Euclid preparation. Simulations and nonlinearities beyond $Λ$CDM.
4. Constraints on $f(R)$ models from the photometric primary probes

Astronomy & Astrophysics manuscript no. KPJC6_paper4
September 6, 2024

Euclid preparation

Simulations and nonlinearities beyond ΛCDM. 4. Constraints on $f(R)$ models from the photometric primary probes

Euclid Collaboration: K. Koyama⋆[1], S. Pamuk[2], S. Casas[2], B. Bose[3], P. Carrilho[3], I. Sáez-Casares[4], L. Atayde[5,6], M. Cataneo[7,8], B. Fiorini[1], C. Giocoli[9,10], A. M. C. Le Brun[4], F. Pace[11,12,13], A. Pourtsidou[3,14], Y. Rasera[4,15], Z. Sakr[16,17,18], H.-A. Winther[19], E. Altamura[20], J. Adamek[21], M. Baldi[22,9,23], M.-A. Breton[24,25,4], G. Rácz[26], F. Vernizzi[27], A. Amara[28], S. Andreon[29], N. Auricchio[9], C. Baccigalupi[30,31,32,33], S. Bardelli[9], F. Bernardeau[27,34], C. Bodendorf[35], D. Bonino[13], E. Branchini[36,37,29], M. Brescia[38,39,40], J. Brinchmann[41,42], A. Caillat[43], S. Camera[11,12,13], V. Capobianco[13], C. Carbone[44], J. Carretero[45,46], M. Castellano[47], G. Castignani[9], S. Cavuoti[39,40], A. Cimatti[48], C. Colodro-Conde[49], G. Congedo[3], C. J. Conselice[20], L. Conversi[50,51], Y. Copin[52], F. Courbin[53,54,55], H. M. Courtois[56], A. Da Silva[6,5], H. Degaudenzi[57], G. De Lucia[31], M. Douspis[58], F. Dubath[57], C. A. J. Duncan[20], X. Dupac[51], S. Dusini[59], M. Farina[60], S. Farrens[61], S. Ferriol[52], P. Fosalba[62,24], M. Frailis[31], E. Franceschi[9], S. Galeotta[31], B. Gillis[3], P. Gómez-Alvarez[63,51], A. Grazian[64], F. Grupp[35,65], L. Guzzo[66,29], M. Hailey[67], S. V. H. Haugan[19], W. Holmes[26], F. Hormuth[68], A. Hornstrup[69,70], P. Hudelot[34], S. Ilić[71,17], K. Jahnke[72], M. Jhabvala[73], B. Joachimi[74], E. Keihänen[75], S. Kermiche[76], A. Kiessling[26], M. Kilbinger[61], B. Kubik[52], M. Kunz[77], H. Kurki-Suonio[78,79], P. B. Lilje[19], V. Lindholm[78,79], I. Lloro[80], G. Mainetti[81], D. Maino[66,44,82], E. Maiorano[9], O. Mansutti[31], O. Marggraf[8], K. Markovic[26], M. Martinelli[47,83], N. Martinet[43], F. Marulli[84,9,23], R. Massey[85], E. Medinaceli[9], S. Mei[86], M. Melchior[87], Y. Mellier[88,34], M. Meneghetti[9,23], E. Merlin[47], G. Meylan[53], M. Moresco[84,9], L. Moscardini[84,9,23], E. Munari[31,30], C. Neissner[89,46], S.-M. Niemi[90], C. Padilla[89], S. Paltani[57], F. Pasian[31], K. Pedersen[91], W. J. Percival[92,93,94], V. Pettorino[90], S. Pires[61], G. Polenta[95], M. Poncet[96], L. A. Popa[97], L. Pozzetti[9], F. Raison[35], A. Renzi[98,59], J. Rhodes[26], G. Riccio[39], E. Romelli[31], M. Roncarelli[9], R. Saglia[65,35], J.-C. Salvignol[90], A. G. Sánchez[35], D. Sapone[99], B. Sartoris[65,31], M. Schirmer[72], T. Schrabback[100], A. Secroun[76], G. Seidel[72], S. Serrano[62,101,24], C. Sirignano[98,59], G. Sirri[23], L. Stanco[59], J. Steinwagner[35], P. Tallada-Crespí[45,46], A. N. Taylor[3], I. Tereno[6,102], R. Toledo-Moreo[103], F. Torradeflot[46,45], I. Tutusaus[17], L. Valenziano[9,104], T. Vassallo[65,31], G. Verdoes Kleijn[105], A. Veropalumbo[29,37,106], Y. Wang[107], J. Weller[65,35], G. Zamorani[9], E. Zucca[9], A. Biviano[31,30], E. Bozzo[57], C. Burigana[108,104], M. Calabrese[109,44], D. Di Ferdinando[23], J. A. Escartin Vigo[35], G. Fabbian[110,111], R. Farinelli[9], F. Finelli[9,104], J. Gracia-Carpio[35], S. Matthew[3], N. Mauri[48,23], A. Pezzotta[35], M. Pöntinen[78], V. Scottez[88,112], M. Tenti[23], M. Viel[30,31,33,32,113], M. Wiesmann[19], Y. Akrami[114,115], S. Anselmi[59,98,4], M. Archidiacono[66,82], F. Atrio-Barandela[116], M. Ballardini[117,9,118], D. Bertacca[98,64,59], A. Blanchard[17], L. Blot[119,4], H. Böhringer[35,120,121], S. Bruton[122], R. Cabanac[17], A. Calabro[47], B. Camacho Quevedo[62,24], G. Cañas-Herrera[90,123], A. Cappi[9,124], F. Caro[47], C. S. Carvalho[102], T. Castro[31,32,30,113], K. C. Chambers[125], S. Contarini[35], A. R. Cooray[126], G. Desprez[127], A. Díaz-Sánchez[128], J. J. Diaz[129], S. Di Domizio[36,37], H. Dole[58], S. Escoffier[76], M. Ezziati[43], A. G. Ferrari[48,23], P. G. Ferreira[130], I. Ferrero[19], A. Finoguenov[78], A. Fontana[47], F. Fornari[104], L. Gabarra[130], K. Ganga[86], J. García-Bellido[114], T. Gasparetto[31], V. Gautard[131], E. Gaztanaga[24,62,1], F. Giacomini[23], F. Gianotti[9], G. Gozaliasl[132,78], C. M. Gutierrez[133], A. Hall[3], H. Hildebrandt[7], J. Hjorth[91], A. Jimenez Muñoz[134], S. Joudaki[1], J. J. E. Kajava[135,136], V. Kansal[137,138], D. Karagiannis[139,140], C. C. Kirkpatrick[75], J. Le Graet[76], L. Legrand[141], J. Lesgourgues[2], T. I. Liaudat[142], S. J. Liu[60], A. Loureiro[143,144], G. Maggio[31], M. Magliocchetti[60], F. Mannucci[145], R. Maoli[146,47], J. Martín-Fleitas[147], C. J. A. P. Martins[148,41], L. Maurin[58], R. B. Metcalf[84,9], M. Miluzio[51,149], P. Monaco[150,31,32,30], A. Montoro[24,62], A. Mora[147], C. Moretti[33,113,31,30,32], G. Morgante[9], C. Murray[86], S. Nadathur[1], Nicholas A. Walton[110], L. Pagano[117,118], L. Patrizii[23], V. Popa[97], D. Potter[21], P. Reimberg[88], I. Risso[106], P.-F. Rocci[58], M. Sahlén[151], E. Sarpa[33,113,32], A. Schneider[21], M. Sereno[9,23], A. Silvestri[123], A. Spurio Mancini[152,67], J. Stadel[21], K. Tanidis[130], C. Tao[76], N. Tessore[74], G. Testera[37], R. Teyssier[153], S. Toft[154,155], S. Tosi[36,37], A. Troja[98,59], M. Tucci[57], J. Valiviita[78,79], D. Vergani[9], G. Verza[156,157], and P. Vielzeuf[76]

(Affiliations can be found after the references)

September 6, 2024

ABSTRACT

We study the constraint on $f(R)$ gravity that can be obtained by photometric primary probes of the *Euclid* mission. Our focus is the dependence of the constraint on the theoretical modelling of the nonlinear matter power spectrum. In the Hu–Sawicki $f(R)$ gravity model, we consider four different predictions for the ratio between the power spectrum in $f(R)$ and that in ΛCDM: a fitting formula, the halo model reaction approach, `ReACT` and two emulators based on dark matter only N-body simulations, `FORGE` and `e-Mantis`. These predictions are added to the `MontePython` implementation to predict the angular power spectra for weak lensing (WL), photometric galaxy clustering and their cross-correlation. By running Markov Chain Monte Carlo, we compare constraints on parameters and investigate the bias of the recovered $f(R)$ parameter if the data are created by a different model. For the pessimistic setting of WL, one dimensional bias for the $f(R)$ parameter, $\log_{10}|f_{R0}|$, is found to be 0.5σ when `FORGE` is used to create the synthetic data with $\log_{10}|f_{R0}| = -5.301$ and fitted by `e-Mantis`. The impact of baryonic physics on WL is studied by using a baryonification emulator `BCemu`. For the optimistic setting, the $f(R)$ parameter and two main baryon parameters are well constrained despite the degeneracies among these parameters. However, the difference in the nonlinear dark matter prediction can be compensated by the adjustment of baryon parameters, and the one-dimensional marginalised constraint on $\log_{10}|f_{R0}|$ is biased. This bias can be avoided in the pessimistic setting at the expense of weaker constraints. For the pessimistic setting, using the ΛCDM synthetic data for WL, we obtain the prior-independent upper limit of $\log_{10}|f_{R0}| < -5.6$. Finally, we implement a method to include theoretical errors to avoid the bias due to inaccuracies in the nonlinear matter power spectrum prediction.

Key words. Cosmology: theory; large-scale structure of Universe; cosmological parameters; dark energy. Gravitational lensing: weak

1. Introduction

In 1998, astronomers made a surprising discovery that the expansion of the Universe is accelerating, not slowing down (Riess et al. 1998; Perlmutter et al. 1999). This late-time acceleration of the Universe has become the most challenging problem in theoretical physics. The main aim of ongoing and future cosmological surveys is to address the key questions about the origin of the late-time acceleration. Is the acceleration driven by a cosmological constant or dark energy that evolves with the expansion of the Universe? Alternatively, there could be no dark energy if General Relativity (GR) itself is in error on cosmological scales. There has been significant progress in developing modified theories of gravity and these have been developed into tests of GR itself via cosmological observations (Koyama 2016; Ishak 2019; Ferreira 2019). Testing gravity is one of the main objectives of stage-IV dark energy surveys (Albrecht et al. 2006).

Of particular importance in these surveys is the *Euclid* mission (Euclid Collaboration: Mellier et al. 2024). The *Euclid* satellite undertakes a spectroscopic survey of galaxies and an imaging survey (targeting weak lensing which can also be used to reconstruct galaxy clustering using photometric redshifts). The combination of these two probes is essential for cosmological tests of gravity (Euclid Collaboration: Blanchard et al. 2020, EC20 hereafter).

Modified gravity models typically include an additional scalar degree of freedom that gives rise to a fifth-force. To satisfy the stringent constraints on deviations from GR in the solar system, many modified gravity models utilise screening mechanisms to hide modifications of gravity on small scales (Joyce et al. 2015; Brax et al. 2021). This is accomplished by nonlinearity in the equation that governs the dynamics of the scalar degree of freedom. This significantly complicates the nonlinear modelling of matter clustering in these models as the nonlinear equation for the scalar mode coupled to nonlinear density fields needs to be solved. A systematic comparison of N-body simulations in $f(R)$ gravity and normal-branch Dvali–Gabadadze–Porratti (nDGP) models was done in Winther et al. (2015). Since then, new simulations have been developed, including those using approximate methods to accelerate simulations (Valogiannis & Bean 2017; Winther et al. 2017). A fitting formula (Winther et al. 2019) and emulators for the nonlinear matter power spectrum have been developed (Arnold et al. 2019; Ramachandra et al. 2021; Sáez-Casares et al. 2023; Fiorini et al. 2023). At the same time, a semi-analytic method based on the halo model to predict the nonlinear matter power spectrum for general dark energy and modified gravity models has been developed (Cataneo et al. 2019; Bose et al. 2021; Bose et al. 2023; Carrilho et al. 2022). These nonlinear predictions were used to study their modifications of the Weak Lensing (WL) observables (Schneider et al. 2020; Harnois-Déraps et al. 2023; Spurio Mancini & Bose 2023; Carrion et al. 2024; Tsedrik et al. 2024), cross-correlation of galaxies and cosmic microwave background (Kou et al. 2023) for example in $f(R)$ gravity.

In Casas et al. (2023), Fisher Matrix forecasts were performed to predict *Euclid*'s ability to constrain $f(R)$ gravity models. In the Hu–Sawicki $f(R)$ model, it was shown that in the optimistic setting defined in EC20, and for a fiducial value of $|f_{R0}| = 5 \times 10^{-6}$, *Euclid* alone will be able to constrain the additional parameter $\log_{10}|f_{R0}|$ at the 3% level, using spectroscopic galaxy clustering alone; at the 1.4% level, using the combination of photometric probes on their own; and at the 1% level, using the combination of spectroscopic and photometric observations. The forecast for photometric probes used the fitting formula for the nonlinear matter power spectrum obtained in Winther et al. (2019). Further Fisher Matrix forecasts have been done for other modified gravity models with linearly scale-independent growth (Frusciante et al. 2023).

