found yet another two different cosmologies (not shown) whose nonlinear power spectra fit that of the reference cosmology very well only at linear and only at nonlinear wave modes, respectively. This fact makes it very clear once again why modern cosmological surveys need to exploit as much information as possible from both regimes, linear and nonlinear.

10 CONCLUSIONS

For this work we have modified `PKDGRAV3` in such a way that DM is not only evolved fully nonlinearly due to self-interaction but also is subject to an additional gravity source due to massive neutrinos, radiation, DE, and the metric field perturbations. The latter four species themselves are, however, only evolved linearly. To this end, `PKDGRAV3` has been interfaced with `CONCEPT` and `CLASS`. While in older simulations, as those used in Euclid Collaboration (2019), the traditional back-scaling approach has been used for the construction of initial conditions of the N-body simulations, now we employ a novel approach taking advantage of the fully correct linear evolution

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Figure 18. Fit of a nonlinear power spectrum of w_0w_aCDM+$\sum m_\nu$ cosmology with a cosmology with only massless neutrinos. The purple line in the upper panel is a fit of the reference data to better than 1 per cent at all scales of interest. The different colours correspond to different redshifts. It is evident that even though the 'fit' cosmology approximates the reference very well at $z = 0$, the degeneracies between the neutrino mass and the other parameters are broken at different redshifts, highlighting the importance of tomographic surveys.

Table 5. Two cosmologies with nonlinear power spectra that agree to better than 1 per cent over all scales $0.01\,h\,\mathrm{Mpc}^{-1} \leq k \leq 10\,h\,\mathrm{Mpc}^{-1}$ at $z = 0$.

	reference	fit
Ω_b	0.049	0.049
Ω_m	0.3194	0.309
$\sum m_\nu$	0.15 eV	0.00 eV
n_s	0.96	0.97
h	0.67	0.67
w_0	− 1.0	− 1.0
w_a	0.0	0.0
A_s	2.1×10^{-9}	2.01×10^{-9}

of particles carried out in Einstein–Boltzmann codes (here CLASS). As a result, PKDGRAV3 recovers linear theory accurately at all redshifts even in the presence of massive neutrinos.

Moreover, we work with transfer functions in the *N*-body gauge instead of the more standard synchronous gauge. In this way, results computed with a purely Newtonian *N*-body code such as PKDGRAV3 can be interpreted within the framework of general relativity without the need of including general relativistic corrections at the *N*-body code level.

In a next step we have performed an extensive convergence study with the goal to pin down the smallest volume and the lowest mass resolution necessary in order simulate CDM+baryon NLC factors that have converged at the 1 per cent level all the way up to $k = 10\,h\,\mathrm{Mpc}^{-1}$. As references, we have used a simulation box of $L = 8192\,h^{-1}$ Mpc for the volume convergence series and a simulation of resolution $\ell^{-1} = N/L = 8\,h\,\mathrm{Mpc}^{-1}$ for the resolution convergence series. We identify $L = 1\,h^{-1}$ Gpc to be just barely enough for the side length of a simulation box necessary to achieve the 1 per cent target accuracy, although this is only true if pairing-and-fixing is used for the construction of the initial conditions. We find that resolution convergence at the targeted level of accuracy is increasingly difficult for higher redshifts. Even at $z = 0$ one only achieves the 1 per cent accuracy at $k = 10\,h\,\mathrm{Mpc}^{-1}$ with simulations of $\ell^{-1} > 4\,h\,\mathrm{Mpc}^{-1}$, which is beyond our capabilities given the minimal box size. From the convergence series one can further extrapolate that a mass resolution of roughly $\ell^{-1} \sim 6\,h\,\mathrm{Mpc}^{-1}$ is required to achieve convergence at the 1 per cent level at $k = 10\,h\,\mathrm{Mpc}^{-1}$ at $z \sim 3$. To put this in context we remind the reader that the Euclid Flagship v2.0 simulation (EFS2), using 4 trillion *N*-body particles, has resolution parameter of $\ell^{-1} = 4.4\,h\,\mathrm{Mpc}^{-1}$.

In order to correct for the power suppression at small scales resulting from the low mass resolution, we present a way to correct the power spectrum (and equivalently the NLC) curves using a cosmology-independent RCF, which can be applied in a post-processing step. The result of applying this correction to a power spectrum measured in a $\ell^{-1} = 3\,h\,\mathrm{Mpc}^{-1}$-simulation is a power spectrum that approximates very closely that obtained from an equivalent simulation with $\ell^{-1} = 8\,h\,\mathrm{Mpc}^{-1}$ up to $k \sim 10\,h\,\mathrm{Mpc}^{-1}$.

We have then produced a set of 127 P+F simulations of $1\mathrm{Gpc}^3/h^3$ with 3000^3 particles, corresponding to a resolution of $\ell^{-1} = 3\,h\,\mathrm{Mpc}^{-1}$. This corresponds to a computational cost of roughly 650 000 node hours which we have invested using the Piz Daint supercomputer located at the Swiss National Scientific Supercomputing Centre (CSCS). At redshift $z = 0$ this implies that the simulations are converged at the 2 per cent level at $k = 10\,h\,\mathrm{Mpc}^{-1}$ (and at 1 per cent up to $k \sim 5\,h\,\mathrm{Mpc}^{-1}$). At redshift $z = 2.76$ we achieve 1 per cent convergence up to $k \sim 2\,h\,\mathrm{Mpc}^{-1}$ and 2 per cent up to $k \sim 3.5\,h\,\mathrm{Mpc}^{-1}$. By applying the RCF, the convergence is subsequently improved to ∼0.5 per cent at $k \sim 10\,h\,\mathrm{Mpc}^{-1}$ and $z = 0$ and to ∼1 per cent at $k \sim 10\,h\,\mathrm{Mpc}^{-1}$ and $z = 2$ at the cost of introducing an additional source of uncertainty (see Fig. 4). We leave the decision about whether or not the resolution correction should be applied to the user of EuclidEmulator2 by not including it in the training data.