For real data analysis, it is imperative to check the effect of the accuracy of the theoretical modelling. This article is part of a series that collectively explores simulations and nonlinearities beyond the Λ Cold Dark Matter (ΛCDM) model:

1. Numerical methods and validation (Euclid Collaboration: Adamek et al. 2024, paper 1 hereafter)
2. Results from non-standard simulations (Euclid Collaboration: Racz et al. 2024, paper 2 hereafter)
3. Cosmological constraints on non-standard cosmologies from simulated *Euclid* probes (D'Amico et al. in prep.)
4. Constraints on $f(R)$ models from the photometric primary probes (this work)

In paper 1, the comparison of N-body simulations performed in Winther et al. (2015) in the Hu–Sawicki $f(R)$ and nDGP models was extended to add more simulations. The comparison was done for the matter power spectrum and the halo mass function. The measurements of these quantities were done using the dedicated pipeline developed by paper 2. In paper 2, additional simulations have been analysed in addition to those used in paper 1. In this paper, we utilise these developments and compare several predictions for the nonlinear matter power spectrum in $f(R)$ gravity. We will perform Markov Chain Monte Carlo (MCMC) simulations using synthetic data for *Euclid* photometric probes, and assess the impact of using different nonlinear models for the

⋆ e-mail: kazuya.koyama@port.ac.uk

matter power spectrum at the level of parameter constraints. In addition, we will add baryonic effects using a baryonification method, which was not included in the Fisher Matrix forecast. Although this paper focuses on $f(R)$ gravity, for which multiple public codes are available to predict the nonlinear matter power spectrum, the methodology and code developed in this paper are readily applicable to other modified gravity models. The validation of the nonlinear models for spectroscopic probes has been done in Bose et al. (2023) and D'Amico et al. (paper 3) will perform a similar analysis to this paper's for the spectroscopic probes.

This paper is organised as follows. In Sect. 2, we summarise theoretical predictions for *Euclid* observables based on EC20 and Casas et al. (2023). In Sect. 3, we introduce the Hu–Sawicki $f(R)$ gravity model and summarise the four nonlinear models for the matter power spectrum. We then discuss the implementation of these models in the `MontePython` code introduced in Casas et al. (2024). In Sect. 4, we compare predictions for the nonlinear matter power spectrum with N-body simulations and study their impact on the angular power spectra for *Euclid* photometric probes. Forecasts for errors from the combination of photometric probes considered in EC20 will be presented based on the synthetic data created by four different nonlinear models. We also compare the result with the Fisher Matrix forecast. We then study the bias in the recovered $f(R)$ gravity parameter when the synthetic data are created by a different nonlinear model. Section 5 is devoted to the study of baryonic effects using the `BCemu` baryonification method. We show how the bias is affected by baryonic effects and obtain the upper bound on $|f_{R0}|$ using the ΛCDM synthetic data. In Sect. 6, we implement theoretical errors to take into account the uncertainties of theoretical predictions. We conclude in Sect. 7.

2. Theoretical predictions for Euclid observables

In this section, we discuss how moving away from the standard GR assumption impacts the predictions for the angular power spectra $C(\ell)$ that will be compared with the photometric data survey. The observables that need to be computed and compared with the data are the angular power spectra for weak lensing (WL), photometric galaxy clustering ($\mathrm{GC_{ph}}$) and their cross-correlation ($\mathrm{XC_{ph}}$). In EC20, these were calculated using the Limber and flat-sky approximations in a flat ΛCDM Universe, as

$$C_{ij}^{XY}(\ell) = c \int_{z_{\rm min}}^{z_{\rm max}} \mathrm{d}z \, \frac{W_i^X(z)\, W_j^Y(z)}{H(z)\, r^2(z)} P_{\delta\delta}(k_\ell, z)\,, \tag{1}$$

where $W_i^X(z)$ is the window function in each tomographic redshift bin i with $X = \{\mathrm{L}, \mathrm{G}\}$ corresponding to WL and $\mathrm{GC_{ph}}$, respectively, $k_\ell = (\ell + 1/2)/r(z)$, $r(z)$ is the comoving distance to redshift z, $P_{\delta\delta}(k_\ell, z)$ is the nonlinear power spectrum of matter density fluctuations, δ, at wave number k_ℓ and redshift z, in the redshift range of the integral from $z_{\rm min} = 0.001$ to $z_{\rm max} = 4$ and $H(z)$ is the Hubble function.

However, when abandoning the assumption of GR, one has to account for changes in the evolution of both the homogeneous background and cosmological perturbations. For WL, it is the Weyl potential ϕ_W that determines the angular power spectrum. The power spectrum of the Weyl potential is related to $P_{\delta\delta}$ as (Casas et al. 2023)

$$P_{\phi_W}(k, z) = \left[-3\,\Omega_{\rm m} \left(\frac{H_0}{c}\right)^2 (1+z)\,\Sigma(k, z)\right]^2 P_{\delta\delta}(k, z)\,, \tag{2}$$

where $\Sigma(k, z)$ is a phenomenological parameterisation to account for deviations from the standard ΛCDM lensing prediction. $\Omega_{\rm m}$ is the density parameter of matter and H_0 is the Hubble parameter – both at the present time. Here we assumed a standard background evolution of the matter component, i.e. $\rho_{\rm m}(z) = \rho_{\rm m,0}(1+z)^3$.

We can, therefore, use the recipe of Eq. (1) with H, r and $P_{\delta\delta}$ provided by a Boltzmann solver, but with the new window functions (Spurio Mancini et al. 2019)

$$W_i^{\rm G}(k, z) = \frac{1}{c}\, b_i(k, z)\, n_i(z)\, H(z)\,, \tag{3}$$

$$W_i^{\rm L}(k, z) = \frac{3}{2}\,\Omega_{\rm m} \left(\frac{H_0}{c}\right)^2 (1+z)\, r(z)\, \Sigma(k, z) \times \int_z^{z_{\rm max}} \mathrm{d}z'\, n_i(z)\, \frac{r(z') - r(z)}{r(z')} + W_i^{\rm IA}(k, z)\,, \tag{4}$$

where $n_i(z)$ is the normalised galaxy number-density distribution in a tomographic redshift bin i such that $\int_{z_{\rm min}}^{z_{\rm max}} n_i(z) dz = 1$, and $W_i^{\rm IA}(k, z)$ encodes the contribution of intrinsic alignments (IA) to the WL power spectrum. We follow EC20 in assuming an effective scale-independent galaxy bias, constant within each redshift bin, and its values b_i are introduced as nuisance parameters in our analysis, with their fiducial values determined by $b_i = \sqrt{1+\bar{z}_i}$, where $\bar{z}_i$ is the mean redshift of each redshift bin.

The IA contribution is computed following the nonlinear alignment model with a redshift dependent amplitude (EC20), in which

$$W_i^{\rm IA}(k, z) = -\frac{\mathcal{A}_{\rm IA}\, C_{\rm IA}\, \Omega_{\rm m}\, \mathcal{F}_{\rm IA}(z)}{\delta(k, z) \,/\, \delta(k, 0)}\, n_i(z)\, \frac{H(z)}{c}\,, \tag{5}$$

where

$$\mathcal{F}_{\rm IA}(z) = (1+z)^{\eta_{\rm IA}}. \tag{6}$$

The parameters $\mathcal{A}_{\rm IA}$ and $\eta_{\rm IA}$ are the nuisance parameters of the model, and $C_{\rm IA}$ is a constant accounting for dimensional units. The galaxy distribution is binned into 10 equipopulated redshift bins with an overall distribution following

$$n(z) \propto \left(\frac{z}{z_0}\right)^2 \exp\left[-\left(\frac{z}{z_0}\right)^{3/2}\right], \tag{7}$$

with $z_0 = 0.9/\sqrt{2}$ and the normalisation set by the requirement that the surface density of galaxies is $\bar{n}_{\rm g} = 30\,\mathrm{arcmin}^{-2}$ (EC20).

Changes in the theory of gravity impact the IA contribution, introducing a scale dependence through the modified perturbations' growth. This is explicitly taken into account in Eq. (5) through the matter perturbation $\delta(k, z)$, while the modifications on the clustering of matter in the $\mathrm{GC_{ph}}$ case are accounted for in the new $P_{\delta\delta}(k_\ell, z)$.

Finally, we present the likelihood that we will use. We follow the approach presented in Casas et al. (2024). We first construct an $(N^{\rm G} + N^{\rm L}) \times (N^{\rm G} + N^{\rm L})$ angular power spectrum matrix for each multipole, where the different N correspond to the number of redshift bins for each probe (WL and $\mathrm{GC_{ph}}$):

$$\mathsf{C}_\ell = \begin{bmatrix} C_{ij}^{\rm LL}(\ell) & C_{ij}^{\rm GL}(\ell) \\ C_{ij}^{\rm LG}(\ell) & C_{ij}^{\rm GG}(\ell) \end{bmatrix}, \tag{8}$$

where lower-case Latin indexes $i, \ldots$ run over all tomographic bins. Similarly, the noise contributions can also be condensed into one noise matrix N:

$$\mathsf{N}_\ell = \begin{bmatrix} N_{ij}^{\rm LL}(\ell) & N_{ij}^{\rm GL}(\ell) \\ N_{ij}^{\rm LG}(\ell) & N_{ij}^{\rm GG}(\ell) \end{bmatrix}, \tag{9}$$

where the noise terms $N_{ij}^{AB}(\ell)$ are given by

$$N_{ij}^{\rm LL}(\ell) = \frac{\delta_{ij}^{\rm K}}{\bar{n}_i}\sigma_\epsilon^2\,, \tag{10}$$

$$N_{ij}^{\rm GG}(\ell) = \frac{\delta_{ij}^{\rm K}}{\bar{n}_i}\,, \tag{11}$$

$$N_{ij}^{\rm GL}(\ell) = 0\,, \tag{12}$$

where $\sigma_\epsilon^2 = 0.3^2$ is the variance of observed ellipticities. We can further define $\hat{\mathsf{C}}_\ell := \mathsf{C}_\ell + \mathsf{N}_\ell$. This is the covariance of the spherical multipole moments a_{lm}.

In the Gaussian approximation, it can be shown that the description in EC20, using the covariance of the angular power spectra, is equivalent to the description found in Casas et al. (2024), using the covariance of the $a_{\ell m}$. The likelihood then can be expressed in terms of the observed covariance

$$\hat{\mathsf{C}}_{ij}^{\rm obs}(\ell) = \frac{1}{2\,\ell+1}\sum_{m=-\ell}^{\ell}(a_{\ell m})_i\,(a_{\ell m})_j^*\,, \tag{13}$$

and the theoretically predicted one $\hat{\mathsf{C}}_{ij}^{\rm th}(\ell)$ as

$$\chi^2 = f^{\rm sky}\sum_{\ell=\ell_{\rm min}}^{\ell_{\rm max}}(2\ell+1)\left[\frac{d^{\rm mix}}{d^{\rm th}} + \ln\left(\frac{d^{\rm th}}{d^{\rm obs}}\right) - N\right]\,, \tag{14}$$

where the determinants d are defined as

$$d^{\rm th}(\ell) = \det\left[\hat{\mathsf{C}}_{ij}^{\rm th}(\ell)\right], \tag{15}$$

$$d^{\rm obs}(\ell) = \det\left[\hat{\mathsf{C}}_{ij}^{\rm obs}(\ell)\right], \tag{16}$$

$$d^{\rm mix}(\ell) = \sum_k^N \det\left[\begin{cases}\hat{\mathsf{C}}_{ij}^{\rm obs}(\ell) & \text{for } k=j\\ \hat{\mathsf{C}}_{ij}^{\rm th}(\ell) & \text{for } k\neq j\end{cases}\right]. \tag{17}$$

Here N is the number of bins and thus either $(N^{\rm G}+N^{\rm L})$ for multipoles where we include the cross correlation, or the respective N for multipoles where we treat the two probes separately. The additional factor $f^{\rm sky}$ comes from an approximation to account for having less available independent ℓ modes due to partial sky coverage. In this paper, we set the observed covariance $\hat{\mathsf{C}}^{\rm obs}$ to the theoretically predicted one computed at the fiducial cosmology. Following EC20 and Casas et al. (2024), we do not include super-sample covariance (SSC) in this work. The SSC impact was shown to be non-negligible and will need to be included in future analyses as shown in Euclid Collaboration: Sciotti et al. (2023).

We consider two different scenarios: an optimistic and a pessimistic case. In the optimistic case, we consider $\ell_{\rm max} = 5000$ for WL, and $\ell_{\rm max} = 3000$ for $\rm GC_{ph}$ and $\rm XC_{ph}$. Instead, in the pessimistic scenario, we consider $\ell_{\rm max} = 1500$ for WL, and $\ell_{\rm max} = 750$ for $\rm GC_{ph}$ and $\rm XC_{ph}$.

In the smallest photometric redshift bin, the galaxy number density distribution $n(z)$ peaks at around $z = 0.25$. Under the Limber approximation for our fiducial cosmology, the corresponding maximum values of k evaluated in the power spectrum corresponding to the pessimistic and optimistic scenario for $\rm GC_{ph}$ are $k_{\rm max} = [0.7, 2.9]\,h\,{\rm Mpc}^{-1}$, respectively, while for the WL, maximum wavenumbers probed are $k_{\rm max} = [1.4, 4.8]\,h\,{\rm Mpc}^{-1}$ at the peak redshift $z = 0.25$ (Casas et al. 2023). Here h denotes the dimensionless Hubble parameter $h := H_0/(100{\rm km\,s^{-1}\,Mpc^{-1}})$. For smaller values of z, the values of k at a given ℓ increase, but the window functions in Eqs. (3) and (4) suppress the power spectrum and we set it to zero after a fixed $k_{\rm max} = 30\,h\,{\rm Mpc}^{-1}$. We list the specific choices of scales and settings used for each observable in Table 1. Although the currently foreseen specification of the Euclid Wide Survey differs from that assumed in EC20, for example in terms of the survey area, we use their results to allow for comparison with earlier forecasts.

Table 1. *Euclid* survey specifications for WL, $\rm GC_{ph}$ and $\rm GC_{sp}$ used by EC20 and Casas et al. (2024).

Survey sky coverage	$f^{\rm sky}$	36.36%
WL		
Number of photo-z bins	$N^{\rm L}$	10
Galaxy number density	$\bar{n}_{\rm gal}$	$30\,{\rm arcmin}^{-2}$
Intrinsic ellipticity σ	σ_ϵ	0.30
Minimum multipole	$\ell_{\rm min}$	10
Maximum multipole	$\ell_{\rm max}$	
– Pessimistic		1500
– Optimistic		5000
$\rm GC_{ph}$		
Number of photo-z bins	$N^{\rm G}$	10
Minimum multipole	$\ell_{\rm min}$	10
Maximum multipole	$\ell_{\rm max}$	
– Pessimistic		750
– Optimistic		3000

3. `MontePython` implementation

In this section, we describe the implementation of the likelihood discussed in Sect. 2 in the `MontePython` code developed in Casas et al. (2024), in the Hu–Sawicki $f(R)$ gravity model.

3.1. Hu–Sawicki $f(R)$ gravity

Modified gravity $f(R)$ models (Buchdahl 1970) are models constructed by promoting the Ricci scalar R in the Einstein–Hilbert action to a generic function of $R \to R + f(R)$, i.e.

$$S = \frac{c^4}{16\pi G}\int {\rm d}^4x\,\sqrt{-g}\,[R + f(R)] + S_{\rm m}[g_{\mu\nu}, \Psi_{\rm m}]\,, \tag{18}$$

where $g_{\mu\nu}$ is the metric tensor, g its determinant and $S_{\rm m}$ represents the matter sector with its matter fields $\Psi_{\rm m}$. For further discussions we refer to Sotiriou & Faraoni (2010), Clifton et al. (2012) and Koyama (2016).

How this modification changes gravity is more easily seen by formulating the theory in the so-called Einstein frame by performing a conformal transformation $g_{\mu\nu} = \tilde{g}_{\mu\nu}A^2(\phi)$ where $A(\phi) = \sqrt{1+f_R} = e^{\phi/\sqrt{6}}$ to obtain

$$S = \frac{c^4}{16\pi G}\int {\rm d}^4x\,\sqrt{-\tilde{g}}\left[\tilde{R} + \frac{1}{2}(\partial\phi)^2 - V(\phi)\right] + S_{\rm m}[A^2(\phi)\,\tilde{g}_{\mu\nu}, \Psi_{\rm m}]\,, \tag{19}$$

with the potential $V = [f_R R - f(R)]/2(1+f_R)^2$ and $f_R := {\rm d}f(R)/{\rm d}R$. This demonstrates that the theory reduces to standard GR together with an extra scalar field that is coupled to the matter sector giving rise to a fifth-force. This fifth-force has a finite range $\lambda_{\rm c}$ and within this range, it mediates a force that has a strength that, in the linear regime, corresponds to 1/3 of the usual gravitational force:

$$\mathbf{F}_{\rm fifth} = \frac{1}{3}\frac{GMm}{r^2}\,(1 + r/\lambda_{\rm c})\,{\rm e}^{-r/\lambda_{\rm c}}\,. \tag{20}$$

This effectively changes $G \to \frac{4}{3}G$ on small scales ($r \ll \lambda_c$) while keeping the usual G on large scales. Such a large modification would be ruled out by observations if it was not for the fact that the theory possesses a screening mechanism (Khoury & Weltman 2004; Brax et al. 2008) which hides the modifications in high-density regions. This screening mechanism effectively suppresses the fifth-force by a factor $\propto f_R/\Phi_N$ where Φ_N is the Newtonian gravitational potential.

Not all $f(R)$ models one can write down possess such a screening mechanism as this places some restrictions on their functional form (see, e.g., Sotiriou & Faraoni 2010). One concrete example of a model that has all the right ingredients is the model proposed by Hu & Sawicki (2007) which in the large-curvature limit is defined by

$$f(R) = -6\Omega_{\rm DE}\frac{H_0^2}{c^2} + |f_{R0}|\frac{\bar{R}_0^2}{R}\,, \tag{21}$$

where

$$\bar{R}_0 = 3\Omega_{\rm m}\frac{H_0^2}{c^2}\left(1 + 4\frac{\Omega_{\rm DE}}{\Omega_{\rm m}}\right), \tag{22}$$

is the Ricci scalar in the cosmological background and $\Omega_{\rm DE}$ is the density parameter of dark energy at the present time. The first term in Eq. (21) corresponds to a cosmological constant and the only free parameter is $|f_{R0}|$. In the limit $|f_{R0}| \to 0$, we recover GR and the ΛCDM model. This parameter controls the range of the fifth-force and, in the cosmological background, we have $\lambda_c \simeq 32\sqrt{|f_{R0}|/10^{-4}}$ Mpc at the present time. Solar System constraints require $|f_{R0}| \lesssim 10^{-6}$, cosmological constraints currently lie around 10^{-6}–10^{-4} depending on the probe in question (see, e.g., Fig. 28 in Koyama 2016, for a summary) while astrophysical constraints at the galaxy scale can be as tight as $\lesssim 10^{-8}$ (Desmond & Ferreira 2020), but see Burrage et al. (2024) for a recent note of caution on such galactic-scale constraints.

The energy density of the scalar field contributes in general to the expansion of the Universe, however for viable models, like the one we consider here, this contribution is tiny (of the order $|f_{R0}|\,\Omega_{\rm DE,0}$) apart from the constant part of the potential which is indistinguishable from a cosmological constant. The background evolution of such models is therefore very close to ΛCDM. Since $f(R)$ models have a conformal coupling, light deflection is weakly affected as follows

$$\Sigma(z) = \frac{1}{1 + f_R(z)}\,. \tag{23}$$

Since the maximum value of $|f_R(z)|$ is given by $|f_{R0}|$, for the values of $|f_{R0}|$ we consider in this paper, we can ignore this effect. Thus gravitational lensing is also not modified in the sense that the lensing potential is sourced by matter in the same way as in GR (though the underlying density field will of course be different). The main cosmological signatures of the model therefore come from having a fifth-force, acting only on small scales $r \lesssim \lambda_c$, in the process of structure formation. The main effect of the screening mechanism is that the prediction for the amount of clustering will generally be much smaller than what naive linear perturbation theory predicts.

3.2. Nonlinear modelling

We implement $\Xi(k,z)$ defined as

$$\Xi(k,z) := \frac{P_{f(R)}(k,z)}{P_{\Lambda\rm CDM}(k,z)}\,, \tag{24}$$

to obtain the nonlinear $f(R)$ matter power spectrum. For the ΛCDM power spectrum $P_{\Lambda\rm CDM}(k,z)$, we use the `halofit` 'Takahashi' prescription (Takahashi et al. 2012) following Casas et al. (2023). It includes the minimum mass for massive neutrinos in $P_{\Lambda\rm CDM}(k,z)$, but ignores the effect of massive neutrinos on $\Xi(k,z)$. This approximation was shown to be well justified for the minimum mass of neutrinos in Winther et al. (2019) using data from Baldi et al. (2014). We describe below four models for $\Xi(k,z)$ used in this paper. We note that we can use any ΛCDM power spectrum prediction in our approach such as `EuclidEmulator2` (Euclid Collaboration: Knabenhans et al. 2021) and `Bacco` (Angulo et al. 2021). The exercise we perform in this paper is a comparison of the different prescriptions for $\Xi(k,z)$. Given this, the ΛCDM nonlinear spectrum prescription does not matter and it is common to all nonlinear models.

3.2.1. Fitting formula

A fitting formula for $\Xi(k,z)$ was developed in Winther et al. (2019) and describes the enhancement in the power spectrum compared to a ΛCDM nonlinear power spectrum as a function of the parameter $|f_{R0}|$. This fitting function has been calibrated using the `DUSTGRAIN` (Giocoli et al. 2018) N-body simulations run by `MG-Gadget` (Puchwein et al. 2013) and the `ELEPHANT` N-body simulations (Cautun et al. 2018) run by `ECOSMOG` (Li et al. 2012; Bose et al. 2017). The main approximation is that the cosmological parameter dependence of $\Xi(k,z)$ is ignored. In Winther et al. (2019), this assumption was checked using simulations with different σ_8, $\Omega_{\rm m}$, as well as the mass of massive neutrinos. Winther et al. (2019) also corrected the fitting formula to account for additional dependence on these parameters. In this paper, we do not include these corrections as the previous forecast paper (Casas et al. 2023) used the fitting formula without these corrections.

The fitting function has in total 54 parameters for the full scale, redshift and $|f_{R0}|$ dependence. This fitting formula is not defined outside the range $10^{-7} < |f_{R0}| < 10^{-4}$. The code is publicly available ( ).

3.2.2. Halo model reaction

We further consider the halo model reaction approach of Cataneo et al. (2019) which combines the halo model and perturbation theory frameworks to model corrections coming from non-standard physics. The nonlinear power spectrum is given by

$$P_{\rm NL}(k,z) = \mathcal{R}(k,z)\,P_{\rm NL}^{\rm pseudo}(k,z)\,, \tag{25}$$

where $P_{\rm NL}^{\rm pseudo}(k,z)$ is called the pseudo-power spectrum and is defined as a nonlinear ΛCDM spectrum with initial conditions tuned such that the linear clustering at the target redshift z matches the beyond-ΛCDM case. This choice ensures the halo mass functions in the beyond-ΛCDM and pseudo-universes are similar, giving a smoother transition of the power spectrum over inter- and intra-halo scales. We model the pseudo-cosmology nonlinear power spectrum using `HMCode2020` (Mead et al. 2015, 2016, 2021) by supplying the code with the linear $f(R)$ power spectrum.

The halo model reaction, $\mathcal{R}(k,z)$, models all the corrections to the pseudo spectrum coming from nonlinear beyond-ΛCDM physics. We refer the reader to Cataneo et al. (2019), Bose et al. (2021) and Frusciante et al. (2023) for the exact expressions for this term. The halo model reaction can be computed efficiently

using the publicly available ReACT (Bose et al. 2020, 2021; Bose et al. 2023,) code.

Despite ReACT being highly efficient, having been used in previous real cosmic shear analyses (Tröster et al. 2021), it is still too computationally expensive for the number of tests we wish to perform. To accelerate our inference pipeline, we create a neural network emulator using the Cosmopower package (Spurio Mancini et al. 2022,) for the halo model reaction-based boost

$$\Xi(k,z) = \frac{\mathcal{R}(k,z)\, P^{\rm pseudo}_{\rm HMCode2020}(k,z)}{P_{\Lambda{\rm CDM}}(k,z)}\,, \tag{26}$$

where $P_{\Lambda{\rm CDM}}(k,z)$ and $P^{\rm pseudo}_{\rm HMCode2020}(k,z)$ are calculated using HMCode2020 (Mead et al. 2021,) while $\mathcal{R}(k,z)$ is calculated using ReACT. We chose the HMCode2020 to model the pseudo-power spectrum as it has been shown to have improved accuracy and does not show suppression of power with respect to ΛCDM ($\Xi < 1$) at high redshifts, which is not expected.

To train the emulator, we follow the procedure of Spurio Mancini & Bose (2023) but widen the parameter priors. We produce $\sim 10^5$ boost predictions in the range $k \in [0.01, 3]\,h\,{\rm Mpc}^{-1}$ and $z \in [0, 2.5]$. We take cosmologies from the Latin hypercube given by the ranges in Table 3 and Table 4, with the massive neutrino density parameter today's range being $\Omega_\nu \in [0.0, 0.00317]$. Emulation of the boost speeds up the computation by 4 orders of magnitude and we find similar accuracy of our emulator as found in Spurio Mancini & Bose (2023). The emulator is publicly available (). Finally, we note another small difference between our emulator and that of Spurio Mancini & Bose (2023). In this work, we assume $P_{\Lambda{\rm CDM}}$ in Eq. (26) with $\Omega_{\rm m}$ equal to the total of the true cosmology, whereas in Spurio Mancini & Bose (2023) they assume $\Omega_{\rm m} = \Omega_{\rm b} + \Omega_{\rm c}$, $\Omega_{\rm b}$ and $\Omega_{\rm c}$ being the baryon and cold dark matter density parameters today respectively. The emulators give the same output for $\Omega_\nu = 0$, which is what we assume in this work for Ξ.