The key goal of this publication was to construct an emulator which is able to quickly and accurately predict the NLC for w_0w_aCDM+$\sum m_\nu$ cosmologies up to scales of $k \sim 10\,h\,\mathrm{Mpc}^{-1}$. The emulator takes inputs from within the parameter box defined in Table 2. In order to investigate the behaviour of such an emulator and its dependencies on various quantities such as training set size or number of principal components taken into account, we created a mock emulator based on HALOFIT data. We project that we can achieve a generalization error of sub-1 per cent inside the axis-aligned hyperellipsoid inscribed in the parameter box if we exclude a problematic region in the (w_0, w_a)-plane in which the first-order principal component weight shows exponential behaviour. We exclude this region from the training set by ignoring all cosmologies with $w_a > 0.5$, reducing the training set to 108 training examples.

Finally, we construct the actual emulator EuclidEmulator2 based on the PKDGRAV3 simulation data containing 108 training cosmologies. We train the emulator using the MATLAB package UQLab within only 9 s. The projected error of below 1 per cent up to scales of $k \sim 10\,h\,\mathrm{Mpc}^{-1}$ is confirmed with a small validation set. We stress that for smaller scales EuclidEmulator2 does

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

not allow the computation of the NLC. At these scales one has to fall back on suitable alternative methods as e.g. HALOFIT. Further, EuclidEmulator2 is compared to multiple other fast predictors such as its predecessor EuclidEmulator1 (Fig. 11), HALOFIT (Fig. 13), HMCode (Fig. 14), and CosmicEmu (Fig. 15). In all comparisons the error bounds as reported on in the corresponding publications have been respected. We have also performed a Sobol' sensitivity analysis (Fig. 8) which clearly revealed that w_a is a parameter that adds considerable complexity to the underlying model while $\sum m_\nu$ is quasi negligible, at least for the relatively narrow range in $\sum m_\nu$ we have considered.

In first benchmark tests using UQLab we have measured that EuclidEmulator2 can be evaluated in ∼0.3 s on a usual laptop. This compares well to the ∼0.4 s per evaluation of EuclidEmulator1 using the python wrapper e2py. We reiterate here that this implies that the computation of the linear power spectrum by CAMB or CLASS is now the bottleneck in the computation of the fully nonlinear power spectrum.

We have applied EuclidEmulator2 to investigate degeneracies of the nonlinear matter power spectrum between the total neutrino mass and the other seven cosmological parameters. We have shown that tomographic surveys exploiting both linear and nonlinear scales are critical as it is possible to find different cosmologies with nonlinear matter power spectra agreeing better than 1 per cent at $z = 0$ (in our case we have tested a massive neutrino and a massless neutrino cosmology).

Further efforts should be taken in multiple directions. While the power spectrum (and thus the NLC) clearly belong to the most used summary statistics of cosmic large-scale structure, higher order statistics are becoming more and more used and thus emulators for their prediction are desirable. An example for such a predictor was recently published (Takahashi et al. 2019). A different, more holistic approach is taken in He et al. (2019) where the displacement field is emulated directly, such that any statistic can be derived from the predicted density field. For simulations of resolution as high as the ones used in this work, it is however questionable to what extent such an approach is practical. Further, in order to assess more deeply the accuracy of the NLC predictions at small scales, it is not only necessary to estimate the generalization error and the convergence of the underlying simulations depending on box size and resolution but also to investigate how well different codes agree with each other at the scales under consideration. While such a study has been performed in Schneider et al. (2016) (and augmented by another code in Garrison et al. 2019) we advocate for new efforts in this direction, as with the new updates to PKDGRAV3 and developments in other codes the situation may have changed significantly.

Disregarding such uncertainties in the underlying *N*-body code, at this point we shall summarize the error contributions to EuclidEmulator2 and their dependence on spatial scales and redshift. At low redshifts, the emulation-only generalization error is virtually zero by construction on large scales ($k < 0.01\,h\,\mathrm{Mpc}^{-1}$) such that in this regime the dominant error contribution in the emulator comes from cosmic variance. Based on the results of Angulo & Pontzen (2016), the cosmic variance is expected to be sub-percent. At small scales ($k \gtrsim 1\,h\,\mathrm{Mpc}^{-1}$), cosmic variance is expected to be irrelevant. In this regime, the dominant error contribution (neglecting additional physics such as baryons) is due to emulation itself. The level of the dominant error at high k is estimated to be at the ∼0.7 per cent level according to Fig. 9. This error is estimated from only a very small sample of validation simulations, however, the error level is also consistent with the estimate in Fig. 6 and hence we regard this error estimate to be representative at small scales. Estimating the overall error level in the intermediate k range is tricky because several effects contribute errors at a similar level: on the one hand it is evident in Fig. 9 that an accurate prediction of the NLC around the BAOs is challenging (the observed accuracy is also at the level of ∼0.6 per cent). At the same time, residual cosmic variance (after pairing-and-fixing) is nonlinearly amplified at these scales. We estimate the error in the intermediate range ($0.01\,h\,\mathrm{Mpc}^{-1} < k < 1\,h\,\mathrm{Mpc}^{-1}$) to be at the level of 1 per cent. The comparison as shown in Fig. 11 suggests an error at the 2 per cent level. This may, however, be overly conservative because the cosmic variance in EuclidEmulator1 is phase-shifted with respect to EuclidEmulator2, leading to an enhancement of errors within this comparison.

The redshift evolution does not greatly change the error contributions discussed above. However, the overall error at intermediate scales is reduced at higher redshifts compared to the low-redshift case. At small scales the resolution effects become the dominant source of error as is visible in Fig. 4. As resolution is currently not corrected in a cosmology-dependent manner, the error is expected to be at the level of 1 per cent at small scales ($k \gtrsim 1\,h\,\mathrm{Mpc}^{-1}$) and higher redshift ($z \sim 3$).

Last but not least we have seen how the large number of dimensions of the parameter space is really starting to become a major challenge regarding the number of simulations required to arrive at the targeted generalization error. As more and more dimensions can be expected to be added in the next couple of years, it may be of interest to also compare different emulation strategies to each other in order to potentially identify the strategy that generalizes best based on only very few examples per dimension.