3.2.3. FORGE

The FORGE emulator was introduced in Arnold et al. (2022). It is based on simulations for 50 combinations of $|f_{R0}|$, $\Omega_{\rm m}$, $\sigma_8^{\Lambda{\rm CDM}}$ and h with all other parameters fixed. We note that $\sigma_8^{\Lambda{\rm CDM}}$ is the σ_8 we would obtain in a ΛCDM model with the same cosmological parameters and initial amplitude $A_{\rm s}$ and not σ_8 in an $f(R)$ gravity model. The emulator accuracy is better than 2.5% around the centre of the explored parameter space, up to scales of $k = 10\,h\,{\rm Mpc}^{-1}$. $f(R)$ simulations are performed by a modified version of the Arepo code, MG-AREPO (Springel 2010; Arnold et al. 2019; Weinberger et al. 2020) that solves the nonlinear scalar field equation using a relaxation method. See paper 1 for more details.

The emulation was made for the ratio between the power spectrum in $f(R)$ and the halofit prediction in ΛCDM. We note that this is different from $\Xi(k,z)$ as the power spectrum in ΛCDM models in these simulations can have slight deviations from the halofit prediction. To account for this effect, the authors provided the ratio of the power spectrum in a reference ΛCDM model to the halofit prediction. This ΛCDM simulation uses $\Omega_{\rm m} = 0.31315$, $h = 0.6737$, $\sigma_8^{\Lambda{\rm CDM}} = 0.82172$. This can be used to obtain $\Xi(k,z)$. However, this correction is provided only in this ΛCDM model. Thus, the assumption here is that this correction is independent of cosmological parameters. The latest version of the FORGE simulations has ΛCDM counterparts to the $f(R)$ simulations using the same seed. These simulations were analysed in paper 2. Hence, it is in principle possible to emulate directly $\Xi(k,z)$ from these simulations. However, as this is not publicly available, we opted for using the original FORGE emulator () for this paper. We will use one of these simulations for validation in Sect. 4.

3.2.4. e-Mantis

We also consider the predictions given by the e-Mantis emulator presented in Sáez-Casares et al. (2023), which can predict the ratio $\Xi(k,z)$ between the nonlinear matter power spectrum in $f(R)$ gravity and in ΛCDM. The predictions are calibrated from a large suite of N-body simulations run with the code ECOSMOG (Li et al. 2012; Bose et al. 2017), a modified gravity version of the Adaptive Mesh Refinement (AMR) N-body code RAMSES (Teyssier 2002; Guillet & Teyssier 2011). The e-Mantis simulation suite covers the $\left\{|f_{R_0}|, \Omega_{\rm m}, \sigma_8^{\Lambda{\rm CDM}}\right\}$ parameter space with 110 cosmological models sampled from a Latin hypercube (McKay et al. 1979). The remaining ΛCDM parameters, h, $n_{\rm s}$ and $\Omega_{\rm b}$ remain fixed, which means that their impact on $\Xi(k,z)$ is ignored. In Sáez-Casares et al. (2023), it was shown that the error made by this approximation is at the sub-percent level. The quantity $\Xi(k,z)$ is measured from pairs of $f(R)$ and ΛCDM simulations, both run with the same initial conditions, which leads to a large cancellation of cosmic variance and numerical resolution errors. The emulation is done through a Gaussian Process regression (see e.g. Rasmussen & Williams 2005). The emulator can give predictions for redshifts $z \in [0, 2]$, wavenumbers $k \in [0.03, 10]\,h\,{\rm Mpc}^{-1}$ and cosmological parameters in the following range: $\left|f_{R_0}\right| \in \left[10^{-7}, 10^{-4}\right]$, $\Omega_{\rm m} \in [0.2365, 0.3941]$ and $\sigma_8^{\Lambda{\rm CDM}} \in [0.6083, 1.0140]$. The estimated accuracy of e-Mantis, including emulation errors and systematic errors in the training data, is better than 3% for scales $k \lesssim 7\,h\,{\rm Mpc}^{-1}$ and across the whole parameter space although in most cases, the accuracy is of order 1%. The code is publicly available ().

3.3. MontePython implementation

Our implementation of the likelihood is based on the MontePython implementation described in Casas et al. (2024). For a more detailed explanation of the numerical implementation, we refer to that work. For our purposes, we have modified this code to include our prediction for $\Xi(k,z)$ to obtain the nonlinear power spectrum in the $f(R)$ gravity model. As the different implementations have different ranges in cosmological parameters, wavenumbers, and redshifts, we have chosen an extrapolation scheme to unify the ranges. We note that the exact implementation does not affect the final results strongly. We checked that the contributions from high redshifts and high wavenumbers which used extrapolations were sub-dominant.

For the modified gravity parameter $|f_{R0}|$, we use flat priors in terms of $\log_{10}|f_{R0}|$ to stay within the validity range of each model. For the largest scales beyond the range of the emulators, we set $\Xi = 1$. This is because the effect of the fifth-force has to vanish on these scales. For small scales, we do a power-law extrapolation. To obtain the spectral index for the extrapolation, we proceed as follows: we calculate Ξ on a grid close to the edges of the emulators; we then fit, for fixed z, a linear function in the log-log space onto this grid. This is done to average out any numerical noise at the edge. We do the same for high redshifts

Table 2. A summary of four nonlinear models that we use in this paper. For the details of N-body simulations codes and simulation suites, see paper 1 and paper 2. The prediction for Ξ remains the same if we vary cosmological parameters which are not included in the prediction.

method	simulations (code)	parameters
`fitting`		
fitting formula	`ELEPHANT` (`ECOSMOG`) `DUSTGRAIN` (`MG-Gadget`)	$\lvert f_{R_0}\rvert$
`ReACT`		
halo model (emulator)	N.A.	$\lvert f_{R_0}\rvert, \Omega_{\rm m}, \Omega_{\rm b}, h, n_s, A_{\rm s}$
`FORGE`		
emulator	`FORGE` (`MG-AREPO`)	$\lvert f_{R_0}\rvert, \Omega_{\rm m}, h, \sigma_8^{\Lambda{\rm CDM}}$
`e-Mantis`		
emulator	(`ECOSMOG`)	$\lvert f_{R_0}\rvert, \Omega_{\rm m}, \sigma_8^{\Lambda{\rm CDM}}$

Table 3. Ranges of wavenumbers, redshifts and $\log_{10}|f_{R0}|$ in different emulators used in this work.

k range $[h\,{\rm Mpc}^{-1}]$	z range	$\log_{10}\lvert f_{R0}\rvert$ range
`fitting`		
$[\ \ 10^{-2}, 10]$	≤ 2	$[\ -7\ \ , -4\ \ ]$
`ReACT`		
$[\ \ 10^{-2},\ \ 3]$	≤ 2.5	$[-10\ \ , -4\ \ ]$
`FORGE`		
$[\ \ 10^{-3}, 10]$	≤ 2	$[\ -6.4\ , -4.5]$
`e-Mantis`		
$[3\times 10^{-2}, 10]$	≤ 2	$[\ -7\ \ , -4\ \ ]$

by fitting the power law for fixed wavenumbers. For regions of both high k and z, we do a constant extrapolation of the spectral index. To keep the extrapolated function $\Xi(k,z)$ from going to non-physical values, we set a hard lower bound of $\Xi = 1$ and an upper bound of $\Xi = 2$. The ranges of the different emulators are found in Table 3. We do a constant extrapolation in the ΛCDM parameters. This can be done because during a typical MCMC the majority of the suggested points are within the ranges of our emulators. The validity range of ΛCDM cosmological parameters is summarised in Table 4.

We also add the effect of baryonic feedback in the form of an additional correction $\Xi^{\rm BFM}$ to our power spectrum prediction. The physics of the baryonic feedback effects is discussed in Sect. 5. We refer the reader to Schneider et al. (2020) and Mead et al. (2021) for further details. We obtain the correction from `BCemu` by Giri & Schneider (2021). This emulator is only trained for redshifts below $z = 2$ and wavenumbers $k < 12.5\,h\,{\rm Mpc}^{-1}$. In this case, we use constant extrapolation for both k and z. We checked that this had little effect on our results. This is because the *Euclid* main probes are mostly sensitive to redshifts around $z \sim 1$. For these redshifts, the extrapolation only affects very high multipoles $\ell \gtrsim 4500$. Thus, the extrapolation has negligible effect on the angular power spectrum. We estimate the overall effect this has on our result to be at most at the percent level. In the absence of modified gravity, to obtain the nonlinear baryonic feedback power spectrum, we multiply the ΛCDM nonlinear power spectrum by $\Xi^{\rm BFM}$. When adding the effect of modified gravity, we combine both boosts as

$$P^{\rm BFM}_{f(R)}(k,z) = \Xi^{\rm BFM}(k,z)\,\Xi(k,z)\,P_{\Lambda{\rm CDM}}(k,z)\,.$$

We can make this approximation if both effects are independent. This was shown to be the case for small deviations from ΛCDM in $f(R)$ gravity models considered in this paper using hydrodynamical simulations by Arnold et al. (2019) and Arnold & Li (2019).

Our final addition to the code is the inclusion of theoretical errors. For this, we have adjusted the prescription of Audren et al. (2013). The theory and numerical implementation is further discussed in Sect. 6.

4. Comparison of different nonlinear models

In this section, we compare four different predictions for the nonlinear dark matter power spectrum in terms of $\Xi(k,z)$ defined in Eq. (24) introduced in Sect. 3. We start from a comparison of the matter power spectrum with N-body simulations. We then compare the angular power spectra in the *Euclid* reference cosmology. We perform MCMC simulations to compare errors and investigate bias due to the difference in the prediction of the nonlinear matter power spectrum using the settings defined in EC20.

4.1. Comparison of predictions

4.1.1. Comparison with N-body simulations

Figure 1 shows a comparison of the ratio of the power spectra between $f(R)$ gravity and ΛCDM measured from N-body simulations with the theoretical predictions for Ξ. These simulations use the same initial conditions and the ratio removes the cosmic variance and the effect of mass resolution. The left-hand side plot shows a comparison using the measurements from paper 1. This is based on the comparison project in Winther et al. (2015). These simulations were run in a ΛCDM cosmology with $\Omega_{\rm m} = 0.269$, $h = 0.704$, $n_{\rm s} = 0.966$ and $\sigma_8 = 0.801$. The simulations have $N_{\rm p} = 512^3$ particles of mass $M_{\rm p} \simeq 8.756\times 10^9\,h^{-1}\,M_\odot$ in a box of size $B = 250\,h^{-1}\,{\rm Mpc}$ and start at redshift $z = 49$. We picked a model called F5 with $|f_{R0}| = 10^{-5}$ and showed the result at $z = 0.667$ that is presented in paper 1. In the comparison, we included the measurements from `MG-AREPO` and `ECOSMOG`, as `FORGE` is based on `MG-AREPO`, while `e-Mantis` is based on `ECOSMOG`. The prediction of `e-Mantis` agrees with `ECOSMOG` very well. On the other hand, the `FORGE` prediction deviates from `MG-AREPO` as well as `ECOSMOG`. We note that the `FORGE` prediction is corrected using the ΛCDM simulation (Node 0) in the `FORGE` simulation suite with $\Omega_{\rm m} = 0.31315$, $h = 0.6737$, $\sigma_8^{\Lambda{\rm CDM}} = 0.82172$ to obtain Ξ. We find that if we use $\Omega_{\rm m} = 0.31315$ and $h = 0.6737$ in the `FORGE` prediction, the agreement with `MG-AREPO` is much better.

To further investigate this issue with `FORGE`, we use the measurement of the power spectrum in one of the nodes in the `FORGE` simulation suite (Node 13) run by `MG-AREPO` that is closest to the *Euclid* reference cosmology that we will use in this paper with non-zero $|f_{R0}|$. This simulation has the following cosmological

Table 4. Ranges for cosmological parameters implemented in the `MontePython` code.

Ω_m	Ω_b	$100h$	n_s	$\ln(10^{10}A_s)$	$\sigma_8^{\Lambda CDM}$
`ReACT`					
[0.24, 0.35]	[0.04, 0.06]	[63,75]	[0.9, 1.01]	[2.83, 3.22]	N.A.
`FORGE`					
[0.18, 0.55]	N.A.	[60,80]	N.A.	N.A.	[0.6, 1.0]
`e-Mantis`					
[0.24, 0.39]	N.A.	N.A.	N.A.	N.A.	[0.6, 1.0]

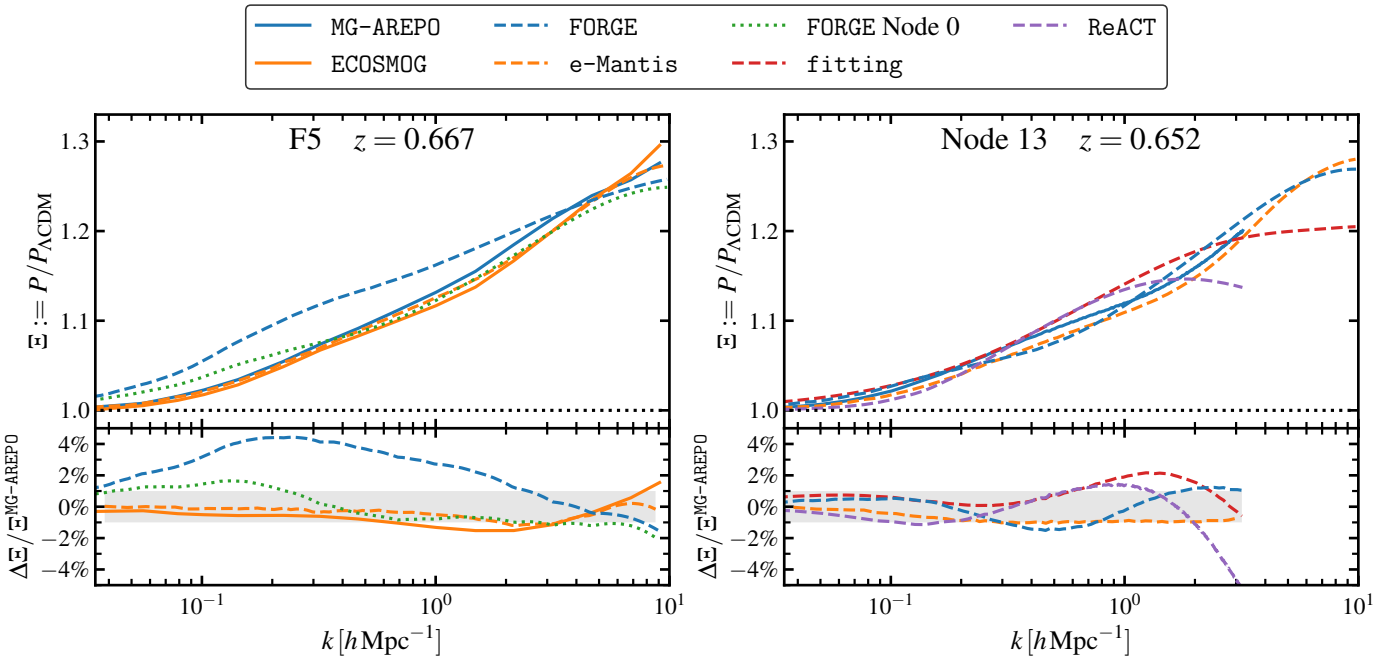

Fig. 1. *Left panel*: Comparisons between N-body simulations, `e-Mantis` and `FORGE` for $|f_{R0}| = 10^{-5}$. N-body data from simulations run by `MG-AREPO` and `ECOSMOG` codes are taken from the latest comparison project presented in paper 1. We also include the prediction of `FORGE` where the cosmological parameters in the fiducial ΛCDM simulation, $\Omega_m = 0.31315$ and $h = 0.6737$, are used to make the prediction (`FORGE` Node 0). *Right panel*: Comparison of four prescriptions with an N-body simulation run by `MG-AREPO`. This simulation is one of the simulations used for training to construct the `FORGE` emulator (Node 13). The measurement of the power spectrum was done in paper 2.

parameters: $\Omega_m = 0.34671$, $h = 0.70056$ and $\sigma_8 = 0.78191$ and we show a comparison at $z = 0.652$. Both ΛCDM and $f(R)$ simulations with $|f_{R0}| = 10^{-4.90056}$ are available so that we can measure Ξ directly. We note that the pipeline developed in paper 2 measures the power spectrum only up $k = 3\,h\,\mathrm{Mpc}^{-1}$. In this case, both `FORGE` and `e-Mantis` agree with `MG-AREPO` within 1% up to $k = 3\,h\,\mathrm{Mpc}^{-1}$. The fitting formula and `ReACT` agree with `MG-AREPO` within 1% up to $k = 1\,h\,\mathrm{Mpc}^{-1}$. Given this result, the large discrepancy between `FORGE` and `MG-AREPO` is likely to be caused by emulation errors as well as calibrations using the `halofit` ΛCDM power spectra predictions.

4.1.2. Comparison in the *Euclid* reference cosmology

In this paper, we consider the model called HS6 in Casas et al. (2023), which has the following parameters

$$\begin{aligned} \boldsymbol{\Theta} &= \{\Omega_m, \Omega_b, h, n_s, \sigma_8, \log_{10}|f_{R0}|\}\,, \\ \mathrm{HS6}: \boldsymbol{\Theta}_{\rm fid} &= \{0.32, 0.05, 0.67, 0.96, 0.853, -5.301\}\,. \end{aligned} \tag{27}$$

The cosmological parameters are the same ones adopted in EC20. Casas et al. (2023) showed that this value of $|f_{R0}| = 5 \times 10^{-6}$ can be well constrained by the *Euclid* photometric probes. Also the range of $|f_{R0}|$ covered by the four models is wider than the errors predicted in Casas et al. (2023). Our fiducial cosmology includes massive neutrinos with a total mass of $\sum m_\nu = 0.06\,\mathrm{eV}$, but we keep $\sum m_\nu$ fixed in the following analysis.

Figure 2 shows a comparison of the power spectrum and the angular power spectrum for WL, $\mathrm{GC_{ph}}$ and their cross-correlation $\mathrm{XC_{ph}}$. In these plots, we show the ratio to the ΛCDM prediction and error bars from the diagonal part of the covariance matrix. For the power spectrum comparison at $z = 0.5$, we see that `e-Mantis` and `fitting` agree best. This is not surprising as these two predictions are based on N-body simulations ran by the same code `ECOSMOG`. On the other hand, `FORGE` overestimates Ξ at $k = 0.1h\,\mathrm{Mpc}^{-1}$. This is similar to the deviation we find in the comparison with the N-body simulation from paper 1 although the deviation is smaller at the 2% level. This is likely due to the fact that Ω_m in HS6 is closer to Ω_m in the fiducial ΛCDM `FORGE` simulation ($\Omega_m = 0.31315$). On the other hand, `ReACT` underestimates Ξ at $k < 0.1h\,\mathrm{Mpc}^{-1}$ and even predicts $\Xi < 1$. As we discussed in the previous section, we enforce $\Xi \geq 1$.

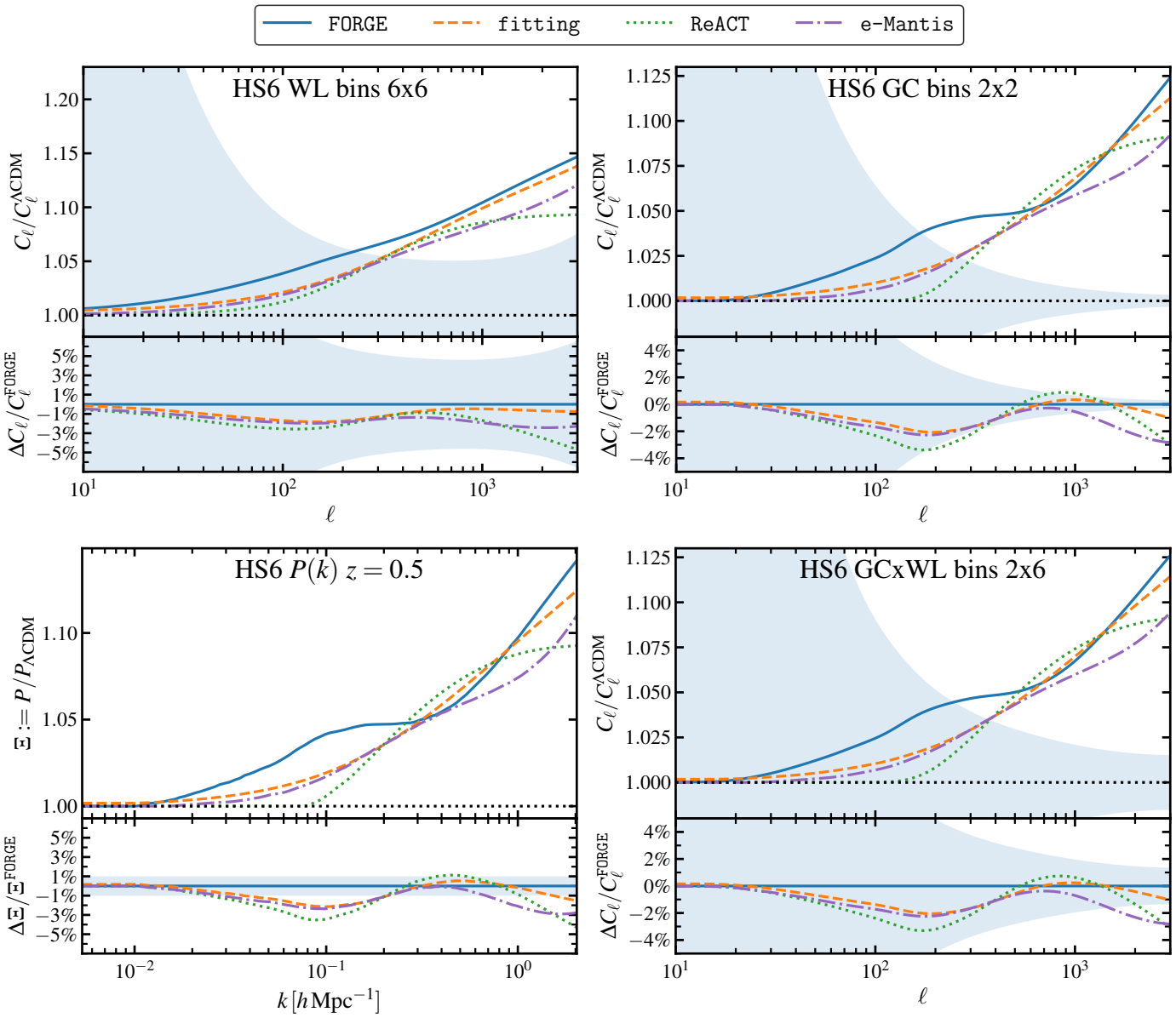

Fig. 2. Comparisons between all models considered for the HS6 fiducial cosmology for angular power spectra of weak lensing (*top left*), galaxy clustering (*top right*) and their cross-correlation (*bottom right*) as well as for the matter power spectrum (*bottom left*). We show these for the weak lensing redshift bin 6 and galaxy clustering bin 2, with the matter power spectrum plotted at $z = 0.5$, the redshift at which the kernels of both observables peak for the chosen bins. For all angular spectra, the error bands taken from the diagonal of the covariance are also shown.

4.2. Forecasting errors and biases

4.2.1. Forecasting errors for WL

We first compare errors obtained by running MCMC using the synthetic data created by one of the four nonlinear models and fitting it by the same model. In this case, by definition, we recover the input parameters that were used to create the synthetic data. Our interests are constraints on the $|f_{R0}|$ parameter and cosmological parameters. We first consider the WL-only case. In this case, we impose a tight Gaussian prior on the spectral index, n_s, taken from the Planck results (Planck Collaboration: Aghanim et al. 2020) and on the baryon density parameter, using Big Bang nucleosynthesis constraints (Pisanti et al. 2021):

$$n_\mathrm{s} = 0.96 \pm 0.004, \quad \Omega_\mathrm{b}h^2 = 0.022445 \pm 0.00036\,, \tag{28}$$

as we do not expect to obtain strong constraints on these parameters from WL alone. The parameters that are used in the MCMC runs are summarised in Table 5, including their fiducial values and prior ranges. As a convergence criterion, we use a Gelman–Rubin (Gelman & Rubin 1992) value of $R - 1 < 0.01$ for each individual sampling parameter using `MontePython`. For post-processing chains, we use `GetDist` (Lewis 2019,).

Figure 3 shows the 2D contours of the constraints on parameters, and Table 6 and Table 7 summarise constraints on $\log_{10}|f_{R0}|$ for the optimistic and pessimistic settings. The constraints on cosmological parameters are consistent among the four different nonlinear models, while we see some notable differences in the constraints on $|f_{R0}|$. In the case of the optimistic setting, `e-Mantis` gives the tightest constraints, which is also closer to Gaussian. The fitting formula agrees with `e-Mantis` for

Table 5. Fiducial values and flat prior ranges for cosmological parameters, $|f_{R0}|$ and IA parameters.

$\Omega_c h^2$	$100h$	$\ln(10^{10}A_s)$	$\mathcal{A}_{IA}$	η_{IA}	$\log_{10}\|f_{R0}\|$
Fiducial					
0.12056	67	3.05685	1.71	−0.41	−5.30103
Prior					
[0.005, 1]	[10, 150]	[2.7, 3.3]	[0, 12.1]	[−7, 6.17]	[−7, −4]

small $\log_{10}|f_{R0}|$ but has a weaker constraint for larger $\log_{10}|f_{R0}|$. Constraints from FORGE agree with the fitting formula for large $\log_{10}|f_{R0}|$, but give a weaker constraint for small $\log_{10}|f_{R0}|$. The degeneracy between $\log_{10}|f_{R0}|$, $\Omega_c h^2$ and $\ln(10^{10}A_s)$ also presents some notable differences. The degeneracy for larger $\log_{10}|f_{R0}|$ is different for the fitting formula when compared with e-Mantis and FORGE. This could be attributed to the fact that the fitting formula does not include any cosmological parameter dependence in the prediction for $\Xi(k, z)$. We observe similar agreements and disagreements for the pessimistic setting, but the agreement among e-Mantis, FORGE and the fitting formula is better, particularly for small $\log_{10}|f_{R0}|$. ReACT gives a weaker constraint on $\log_{10}|f_{R0}|$, but the constraints on cosmological parameters are consistent among the four different models. This is due to the weak cosmology dependence of $\Xi(k, z)$, as discussed in Winther et al. (2019), and the fact that the constraint on cosmological parameters is coming from the ΛCDM power spectrum, which is common in all these four models.