EuclidEmulator2 is the successor of EuclidEmulator1 and will again be published on https://github.com/miknab/EuclidEmulator2.

GLOSSARY

Codes:

CAMB	Code for anisotropies in the microwave background 5, 23
CLASS	Cosmological linear anisotropy solving system 3, 5-7, 15-18, 21, 23
CONCEPT	Cosmological *N*-body code in Python 5-7, 21
CosmicEmu	Cosmic emulator based on the Mira-Titan cosmological simulation suite (successor of FrankenEmu based on the Coyote simulation suite). 2, 16, 19, 20, 22, 29
EuclidEmulator1	Version 1 of EuclidEmulator. This is a code to predict the nonlinear corrections to DM power spectra for w_0CDM cosmologies. 2, 5-7, 14-19, 22, 23
EuclidEmulator2	Version 2 of EuclidEmulator. This is a code to predict the nonlinear corrections to DM power spectra for w_0, w_aCDM+$\sum m_\nu$. 2-5, 9-12, 14-23, 26, 29, 32
HALOFIT	Analytical code to produce nonlinear power spectra 2, 3, 10-12, 14-16, 18, 20, 22, 26, 29
HMCode	Fast predictor for the nonlinear power spectrum based on the halomodel approach. 2, 16, 18-20, 22, 29
NGenHalofit	Code to produce nonlinear power spectra using a semi-analytical approach for large and a smoothing-spline-fit model for small scales 2
	PKDGRAV3 parallel k-D tree gravity code (version 3); Cosmological *N*-body tree code 3, 5-8, 16, 17, 21-23
UQLab	MATLAB-based uncertainty quantification framework 12, 22, 23
e2py	Python wrapper for EuclidEmulator 23

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Acronyms:

1LPT	first-order Lagrangian perturbation theory 5, 8
2LPT	second-order Lagrangian perturbation theory 8
BAO	baryon accoustic oscillations 17-19, 21, 23
CDM	cold dark matter 3, 5, 6, 10, 19, 21, 23
CMB	cosmic microwave background 2, 10
DE	dark energy 2-6, 10, 12, 17, 18, 21
DM	dark matter 2-6, 21
ED	experimental design 9, 15, 16
EFS2	Euclid Flagship v2.0 Simulation 8, 17, 18, 22, 27
EoS	equation of state 4, 10, 12, 17
FMM	fast multipole method 5, 6
GR	general theory of relativity 4, 5
GRF	Gaussian random field 8, 27
IC	initial condition 5, 6, 8
LAR	Least angle regression 12
LH	Latin hypercube 12, 16
LHS	Latin hypercube sampling 12, 14, 15
LOO	leave-one-out 13
MCMC	Markov chain Monte Carlo 2, 11
ML	machine learning 13
NLC	nonlinear correction 2, 5-8, 10-13, 15-23
P+F	paired-and-fixed 5, 8, 15, 17, 18, 22, 26
PCA	principal component analysis 12-14
PCE	polynomial chaos expansion 12-15
PM	particle-mesh 5, 18
PPF	parametrized post-Friedmann 4
RCF	resolution correction factor 6, 9-11, 17, 27
UQ	uncertainty quantification 13

ACKNOWLEDGEMENTS

MK acknowledges support from the Swiss National Science Foundation (SNF) grant 200020_149848 and the Forschungskredit of the University of Zurich, grant no. K-76102-01-01. Simulations were performed on the PizDaint supercomputer at the Swiss National Scientific supercomputing center CSCS and on the zBox4+ cluster at the University of Zurich. The Euclid Consortium acknowledges a number of agencies and institutes that have supported the development of *Euclid*, in particular the Academy of Finland, the Agenzia Spaziale Italiana, the Belgian Science Policy, the Canadian Euclid Consortium, the Centre National d'Etudes Spatiales, the Deutsches Zentrum für Luft- und Raumfahrt, the Danish Space Research Institute, the Fundação para a Ciência e a Tecnologia, the Ministerio de Economia y Competitividad, the National Aeronautics and Space Administration, the Netherlandse Onderzoekschool Voor Astronomie, the Norwegian Space Agency, the Romanian Space Agency, the State Secretariat for Education, Research and Innovation (SERI) at the Swiss Space Office (SSO), and the United Kingdom Space Agency. A complete and detailed list is available on the *Euclid* web site (`http://www.euclid-ec.org`).

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Adamek J., Daverio D., Durrer R., Kunz M., 2016, J. Cosmol. Astropart. Phys., 2016, 053
Adamek J., Durrer R., Kunz M., 2017, J. Cosmol. Astropart. Phys., 2017, 004
Agarwal S., Feldman H. A., 2011, MNRAS, 410, 1647
Ahmed S. N. et al., 2004, Phys. Rev. Lett., 92, 181301
Akeson R. et al., 2019, preprint (arXiv:1902.05569)
Angulo R. E., Pontzen A., 2016, MNRAS, 462, L1
Angulo R. E., Zennaro M., Contreras S., Aricò G., Pellejero-Ibañez M., Stücker J., 2020, preprint (arXiv:2004.06245)
Audren B., Lesgourgues J., Bird S., Haehnelt M. G., Viel M., 2013, J. Cosmol. Astropart. Phys., 026, 23
Banerjee A., Powell D., Abel T., Villaescusa-Navarro F., 2018, J. Cosmol. Astropart. Phys., 2018 028
Becker-Szendy R. et al., 1992, Phys. Rev. D, 46, 3720
Betoule M. et al., 2014, A&A, 568, 32
Bird S., Viel M., Haehnelt M. G., 2012, MNRAS, 420, 2551
Bird S., Ali-Haïmoud Y., Feng Y., Liu J., 2018, MNRAS, 481, 1486
Blas D., Lesgourgues J., Tram T., 2011, J. Cosmol. Astropart. Phys., 07, 34
Blatman G., 2009, PhD thesis, Université de Clermont-Ferrand
Blatman G., Sudret B., 2009, in 19ème Congrès Français de Mécanique. AFM, Maison de la Mécanique, 39/41 rue Louis Blanc - 92400 Courbevoie, Marseille, p. 1
Blatman G., Sudret B., 2010, Probabilistic Eng. Mech., 25, 183
Blatman G., Sudret B., 2011, J. Comput. Phys., 230, 2345
Blatman G., Sudret B., 2013, in Deodatis G., Ellingwood B., Frangopol D., eds, Safety, Reliability, Risk and Life-Cycle Performance of Structures and Infrastructures - Proceedings of the 11th International Conference on Structural Safety and Reliability, ICOSSAR 2013. CRC Press (Taylor & Francis Group), London, p. 3245
Brandbyge J., Rampf C., Tram T., Leclercq F., Fidler C., Hannestad S., 2016, MNRAS, 466, L68–L72
Chisari N. E., Zaldarriaga M., 2011, Phys. Rev. D, 83, 123505
Copeland D., Taylor A., Hall A., 2020, MNRAS, 493, 1640
Coulton W. R., Liu J., Madhavacheril M. S., Böhm V., Spergel D. N., 2019, J. Cosmol. Astropart. Phys., 2019, 043
Crombecq K., Laermans E., Dhaene T., 2011, Eur. J. Oper. Res., 214, 683
Dakin J., 2015, preprint (arXiv:1510.07621)
Dakin J., Hannestad S., Tram T., 2019a, J. Cosmol. Astropart. Phys., 06, 1
Dakin J., Hannestad S., Tram T., Knabenhans M., Stadel J., 2019b, J. Cosmol. Astropart. Phys., 08, 1
Damblin G., Couplet M., Iooss B., 2013, J. Simul., 7, 276
Dawson K. S. et al., 2016, AJ, 151, 44
DeRose J. et al., 2019, ApJ, 875, 69
DESI Collaboration, 2016, preprint (arXiv:1611.00036)
DUNE Collaboration, 2015, preprint (arXiv:1512.06148)
Efron B., Hastie T., Johnstone I. et al., 2004, Ann. Stat., 32, 407
Euclid Collaboration, 2019, MNRAS, 484, 5509
Fang W., Hu W., Lewis A., 2008, Phys. Rev. D, 78, 4
Fidler C., Rampf C., Tram T., Crittenden R., Koyama K., Wands D., 2015, Phys. Rev. D, 92 123517
Fraenkle F., 2008, J. Phys. Conf. Ser., 120, 052028
Fukuda Y. et al., 1998a, Phys. Rev. Lett., 81, 1158
Fukuda Y. et al., 1998b, Phys. Rev. Lett., 81, 1562
Garg V. V., Stogner R. H., 2017, J. Am. Stat. Assoc., 112, 673
Garrison L. H., Eisenstein D. J., Ferrer D., Metchnik M. V., Pinto P. A., 2016, MNRAS, 461, 4125
Garrison L. H., Eisenstein D. J., Ferrer D., Tinker J. L., Pinto P. A., Weinberg D. H., 2018, ApJS, 236, 43
Garrison L. H., Eisenstein D. J., Pinto P. A., 2019, MNRAS, 485, 3370
Giblin B., Cataneo M., Moews B., Heymans C., 2019, MNRAS, 490, 4826
Gration A., Wilkinson M. I., 2019, MNRAS, 485, 4878
Hannestad S., Haugbølle T. C. S., 2012, J. Cosmol. Astropart. Phys., 2012, 1
Hannestad S., Upadhye A., Wong Y., 2020, J. Cosmol. Astropart. Phys., 11, 62