We compare the constraints from the MCMC analysis with the Fisher Matrix forecast and show this in Fig. 3. The MontePython pipeline can be used as a Fisher Matrix forecast tool (Casas et al. 2024), which was shown to agree very well with previous Fisher Matrix forecasts given by EC20. We note that in Casas et al. (2023), no prior was imposed on n_s and $\Omega_b h^2$, thus no direct comparison is possible. Instead, we use MontePython as a Fisher Matrix forecast tool and compare the result with the MCMC analysis to validate the Fisher Matrix forecast. To be consistent with Casas et al. (2023), we use the fitting formula as the nonlinear model. The 1σ error is very consistent: the Fisher Matrix forecast gives 0.111 while the MCMC analysis gives 0.116. The constraint from MCMC is non-Gaussian and the 1D posterior has a slightly longer tail for large $\log_{10}|f_{R0}|$. The constraints on cosmological parameters agree very well between the Fisher Matrix forecast and the MCMC result.

4.2.2. Assessing biases for WL

Next, we create the synthetic data using FORGE and we fit it by different nonlinear models to assess the biases in the recovered parameters due to the difference in the nonlinear modelling. We selected FORGE to create the data because it has a relatively narrow prior range for $|f_{R0}|$ and we encountered a problem with that range when using FORGE as the model to fit when including baryonic effects. We should note that FORGE has a larger discrepancy from other nonlinear models. As we discussed before, this could be attributed to emulation errors and calibration with halofit. In particular, this choice is disadvantageous to ReACT as the difference of the prediction for Ξ from FORGE is the largest. Thus, the estimation of bias in recovered parameters presented in this section is conservative, particularly for ReACT

Figure 4 shows the 2D contours of the constraints on parameters, and Table 8 and Table 9 summarise constraints on $\log_{10}|f_{R0}|$ for the optimistic and pessimistic settings.

We first start from the optimistic setting. Since ReACT is valid only up to $k = 3h\,\mathrm{Mpc}^{-1}$, we will not include ReACT in this case. The fitting formula and e-Mantis recover the input parameters within 1σ. In the case of e-Mantis, cosmological parameters are well recovered, but there is a slight bias in the recovered $\log_{10}|f_{R0}|$. On the other hand, a slight bias appears in h in the case of the fitting formula. This may be attributed to the fact that the fitting formula does not include cosmological parameter dependence in the prediction of Ξ. To quantify the bias, we define a 1D bias as

$$B_{1D} = \frac{\mu - \mu_{\mathrm{FORGE}}}{\sigma_{\mathrm{FORGE}}}, \tag{29}$$

where μ and σ are the mean and 1σ error computed from the 1D marginalised posteriors. If $\mu > \mu_{\mathrm{FORGE}}$ ($\mu < \mu_{\mathrm{FORGE}}$), we use the upper (lower) 68.3% confidence interval to obtain σ_{FORGE}. The 1D bias for $\log_{10}|f_{R0}|$ is $B_{1D} = 0.273$ and 0.602 for the fitting formula and e-Mantis, respectively.

We see a similar result in the pessimistic setting for e-Mantis and the fitting formula. In this case, the input $\log_{10}|f_{R0}|$ is well within 1σ although h is again slightly biased. The 1D bias for $\log_{10}|f_{R0}|$ is $B_{1D} = 0.441$ and 0.518 for the fitting formula and e-Mantis, respectively while the 1D bias for h is 0.988 and 0.812 for the fitting formula and e-Mantis, respectively. On the other hand, we find a bias in the recovered $\log_{10}|f_{R0}|$ when ReACT is used. As mentioned above, we should bear in mind that the choice of FORGE as a fiducial is disadvantageous for ReACT. In the case of ReACT, the difference from FORGE leads to biases also in the cosmological parameters leading to larger $\Omega_c h^2$ and smaller $\ln(10^{10}A_s)$ and h. This is due to the different k dependence of Ξ between ReACT and FORGE and this is compensated by adjusting cosmological parameters as well as $\log_{10}|f_{R0}|$ leading to a stronger bias. The 1D bias for $\log_{10}|f_{R0}|$ reaches $B_{1D} = 3.11$. We will discuss how to mitigate this bias by correcting the prediction for the fiducial cosmology and including theoretical errors in Sect. 6.

Since the FORGE prediction deviates from the other three models, we also performed the same analysis using e-Mantis as fiducial data for the pessimistic setting. We show the results in Appendix A. Qualitatively, we obtain consistent results. As we can see from Fig. A.1, the 2D contours overlap in the same way as the case where FORGE is used as fiducial, while the mean of $\log_{10}|f_{R0}|$ obtained by ReACT is closer to the input value. However, we still find that the 1D bias is at 3 sigma level due to the smaller errors of e-Mantis compared with FORGE.

4.2.3. 3×2pt analysis

We consider the constraints from 3×2pt statistics by adding GC_{ph} and its cross-correlation XC_{ph} to WL. In this paper, we use a simple model of the scale-independent linear bias as in ΛCDM. This assumption needs to be reexamined in $f(R)$ gravity as the scale-dependent growth will lead to a scale-dependent bias. However, the linear bias assumption also needs to be relaxed even in ΛCDM and it is beyond the scope of this paper to implement a more complete bias description. For this reason, we consider the pessimistic case only and focus our attention on a

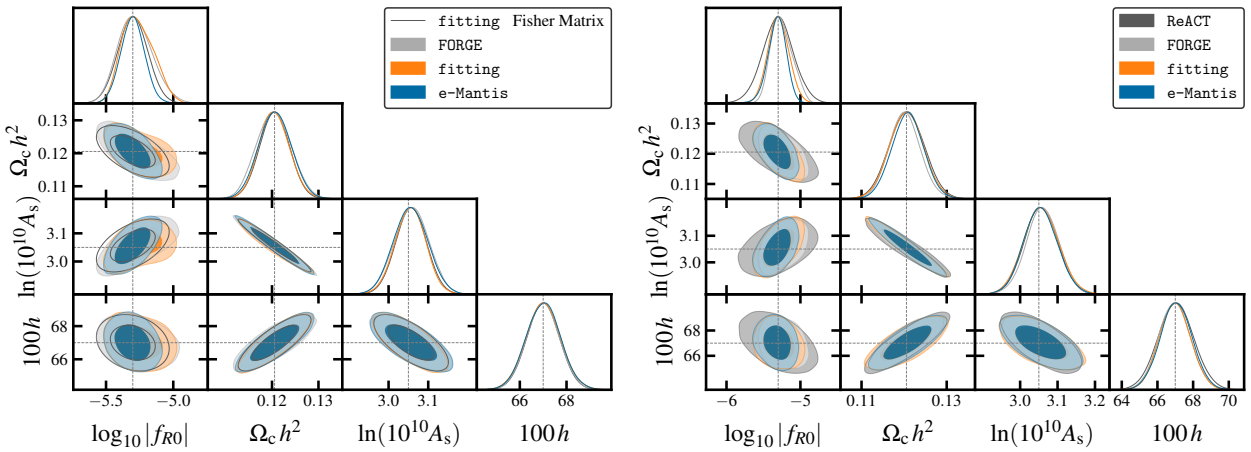

Fig. 3. Constraints on parameters where the same model is used to create the synthetic data and perform the fitting such that the input parameters are reproduced. *Left panel*: WL optimistic case. *Right panel*: WL pessimistic case.

Table 6. The mean, standard deviation and 68.3% upper and lower limit of $\log_{10}|f_{R0}|$ for the WL optimistic setting where the same prediction is used to create the data and the model.

WL optimistic				
	Mean	S.d.	Lower	Upper
fitting	−5.264	0.116	−5.397	−5.160
e-Mantis	−5.301	0.088	−5.388	−5.214
FORGE	−5.280	0.127	−5.419	−5.167
fitting (Fisher)	−5.301	0.111	−5.412	−5.190

Table 7. The mean, standard deviation and 68.3% upper and lower limit of $\log_{10}|f_{R0}|$ for the WL pessimistic setting where the same model is used to create the data and the model.

WL pessimistic				
	Mean	S.d.	Lower	Upper
fitting	−5.252	0.144	−5.419	−5.130
e-Mantis	−5.318	0.127	−5.430	−5.186
FORGE	−5.296	0.143	−5.441	−5.159
ReACT	−5.305	0.216	−5.516	−5.095

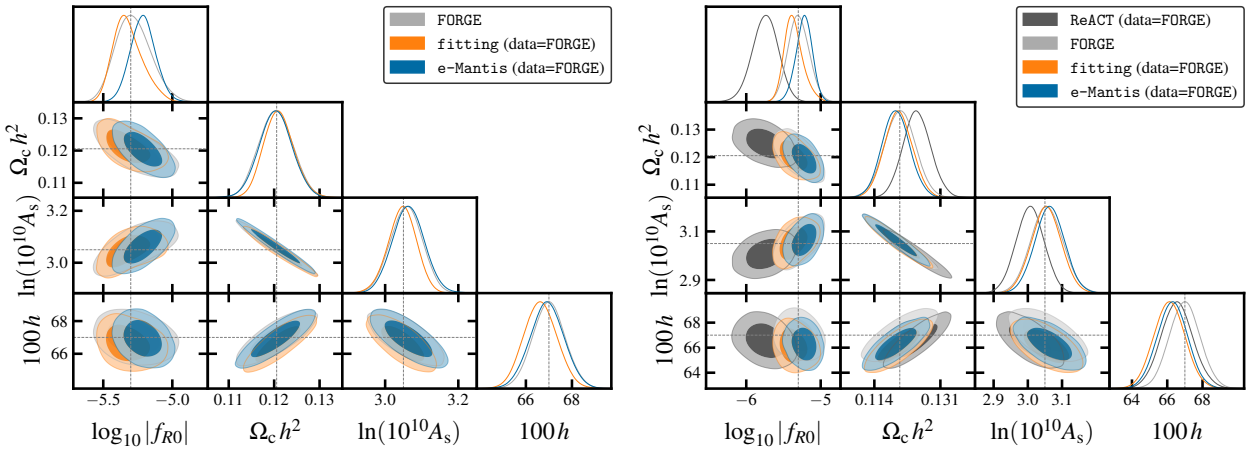

Fig. 4. Bias due to different nonlinear modelling. The synthetic data are created by FORGE and fitted by four different models. *Left panel*: WL optimistic case. *Right panel*: WL pessimistic case.

Table 8. The mean, standard deviation and 68.3% lower and upper limit of $\log_{10}|f_{R0}|$ for the optimistic setting where the data are generated by FORGE and it is fitted by different nonlinear models.

WL optimistic					
	Mean	S.d.	Lower	Upper	B_{1D}
Data = FORGE					
fitting	−5.318	0.108	−5.443	−5.236	0.273
e-Mantis	−5.212	0.092	−5.302	−5.121	0.602

Table 9. The mean, standard deviation and 68.3% lower and upper limit of $\log_{10}|f_{R0}|$ for the pessimistic setting where the data are generated by FORGE and it is fitted by different nonlinear models.

WL pessimistic					
	Mean	S.d.	Lower	Upper	B_{1D}
Data = FORGE					
fitting	−5.360	0.129	−5.503	−5.258	0.441
e-Mantis	−5.225	0.111	−5.326	−5.109	0.518
ReACT	−5.747	0.178	−5.921	−5.572	3.110

comparison with the Fisher Matrix forecast and check if the bias when FORGE is used to create the data becomes worse with the increasing statistical power. We also do not include ReACT as we already observe a significant bias with WL only.

We have 10 scale-invariant bias parameters, one for each redshift bin. For the 3×2pt analysis, we vary n_s, but we still impose a tight Gaussian prior on $\Omega_b h^2$ as it is unlikely to get a tighter constraint than the Big Bang nucleosynthesis constraints.

The parameters used in addition to the WL analysis are listed in Table 10. We did not change the fiducial bias parameters from EC20 as there is no prediction of the linear bias in $f(R)$ gravity. The fiducial bias adopted in EC20 will need to be improved even in ΛCDM. Our prime focus is the study of the effect of nonlinear models and we vary the linear bias in the MCMC analysis resulting in different constraints depending on the nonlinear model used. The observed covariance is built from synthetic data. Thus the effect of $f(R)$ gravity is included in the covariance. We independently checked that the constraints on parameters do not change by using the ΛCDMcovariance. Thus we expect that the effect of $f(R)$ on galaxy bias has also negligible effects on the covariance.

The left-hand side of Fig. 5 shows the 2D contours of the constraints on parameters, and Table 11 summarises constraints on $\log_{10}|f_{R0}|$ where the same nonlinear model is used for the data and the model. The right-hand side of Fig. 5 shows the 2D contours of the constraints on parameters, and Table 12 summarises constraints on $\log_{10}|f_{R0}|$ as well as 1D bias for $\log_{10}|f_{R0}|$ when `FORGE` is used to create the data.