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

He S., Li Y., Feng Y., Ho S., Ravanbakhsh S., Chen W., Póczos B., 2019, Proc. Natl. Acad. Sci. USA, 116, 13825
Heitmann K., Higdon D., White M., Habib S., Williams B. J., Lawrence E., Wagner C., 2009, ApJ, 705, 156
Heitmann K., White M., Wagner C., Habib S., Hingon D., 2010, ApJ, 715, 104
Heitmann K., Lawrence E., Kwan J., Habib S., Hindon D., 2014, ApJ, 780, 111
Heitmann K. et al., 2016, ApJ, 820, 108
Hu W., Sawicki I., 2007, Phys. Rev. D, 76, 104043
Ichiki K., Takada M., Takahashi T., 2009, Phys. Rev. D, 79 023520
Kuijken K. et al., 2019, A&A, 625, A2
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Lawrence E., Heitmann K., White M., Higdon D., Wagner C., Habib S., Williams B., 2010, ApJ, 713, 1322
Lawrence E. et al., 2017, ApJ, 847, 12
Lesgourgues J., Tram T., 2011, J. Cosmol. Astropart. Phys., 9, 32
Liefvendahl M., Stocki R., 2006, J. Stat. Plan. Inference, 136, 3231
LSST Science Collaboration, 2009, preprint (arXiv:0912.0201)
Marelli S., Sudret B., 2014, in Vulnerability, Uncertainty, and Risk (Proc. 2nd Int. Conf. on Vulnerability, Risk Analysis and Management (ICVRAM2014). American Society of Civil Engineers, Liverpool, United Kingdom, p. 2554
Marelli S., Sudret B., 2017, Technical report, UQLab user manual - Polynomial Chaos Expansion. Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich, Zurich
Marelli S., Sudret B., 2018, Struct. Saf., 75, 67
Marelli S., Lamas C., Sudret B., 2017, Technical report, UQLab user manual - Sensitivity analysis. Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich, Zurich
Massara E., Villaescusa-Navarro F., Viel M., 2014, J. Cosmol. Astropart. Phys., 12, 053
McClintock T. et al., 2019, ApJ, 872, 53
McKay M. D., Beckmann R. J., Conover W. J., 1979, Technometrics, 21, 239
Mead A. J., Heymans C., Lombriser L., Peacock J. A., Steele O. I., Winther H. A., 2016, MNRAS, 459, 1468
Nishimichi T. et al., 2019, ApJ, 884, 29
Particle Data Group: Patrignani C. et al., 2016, Chin. Phys. C, 40, 1
Particle Data Group: Tanabashi M. et al., 2018, Phys. Rev. D, 98, 030001
Planck Collaboration, 2006, preprint (arXiv:astro-ph/0604069)
Planck Collaboration XIII, 2015, A&A, 594, 63
Potter D., Stadel J. G., 2016, Astrophysics Source Code Library, record ascl:1102.027
Potter D., Stadel J. G., Teyssier R., 2017, Comput. Astrophys. Cosmol., 4, 2
Rizzo L. A., Villaescusa-Navarro F., Monaco P., Munari E., Borgani S., Castorina E., Sefusatti E., 2017, J. Cosmology Astropart. Phys., 2017, 008
Rogers K. K., Peiris H. V., Pontzen A., Bird S., Verde L., Font-Ribera A., 2019, J. Cosmol. Astropart. Phys., 02, 1
Schneider A. et al., 2016, J. Cosmol. Astropart. Phys., 04, 047
Schneider A., Stoira N., Refregier A., Weiss A. J., Knabenhans M., Stadel J., Teyssier R., 2020a, J. Cosmol. Astropart. Phys., 4, 19
Schneider A. et al., 2020b, J. Cosmol. Astropart. Phys., 4, 20
Schwetz T., Tórtola M., Valle J. W. F., 2008, New J. Phys., 10, 113011
Sheikholeslami R., Razavi S., 2017, Environ. Modelling Softw., 93, 109
Smith R. E., Angulo R. E., 2019, MNRAS, 486, 1448
Springel V., Pakmor R., Zier O., Reinecke M., 2020, MNRAS preprint (arXiv: 2010.03567)
Stadel J. G., 2001, PhD thesis, University of Washington. Available at: https://search.proquest.com/docview/251727174
Swiler L. P., Slepoy R., Giunta A., 2006, in 47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. Newport, Rhode Island, p. 1 doi:10.2514/6.2006-1827
Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152
Takahashi R., Nishimichi T., Namikawa T., Taruya A., Kayo I., Osato K., Kobayashi Y., Shirasaki M., 2020, ApJ, 895, 113
Tang B., 1993, J. Am. Stat. Assoc., 88, 1392
Teyssier R., 2002, A&A, 385, 337
Torre E., Marelli S., Embrechts P., Sudret B., 2019, J. Comput. Phys., 388, 601
Tram T., Brandbyge J., Dakin J., Hannestad S., 2019, J. Cosmol. Astropart. Phys., 3, 1
Valcin D., Villaescusa-Navarro F., Verde L., Raccanelli A., 2019, J. Cosmol. Astropart. Phys., 2019, 057
Valle J. W. F., 2005, in Kiwoon C., Son J. E., Son D., eds, AIP Conf. Proc. Vol. 805, PArticles, Strings, and Cosmology: 11th International Symposium on Particles, Strings, and Cosmology; PASCOS 2005. Am. Inst. Phys., New York, p.128
Viel M., Haehnelt M. G., Springel V., 2010, J. Cosmol. Astropart. Phys., 06, 15
Wang G. G., 2003, J. Mech. Design, 125, 210
Weinheimer C., 2002, Prog. Part. Nucl. Phys., 48, 141
Winther H. A., Casas S., Baldi M., Kazuya K., Baojiu L., Lucas L., Gong-Bo Z., 2019, Phys. Rev. D, 100, 1
Wolf J., Katrin Collaboration, 2010, Nucl. Instrum. Methods Phys. Res., 623, 442
Xiu D., Karniadakis G. E., 2002, SIAM J. Sci. Comput., 24, 619
Yang M., Liu Z., Li W., 2017, J. Stat. Comput. Simul., 87, 3398
Zhai Z. et al., 2019, ApJ, 874, 12

APPENDIX A: SIMULATION TABLE

Here, we summarize all simulations that we have produced specifically for this paper (see Table A1). For each simulation, its unique ID as well as its specifications are listed. The specifications consist of the box size (L), the number of particles per side length (N), whether it is a P+F run (PF yes/no), what order of Lagrangian perturbation theory (LPT) was used to construct the initial conditions, the number of runs, the run time in node hours and on what machine the simulation was executed. Simulations T001 to T127 are the runs that form the actual training set of `EuclidEmulator2`, while HRV001–HRV003 were used for the end-to-end test reported in Section 8.2.1. The runs VCT1-VT5 were used for the volume convergence test and RCT1-RCT5 for the resolution convergence test (see Fig. 2 and Fig. 3). We used the RES3 and RES8 simulations in order to estimate the variance of the cosmology dependence in the RCF (see Fig. 4). The PF simulation was used in the comparison to the simulations GRF1–GRF50 in order to investigate the cosmic sampling variance in P+F simulations (see Fig. 12). The PV runs were used to estimate the output variance on both boost factor and power spectrum level when one of the parameters A_s, w_a, or $\sum m_\nu$ is varied based on which the parameter box of `EuclidEmulator2` was chosen. The total run time for all simulations sums up to over 700 000 node hours.

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Table A1. Simulations used for this publication. The table is organized as follows: The training simulations are in the first row followed by the validation simulations in the second and third row. In rows 4 to 13 we list all simulations used in the volume convergence tests (VCT) and the resolution convergence tests (RCT), respectively. Then, in rows 14 and 15 we list the simulations used to investigate the RCF dependence on cosmology. In rows 16 and 17 the simulations used to compare the P+F approach to the traditional ensemble averaging of GRF simulations are listed. Ultimately, we mention the simulations used to estimate the parameter ranges in the rows 18–21.