In the case where the same nonlinear model is used for the data and the model, errors are consistent between the fitting formula and `e-Mantis`, although we still observe the same longer tail for larger $\log_{10}|f_{R0}|$ for the fitting formula as we observe for WL. `FORGE` gives slightly tighter constraints on $\log_{10}|f_{R0}|$. Constraints on cosmological parameters are very consistent among the three different nonlinear models. The Fisher Matrix forecast using the fitting formula is also very consistent with the MCMC results. We note that errors from 3×2pt analysis in the pessimistic setting are comparable to or better than WL alone in the optimistic setting. This demonstrates the strength of the 3×2pt analysis although we should bear in mind the limitation of the bias model used in this analysis.

We find that the increased statistical power does not degrade the agreement between `FORGE` and `e-Mantis`. When `FORGE` is used to create the data, the input parameters are well recovered by `e-Mantis`, although h is slightly biased as in the WL-only case. The 1D bias for $\log_{10}|f_{R0}|$ is given by 0.150. In the case of the fitting formula, the bias in cosmological parameters becomes worse compared with WL. This could be attributed to the fact that the fitting formula does not take into account the cosmological parameter dependence of Ξ as mentioned before. The 1D bias for $\log_{10}|f_{R0}|$ also becomes slightly larger, with a value of $B_{1D} = 0.667$.

5. Baryonic effects

5.1. Adding baryonic effects using `BCemu`

In this section, we study the impact of baryonic effects on the constraints on the $f(R)$ parameter. We use the seven-parameter emulator of baryonic effects called `BCemu` (Giri & Schneider 2021) and we assume that the baryonic effects and the modified gravity effects can be treated independently as discussed in Sect. 3 (Arnold & Li 2019). The `BCemu` parameters govern the gas profiles and stellar abundances in haloes. It is not our intention to study in detail the baryonic effects. We use the default values for the baryonic parameters provided with `BCemu` and use the full prior range. We choose not to include any redshift dependence. This is because we need to impose a tight prior range for these parameters at $z = 0$ to keep them within the prior range and we may miss important degeneracies between baryonic parameters and $\log_{10}|f_{R0}|$. We found that two baryonic parameters, $M_c^{\rm Giri\ etal.}$ and θ_{ej}, have the strongest effects on the matter power spectrum and they are well constrained in the presence of the $f(R)$ parameter. Thus, we only vary these two parameters. In `BCemu`, the gas profile is modelled as a cored double power law. $M_c^{\rm Giri\ et.al.}$ controls the dark matter halo mass dependence of the logarithmic slope of the first-cored power law. It allows the profile to become less steep than the Navarro et al. (1996) one for $M < M_c^{\rm Giri\ et.al.}$. The parameter θ_{ej} determines the scale radius (with respect to the virial radius) of the second-cored power law. Table 13 summarises the fiducial values and priors for baryon parameters. We will use the dimensionless parameter M_c that is related to the one defined in Giri & Schneider (2021) via $M_c := M_c^{\rm Giri\ et.al.}/M_\odot$ where $M_\odot$ is the solar mass.

5.2. Effects of adding baryons

Figure 6 shows the effect of changing $\log_{10}|f_{R0}|$, h and two baryon parameters on the ratio of the WL angular power spectrum to the ΛCDM one. Baryonic effects introduce scale-dependent modifications to the angular power spectrum that are similar to the effect of $f(R)$ at $\ell > 100$. On the other hand, h changes the overall amplitude of the angular power spectrum as it changes Ω_m. We will see that the interplay between these parameters leads to interesting degeneracies.

Figure 7 shows the 2D contours of the WL constraints on parameters for the optimistic settings with baryons and Table 14 summarises constraints on $\log_{10}|f_{R0}|$.

In the case of the optimistic setting, when `e-Mantis` is used as the data and the model, the two baryonic parameters degrade the constraints on $\log_{10}|f_{R0}|$: the 1σ error becomes 0.281 while it was 0.088 without baryons. However, it is important to stress that $\log_{10}|f_{R0}|$ and the two baryonic parameters are still constrained well within the prior range of these parameters. It means that the impact of modified gravity on the high-k tail of the matter power spectrum is not washed out by the baryonic feedback, and we can still distinguish between the effect of $f(R)$ and baryons.

Nevertheless, if `FORGE` is used as the data, the difference between `FORGE` and `e-Mantis` in the matter power spectrum can be absorbed by the shift in $\log_{10} M_c$ and h, and the constraint on $\log_{10}|f_{R0}|$ is shifted along the degeneracy direction between $\log_{10}|f_{R0}|$ and $\log_{10} M_c$ as well as h. This can be understood from Fig. 6. The difference of the scale dependence between `e-Mantis` and `FORGE` at $\ell > 100$ can be adjusted by decreasing $\log_{10}|f_{R0}|$ and decreasing $\log_{10} M_c$. This leads to a lower amplitude that can be adjusted by decreasing h. Due to the combination of these effects, the best-fit $\log_{10}|f_{R0}|$ becomes smaller. We note that `e-Mantis` does not include the h dependence in the prediction for $\Xi(k,z)$ explicitly, but it depends on h implicitly through $\sigma_8^{\Lambda\rm CDM}$. However, this dependence of h on $\Xi(k,z)$ is very weak. To compute the power spectrum in $f(R)$ we use $P_{f(R)} = \Xi(k,z)P_{\Lambda\rm CDM}(k,z)$. The `e-Mantis` emulator provides us with the first factor, $\Xi(k,z)$, and a ΛCDM emulator provides the last factor. The h dependence comes from the ΛCDM power spectrum not from $\Xi(k,z)$.

As a result, we obtain a 95.5% upper limit of $\log_{10}|f_{R0}|$ as $\log_{10}|f_{R0}| < -5.477$, which is incompatible with the input value of $\log_{10}|f_{R0}| = -5.301$. This is partly due to prior volume effects shifting the contour to lower values of $\log_{10}|f_{R0}|$, caused by the strong degeneracy between $\log_{10}|f_{R0}|$ and $\log_{10} M_c$ as well as h.

To confirm this, we obtain a profile likelihood for $\log_{10}|f_{R0}|$ by fixing $\log_{10}|f_{R0}|$ and finding a minimum χ^2 by varying other

Table 10. Additional parameters that are varied in the 3×2pt analysis.

n_s	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}
Fiducial										
0.96	1.0998	1.2202	1.2724	1.3166	1.3581	1.3998	1.4446	1.4965	1.5652	1.7430
Prior										
[0.8, 1.2]	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.

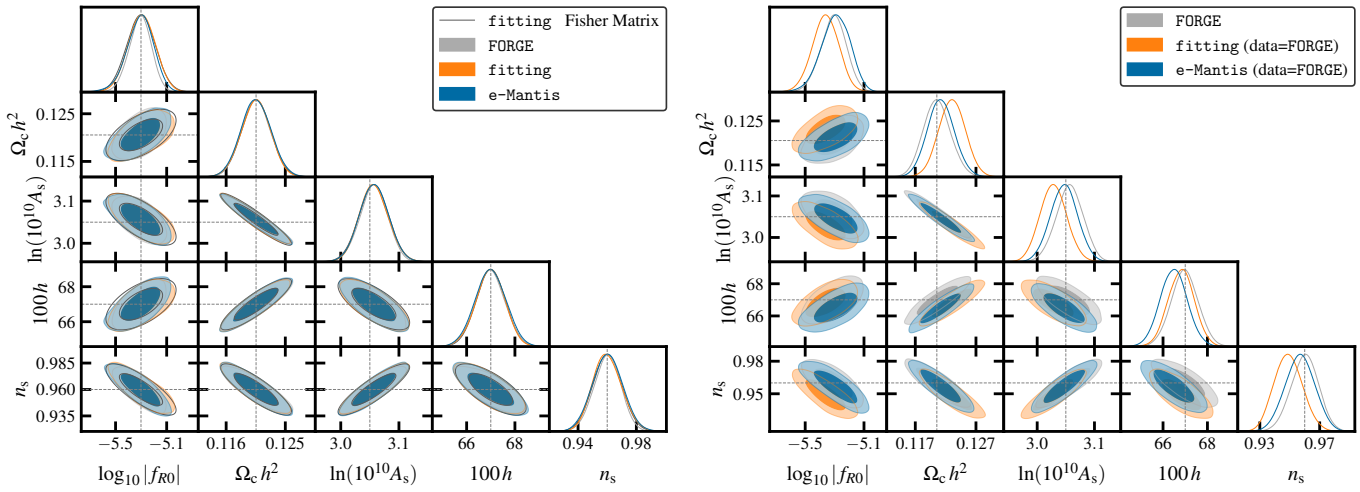

Fig. 5. Constraints from 3×2pt analysis in the pessimistic setting. In the left panel, the same nonlinear model is used for the data and the fitting while in the right panel, the data are generated by FORGE.

Table 11. The mean, standard deviation and 68.3% lower and upper limit of $\log_{10}|f_{R0}|$ for the pessimistic setting where the same prediction is used to create the data and the model.

3×2pt pessimistic				
	Mean	S.d.	Lower	Upper
fitting	−5.297	0.107	−5.404	−5.191
e-Mantis	−5.309	0.107	−5.405	−5.198
FORGE	−5.308	0.087	−5.389	−5.215
fitting Fisher	−5.301	0.112	−5.413	−5.189

Table 12. The mean, standard deviation and 68.3% lower and upper limit of $\log_{10}|f_{R0}|$ for the pessimistic setting where the data are generated by FORGE and it is fitted by different nonlinear models.

3×2pt pessimistic					
	Mean	S.d.	Lower	Upper	B_{1D}
Data = FORGE					
fitting	−5.362	0.091	−5.451	−5.271	0.667
e-Mantis	−5.294	0.099	−5.382	−5.187	0.151

parameters. The profile likelihood was obtained by PROSPECT (Holm et al. 2023,). This is shown in Fig. 8. The global best-fit value for $\log_{10}|f_{R0}|$ is obtained as $\log_{10}|f_{R0}| = -5.765$, which is larger than the mean of the 1D marginalised posterior $\log_{10}|f_{R0}| = -6.087$. Figure 6 also shows the prediction of e-Mantis with the best-fit values, which agrees well with FORGE. The $\Delta\chi^2$ for the input value of $\log_{10}|f_{R0}| = -5.301$ is found as $\Delta\chi^2 = 2.657$, thus it is still within 2σ. We also observe that the $\Delta\chi^2$ curve becomes flat for smaller values of $\log_{10}|f_{R0}|$.

We also tested the fitting formula as a model to fit to the data generated by FORGE. In this case, the posterior is highly non-Gaussian and the chains did not converge well.

This implies that we can break the degeneracy between $f(R)$ gravity and baryonic effects, but we need to have an accurate nonlinear model for the $f(R)$ gravity model. Otherwise, we could obtain a significantly biased result in terms of the 1D marginalised constraint. We note that if we have a physical prior on $\log_{10} M_c$ from baryonic physics, we could break the degeneracy between $\log_{10}|f_{R0}|$ and $\log_{10} M_c$. For example, in the case where FORGE was used as the data and they were modelled by e-Mantis, the mean of $\log_{10} M_c$ is significantly shifted to a smaller value as we can see in Fig. 7. If we had a prior on $\log_{10} M_c$ to prevent this, this bias could be avoided. To test this idea, we also ran an analysis imposing a prior on $\log_{10} M_c$ and showed the result in Fig. 7. We used a Gaussian prior with the width of 0.2, which was estimated from the weak lensing informed gas and stellar mass fraction measurements of massive haloes by Grandis et al. (2024). This prior is consistent with the error on $\log_{10} M_c$ that we obtained by using e-Mantis both for data and model. We observe that the bias is relaxed slightly, however, the degeneracy between $\log_{10} M_c$ and $\log_{10}|f_{R0}|$ is quite strong and this prior is not enough to alleviate the bias in the marginalised constraints. An improved prior from external data is needed to fully break this degeneracy.

We assumed that the effect of $f(R)$ gravity and baryonic effect can be treated independently. Recently, the coupling between baryonic feedback and cosmology has been studied showing that the combined effect of baryonic and non-baryonic suppression mechanisms is greater than the sum of its parts for decaying dark matter (Elbers et al. 2024). The effect of this coupling on the degeneracy needs to be studied in the future.