Simulation identifier	L [h^{-1} Mpc]	N	PF	LPT	Number of runs	Total runtime [node hours]	Machine
T001–T127	1000	3000	yes	1LPT	254	∼500 000	Piz Daint (⋆)
EFS2	3600	16000	no	1LPT	1	∼1000 000	Piz Daint (⋆)
HRV001–HRV003	1000	3000	yes	1LPT	6	∼37 000	zBox4+
VCT1	512	170	yes	1LPT	2	10	zBox4+
VCT2	1024	342	yes	1LPT	2	24	zBox4+
VCT3	2048	682	yes	1LPT	2	118	zBox4+
VCT4	4096	1356	yes	1LPT	2	435	zBox4+
VCT5	8192	2730	yes	1LPT	2	2780	zBox4+
RCT1	512	512	no	1LPT	1	37	zBox4+
RCT2	512	1024	no	1LPT	1	212	zBox4+
RCT3	512	1536	no	1LPT	1	919	zBox4+
RCT4	512	2048	no	1LPT	1	1987	zBox4+
RCT5	512	4046	no	1LPT	1	10 353	zBox4+
RES3_1-RES3_20	128	384	yes	1LPT	40	∼320	zBox4+
RES8_1-RES8_20	128	1024	yes	1LPT	40	∼4800	zBox4+
PF	1024	980	yes	1LPT	2	69	zBox4+
GRF1-GRF50	1024	980	no	1LPT	50	∼3400	zBox4+
PV_{A_s}	640	1024	no	2LPT	6	∼1100	zBox4+
PV_{w_a}	640	1024	no	2LPT	6	∼1100	zBox4+
$\mathrm{PV}_{\sum m_\nu}$	640	1024	no	2LPT	6	∼670	zBox4+
$\mathrm{PV}_{\mathrm{center}}$	640	1024	no	2LPT	1	77	zBox4+

Note. ⋆with GPUs

APPENDIX B: ERROR MAPS OF THE HALOFIT-BASED MOCK EMULATOR

In this appendix we plot error maps for a two exemplary coordinate planes of the 8D parameter box: the $(\sum m_\nu, h)$- and the (w_0, w_a)-plane (see Fig. B1). The errors are defined as follows:

$$\varepsilon = \left| \max_{k\in[0.01,10.0]\,h\,\mathrm{Mpc}^{-1}} \left(\frac{B_{\mathrm{emu}}(k, z=0) - B_{\mathrm{sim}}(k, z=0)}{B_{\mathrm{sim}}(k, z=0)} \right) \right| \tag{B1}$$

The emulator for this investigation was trained with HALOFIT based on exactly the same 108 cosmologies that were used for the actual, simulation-based EuclidEmulator2. We stress that the errors all are measured at $z = 0$. The colour bars are ranging from 0 per cent to 8 per cent for both plot panels. The hyperellipsoid inscribed in the parameter box is shown. Notice that the vast majority of cosmologies inside this region features errors at the 2 per cent level (or even lower). There are, however, also regions with larger errors. This does not contradict the result reported on in Fig. 6 as the error metric in that figure was averaged over all cosmologies. It is not surprising that particularly validation cosmologies with $w_a > 0.5$ feature fairly large errors as in this region there are no training cosmologies. This cut is indicated by a grey, dashed line in all plots with w_a as one of the two dimensions. All parameter planes that have w_a as one of the two dimensions exhibit larger errors for larger values of w_a.

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Figure B1. Error map of the (Σm_ν, h)-plane (top) and the (w_0, w_a)-plane (bottom). These two error maps represent also those of the remaining 26 parameter planes. Most error maps feature only very low errors like the top panel (all errors in the top panel are $\lesssim$1 per cent) in this figure.

APPENDIX C: LIST OF VALIDATION COSMOLOGIES

In this appendix we list the validation cosmologies in Table C1 and visualize their locations inside the parameter space in Fig. C1.

Figure C1. Distribution of 84 validation cosmologies listed in Table C1. Notice that some of the validation cosmologies are located outside the hyperellipse (grey). In fact, some of the validation cosmologies even have a w_a value above 0.5, indicating that `EuclidEmulator2` is even able to extrapolate into this region.

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Table C1. List of validation cosmologies used in Sections 8.2.3, 8.2.4, and 8.2.5 to compare `EuclidEmulator2` with `HALOFIT`, `HMCode`, and `CosmicEmu`.

Ω_b	Ω_m	$\sum m_\nu$ [eV]	n_s	h	w_0	w_a	A_s
5.23×10^{-2}	2.82×10^{-1}	6.70×10^{-2}	9.95×10^{-1}	6.67×10^{-1}	-8.16×10^{-1}	2.66×10^{-1}	2.44×10^{-9}
5.64×10^{-2}	3.40×10^{-1}	1.50×10^{-1}	9.43×10^{-1}	6.30×10^{-1}	-1.19×10^{-2}	-5.44×10^{-1}	1.74×10^{-9}
5.73×10^{-2}	3.38×10^{-1}	4.27×10^{-2}	9.29×10^{-1}	6.32×10^{-1}	-7.46×10^{-1}	-1.08×10^{-1}	2.23×10^{-9}
4.73×10^{-2}	3.14×10^{-1}	2.66×10^{-2}	9.29×10^{-1}	6.93×10^{-1}	-9.88×10^{-1}	5.95×10^{-1}	1.91×10^{-9}
5.51×10^{-2}	3.54×10^{-1}	1.57×10^{-3}	9.49×10^{-1}	6.45×10^{-1}	-8.96×10^{-1}	6.70×10^{-1}	2.15×10^{-9}
5.41×10^{-2}	3.16×10^{-1}	7.48×10^{-2}	9.82×10^{-1}	6.39×10^{-1}	-8.59×10^{-1}	-1.33×10^{-1}	2.36×10^{-9}
5.22×10^{-2}	3.46×10^{-1}	5.33×10^{-2}	9.56×10^{-1}	6.65×10^{-1}	-1.01×10^{-2}	-3.17×10^{-1}	1.91×10^{-9}
5.21×10^{-2}	3.30×10^{-1}	7.66×10^{-2}	9.84×10^{-1}	6.70×10^{-1}	-1.12×10^{-2}	-3.68×10^{-1}	1.95×10^{-9}
4.40×10^{-2}	2.97×10^{-1}	1.11×10^{-1}	9.75×10^{-1}	7.12×10^{-1}	-8.76×10^{-1}	8.45×10^{-2}	2.22×10^{-9}
4.58×10^{-2}	2.72×10^{-1}	8.09×10^{-2}	9.78×10^{-1}	7.10×10^{-1}	-7.26×10^{-1}	-6.27×10^{-1}	1.85×10^{-9}
5.25×10^{-2}	3.44×10^{-1}	3.44×10^{-2}	9.37×10^{-1}	6.53×10^{-1}	-1.06×10^{-2}	6.64×10^{-1}	2.24×10^{-9}
4.37×10^{-2}	2.96×10^{-1}	7.36×10^{-2}	9.25×10^{-1}	7.19×10^{-1}	-1.04×10^{-2}	4.92×10^{-1}	2.19×10^{-9}
4.59×10^{-2}	3.24×10^{-1}	8.37×10^{-2}	9.20×10^{-1}	6.90×10^{-1}	-9.16×10^{-1}	3.53×10^{-1}	2.35×10^{-9}
5.32×10^{-2}	3.22×10^{-1}	9.43×10^{-2}	9.53×10^{-1}	6.57×10^{-1}	-1.26×10^{-2}	-2.22×10^{-1}	1.92×10^{-9}
5.47×10^{-2}	3.81×10^{-1}	7.08×10^{-2}	9.56×10^{-1}	6.35×10^{-1}	-7.70×10^{-1}	2.96×10^{-1}	2.10×10^{-9}
4.88×10^{-2}	2.78×10^{-1}	1.46×10^{-1}	9.99×10^{-1}	6.75×10^{-1}	-1.26×10^{-2}	-2.12×10^{-1}	2.01×10^{-9}
5.17×10^{-2}	3.06×10^{-1}	1.35×10^{-1}	9.41×10^{-1}	6.48×10^{-1}	-7.20×10^{-1}	-4.10×10^{-1}	2.26×10^{-9}
4.84×10^{-2}	3.29×10^{-1}	2.41×10^{-2}	9.92×10^{-1}	6.81×10^{-1}	-7.91×10^{-1}	-1.57×10^{-1}	1.92×10^{-9}
4.80×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.20×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	2.72×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	2.88×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.04×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.36×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	0.00	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	1.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	3.00×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	4.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	6.00×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	9.00×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	1.05×10^{-1}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	1.20×10^{-1}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	1.35×10^{-1}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	1.50×10^{-1}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.20×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.28×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.36×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.44×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.52×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.68×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.76×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.84×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.92×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	1.00×10^{-2}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.58×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.82×10^{-1}	-1.00×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.30×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.24×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.18×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.12×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.06×10^{-2}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-9.40×10^{-1}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-8.80×10^{-1}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-8.20×10^{-1}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-7.60×10^{-1}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-7.00×10^{-1}	0.00	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	-7.00×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	-5.60×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	-4.20×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	-2.80×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	-1.40×10^{-1}	2.10×10^{-9}

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

Table C1 – *continued*

Ω_b	Ω_m	$\sum m_\nu$ [eV]	n_s	h	w_0	w_a	A_s
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	1.40×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	2.80×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	4.20×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	5.60×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	7.00×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	1.70×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	1.78×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	1.86×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	1.94×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.02×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.18×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.26×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.34×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.42×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	7.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.00×10^{-2}	0.00	2.50×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	0.00	9.60×10^{-1}	6.70×10^{-1}	-1.30×10^{-2}	-7.00×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	1.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.24×10^{-2}	-5.60×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	3.00×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.18×10^{-2}	-4.20×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	4.50×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.12×10^{-2}	-2.80×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	6.00×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-1.06×10^{-2}	-1.40×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	9.00×10^{-2}	9.60×10^{-1}	6.70×10^{-1}	-9.40×10^{-1}	1.40×10^{-1}	2.10×10^{-9}
5.00×10^{-2}	3.20×10^{-1}	1.05×10^{-1}	9.60×10^{-1}	6.70×10^{-1}	-8.80×10^{-1}	2.80×10^{-1}	2.10×10^{-9}

[1]*Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland*
[2]*Department of Physics and Astronomy, University of Aarhus, Ny Munkegade 120, DK-8000 Aarhus C, Denmark*
[3]*Chair of Risk, Safety and Uncertainty Quantification, Department of Civil Engineering, ETH Zurich, Stefano-Franscini-Platz 5, CH-8093 Zurich, Switzerland*
[4]*Institute for Particle Physics and Astrophysics, Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 27, CH-8093 Zurich, Switzerland*
[5]*Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, E-08193 Barcelona, Spain*
[6]*INAF-Osservatorio Astronomico di Brera, Via Brera 28, I-20122 Milano, Italy*
[7]*INAF-Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, I-40129 Bologna, Italy*
[8]*SISSA, International School for Advanced Studies, Via Bonomea 265, I-34136 Trieste TS, Italy*
[9]*INFN, Sezione di Trieste, Via Valerio 2, I-34127 Trieste TS, Italy*
[10]*INAF-Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, I-34131 Trieste, Italy*
[11]*Universidad de la Laguna, E-38206 San Cristóbal de La Laguna, Tenerife, Spain*
[12]*Instituto de Astrofísica de Canarias, Calle Vía Làctea s/n, E-38204 San Cristóbal de la Laguna, Tenerife, Spain*
[13]*Istituto Nazionale di Astrofisica (INAF) - Osservatorio di Astrofisica e Scienza dello Spazio (OAS), Via Gobetti 93/3, I-40127 Bologna, Italy*
[14]*Dipartimento di Fisica e Astronomia, Universitá di Bologna, Via Gobetti 93/2, I-40129 Bologna, Italy*
[15]*INFN-Sezione di Bologna, Viale Berti Pichat 6/2, I-40127 Bologna, Italy*
[16]*INAF-IASF Bologna, Via Piero Gobetti 101, I-40129 Bologna, Italy*
[17]*Fakultät für Physik, Universitäts-Sternwarte München, Ludwig-Maximilians-Universität München, Scheinerstrasse 1, D-81679 München, Germany*
[18]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstr 1, D-85748 Garching, Germany*
[19]*IFPU, Institute for Fundamental Physics of the Universe, via Beirut 2, I-34151 Trieste, Italy*
[20]*Department of Astronomy, University of Geneva, ch. d'Écogia 16, CH-1290 Versoix, Switzerland*
[21]*INFN-Sezione di Roma Tre, Via della Vasca Navale 84, I-00146 Roma, Italy*
[22]*Department of Mathematics and Physics, Roma Tre University, Via della Vasca Navale 84, I-00146 Rome, Italy*
[23]*INAF-Osservatorio Astronomico di Roma, Via Frascati 33, I-00078 Monteporzio Catone, Italy*
[24]*INAF-Osservatorio Astronomico di Capodimonte, Via Moiariello 16, I-80131 Napoli, Italy*
[25]*Dipartimento di Fisica e Scienze della Terra, Universitá degli Studi di Ferrara, Via Giuseppe Saragat 1, I-44122 Ferrara, Italy*
[26]*INAF, Istituto di Radioastronomia, Via Piero Gobetti 101, I-40129 Bologna, Italy*
[27]*Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, CNRS, UPS, CNES, 14 Av. Edouard Belin, F-31400 Toulouse, France*
[28]*INFN-Sezione di Torino, Via P. Giuria 1, I-10125 Torino, Italy*
[29]*Dipartimento di Fisica, Universitá degli Studi di Torino, Via P. Giuria 1, I-10125 Torino, Italy*
[30]*INAF-Osservatorio Astrofisico di Torino, Via Osservatorio 20, I-10025 Pino Torinese (TO), Italy*
[31]*Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, F-06304 Nice cedex 4, France*
[32]*INAF-IASF Milano, Via Alfonso Corti 12, I-20133 Milano, Italy*
[33]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra (Barcelona), Spain*
[34]*Faculdade de Ciências, Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa, Tapada da Ajuda, P-1349-018 Lisboa, Portugal*
[35]*Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain*
[36]*AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, F-91191 Gif-sur-Yvette, France*
[37]*Observatoire de Sauverny, Ecole Polytechnique Fédérale de Lau- sanne, CH-1290 Versoix, Switzerland*
[38]*Department of Physics 'E. Pancini', University Federico II, Via Cinthia 6, I-80126 Napoli, Italy*
[39]*INFN section of Naples, Via Cinthia 6, I-80126 Napoli, Italy*
[40]*Centre National d'Etudes Spatiales, Toulouse, France*
[41]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021

[42]*University of Nottingham, University Park, Nottingham NG7 2RD, UK*
[43]*European Space Agency/ESRIN, Largo Galileo Galilei 1, I-00044 Frascati, Roma, Italy*
[44]*ESAC/ESA, Camino Bajo del Castillo, s/n., Urb. Villafranca del Castillo, E-28692 Villanueva de la Cañada, Madrid, Spain*
[45]*Univ Lyon, Univ Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, UMR 5822, F-69622 Villeurbanne, France*
[46]*University of Lyon, UCB Lyon 1, CNRS/IN2P3, IUF, IP2I Lyon, France*
[47]*Faculdade de Ciências, Departamento de Física, Universidade de Lisboa, Edifício C8, Campo Grande, P-1749-016 Lisboa, Portugal*
[48]*Faculdade de Ciências, Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa, P-1749-016 Lisboa, Portugal*
[49]*Aix-Marseille Univ, CNRS, CNES, LAM, Marseille, France*
[50]*Department of Physics, Oxford University, Keble Road, Oxford OX1 3RH, UK*
[51]*Department of Physics & Astronomy, University of Sussex, Brighton BN1 9QH, UK*
[52]*INFN-Bologna, Via Irnerio 46, I-40126 Bologna, Italy*
[53]*Department of Physics, University of Helsinki, PO Box 64, FI-00014 Helsinki, Finland*
[54]*Department of Physics and Helsinki Institute of Physics, University of Helsinki, Gustaf Hällströmin katu 2, FI-00014 Helsinki, Finland*
[55]*Dipartimento di Fisica 'Aldo Pontremoli', Universitá degli Studi di Milano, Via Celoria 16, I-20133 Milano, Italy*
[56]*INFN-Sezione di Milano, Via Celoria 16, I-20133 Milano, Italy*
[57]*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA*
[58]*von Hoerner & Sulger GmbH, SchloßPlatz 8, D-68723 Schwetzingen, Germany*
[59]*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[60]*Aix-Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France*
[61]*Institut de Physique Nucléaire de Lyon, 4, rue Enrico Fermi, F-69622 Villeurbanne cedex, France*
[62]*Département de Physique Théorique and Centre for Astroparticle Physics, Université de Genève, 24 quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland*
[63]*Institute of Theoretical Astrophysics, University of Oslo, PO Box 1029 Blindern, N-0315 Oslo, Norway*
[64]*NOVA optical infrared instrumentation group at ASTRON, Oude Hoogeveensedijk 4, NL-7991PD Dwingeloo, the Netherlands*
[65]*Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany*
[66]*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*
[67]*Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK*
[68]*Space Science Data Center, Italian Space Agency, via del Politecnico snc, I-00133 Roma, Italy*
[69]*INFN-Padova, Via Marzolo 8, I-35131 Padova, Italy*
[70]*Dipartimento di Fisica e Astronomia 'G.Galilei', Universitá di Padova, Via Marzolo 8, I-35131 Padova, Italy*
[71]*Departamento de Física, FCFM, Universidad de Chile, Blanco Encalada 2008, Santiago, Chile*
[72]*Institut d'Astrophysique de Paris, 98bis Boulevard Arago, F-75014 Paris, France*
[73]*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Avenida Complutense 40, E-28040 Madrid, Spain*
[74]*Departamento de Electrónica y Tecnología de Computadoras, Universidad Politécnica de Cartagena, E-30202 Cartagena, Spain*
[75]*Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA*
[76]*Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK*
[77]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK*

This paper has been typeset from a TEX/LATEX file prepared by the author.

Downloaded from https://academic.oup.com/mnras/article/505/2/2840/6275725 by California Institute of Technology user on 04 August 2021