Next, we consider the pessimistic setup. Figure 9 shows the 2D contours of the WL constraints on parameters for the optimistic settings with baryons and Table 15 summarises constraints on $\log_{10}|f_{R0}|$. In this case, due to the weaker constraining power, the 99.7% confidence level upper bound on $\log_{10}|f_{R0}|$ reaches the prior boundary of $\log_{10}|f_{R0}| = -7$. Also, the 95.5% confidence level upper bound of θ_{ej} is bounded by the prior. If the

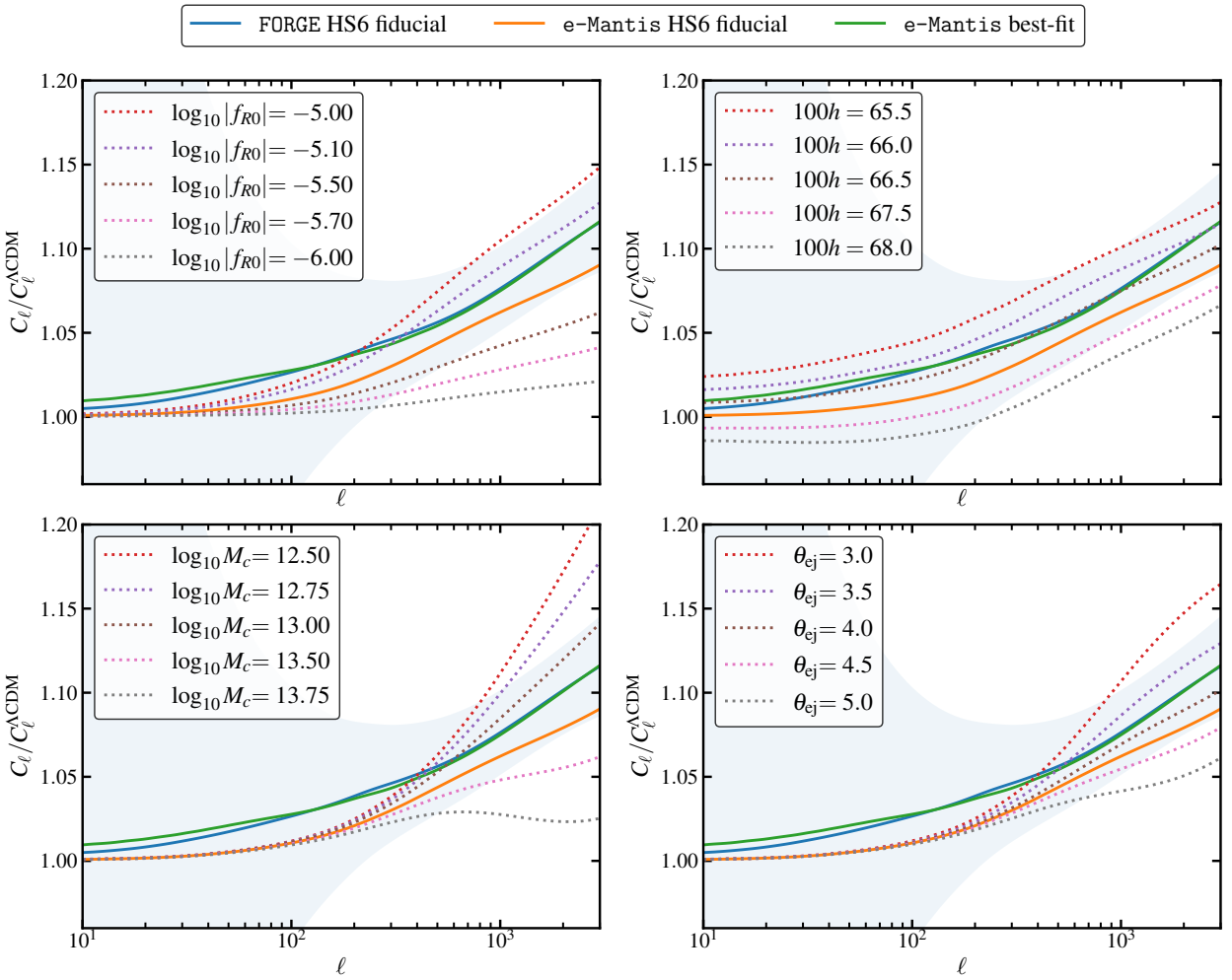

Fig. 6. The effect of $\log_{10}|f_{R0}|$, two baryonic parameters $\log_{10} M_c$ and θ_{ej}, and h on the WL angular power spectrum in bin 10. The dotted lines show predictions of `e-Mantis` with varying these parameters individually. The plot shows the ratio to the ΛCDM power spectrum. The plot includes the prediction of `FORGE` and `e-Mantis` with the fiducial parameters and the `e-Mantis` prediction using the best-fit values ($\log_{10}|f_{R0}| = -5.765, 100h = 65.736, \log_{10}(M_c) = 12.999$) with the data created by `FORGE`. We can see that the `e-Mantis` prediction can fit the `FORGE` by adjusting these parameters. We note that the cosmological parameters for the ΛCDM power spectrum $C_\ell^{\Lambda\mathrm{CDM}}$ is fixed to be the fiducial ones when we vary $100h$.

Table 13. Fiducial values and priors for `BCemu` parameters.

$\log_{10} M_c$	θ_{ej}	f_b	μ	γ	δ	η	η_δ
Fiducial							
13.32	4.235	0.186	0.93	2.25	6.40	0.15	0.14
Prior							
[11, 15]	[2 8]	[0.10, 0.25]	N.A.	N.A.	N.A.	N.A.	N.A.

Table 14. The mean, standard deviation and 68.3% lower and upper limit of $\log_{10}|f_{R0}|$ for the WL optimistic setting with baryons.

WL optimistic				
	Mean	S.d.	Lower	Upper
e-Mantis	−5.353	0.281	−5.586	−5.058
Data = FORGE				
e-Mantis	−6.087	0.424	−6.417	−5.536

Table 15. The mean, standard deviation and 68.3% lower and upper limit of $\log_{10}|f_{R0}|$ for the WL pessimistic setting with baryons.

WL pessimistic				
	Mean	S.d.	Lower	Upper
e-Mantis	−5.482	0.615	−5.909	−4.699
Data = FORGE				
e-Mantis	−5.951	0.575	−6.549	−5.262
fitting	−5.830	0.586	−6.305	−5.045
ReACT	−6.147	0.449	−6.581	−5.593

data are generated by `FORGE`, strong degeneracies appear among $\log_{10}|f_{R0}|$ and two baryonic parameters. The input $\log_{10}|f_{R0}|$ is consistent within 1σ for the fitting formula and `e-Mantis` although the prior bound $\log_{10}|f_{R0}| = -7$ is reached at 95.5% confidence level lower bound. On the other hand, for `ReACT`, the input value is outside the 95.5% confidence level upper bound.

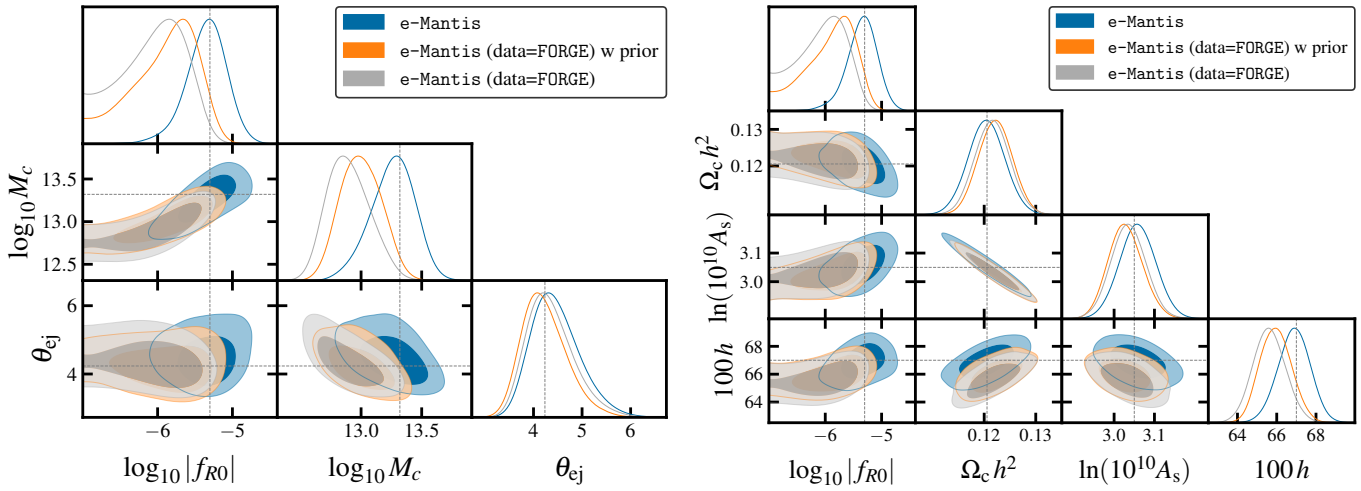

Fig. 7. Constraints on parameters from the WL optimistic setting with baryons. We show the case where `e-Mantis` is used to create both the data and fitting and the case where `FORGE` is used to create the data and it is fitted by `e-Mantis`. We also include the case where we impose a Gaussian prior on $\log_{10} M_c$ with the width of 0.2 in the case of `FORGE` fiducial data. *Left panel*: Constraints on baryon parameters. *Right panel*: Constraints on cosmological parameters.

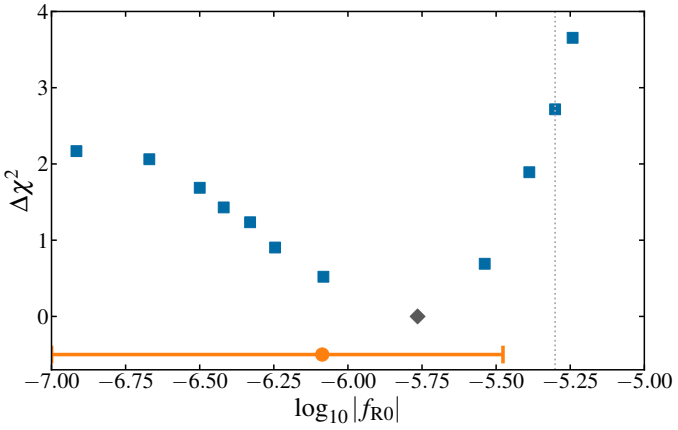

Fig. 8. The profile likelihood for $\log_{10}|f_{R0}|$ where `FORGE` is used to create data fitted by `e-Mantis` with baryons in the optimistic setting of WL. The square points show $\Delta\chi^2$ from the best-fit indicated by a diamond. The 95.5% confidence interval from the 1D marginalised posterior is shown as the error bar. The vertical dotted line indicates the input value of $-\log_{10}|f_{R0}| = 5.301$. We see that the 1D marginalised posterior is shifted to smaller values of $\log_{10}|f_{R0}|$.

As we already observed without baryons, h is biased to a lower value.

We also checked the case where fiducial data in generated by `e-Mantis`. As we can see in Fig. A.2, unlike the case without baryons, the agreements between `e-Mantis`, `fitting` and `ReACT` are improved significantly compared with the case where `FORGE` is used for fiducial data. Particularly, for `ReACT`, the mean of $\log_{10}|f_{R0}|$ is now consistent with the input value within 1σ.

5.3. Constraints on $|f_{R0}|$ with the ΛCDM data

Based on these results, we consider the pessimistic setting and use `e-Mantis` to obtain an upper limit of $\log_{10}|f_{R0}|$ in the presence of baryonic effects to be conservative.

Figure 10 shows the 2D contours of the WL constraints on parameters for the pessimistic settings where ΛCDM is used to create the data. We find that the recovered cosmological parameters are consistent with the input values. The mean of 1D marginalised constraints of baryonic parameters are slightly biased due to the strong degeneracy between them, but it is still consistent within 1σ. We obtain the 95.5% confidence level upper bound on $\log_{10}|f_{R0}|$ as

$$\log_{10}|f_{R0}| < -5.21\,. \tag{30}$$

We note that this bound depends on the prior $\log_{10}|f_{R0}| = -7$. To obtain the prior-independent bound, we follow the approach presented in Gordon & Trotta (2007), Piga et al. (2023) and Kou et al. (2023). The ratio of the marginalised posterior and prior is given by

$$b(x \mid d, p) = \frac{\mathcal{P}(x \mid d, p)}{p(x)}\,, \tag{31}$$

where x is the parameter we are interested in (i.e. $\log_{10}|f_{R0}|$), d is the data and p is the prior, and $\mathcal{P}$ is the Posterior. The Bayes factor $B(x_1, x_2)$, which quantifies the support of the models with $x = x_1$ over the models with $x = x_2$ is given by

$$B(x_1, x_2) = \frac{b(x_1 \mid d, p)}{b(x_2 \mid d, p)} = \frac{\mathcal{L}(d \mid x_1)}{\mathcal{L}(d \mid x_2)}, \tag{32}$$

where $\mathcal{L}(d \mid x)$ is the marginalised likelihood of the data for parameter x. For our purpose, we choose $x = \log_{10}|f_{R0}|$ and fix x_1 to be the upper bound of the prior, $x_1 = -7$. Following Gordon & Trotta (2007), we then use $B(x_1, x_2) = 2.5$ to find x_2 so that the model with $x = x_1$ is favoured compared to the model with $x = x_2$ at 95.5% confidence level. This gives the 95.5% confidence interval of $\log_{10}|f_{R0}|$ that does not depend on the prior. We note that this method applies only to a $\mathcal{L}(d \mid x)$ that is a monotonic function of x. Employing this technique, we obtain

$$\log_{10}|f_{R0}| < -5.58\,. \tag{33}$$

Finally, in Fig. 11, we present the constraints on Ω_m and $\log_{10}|f_{R0}|$ so that we can compare these with those in the literature (Schneider et al. 2020; Harnois-Déraps et al. 2023; Spurio Mancini & Bose 2023). It is not possible to make a direct

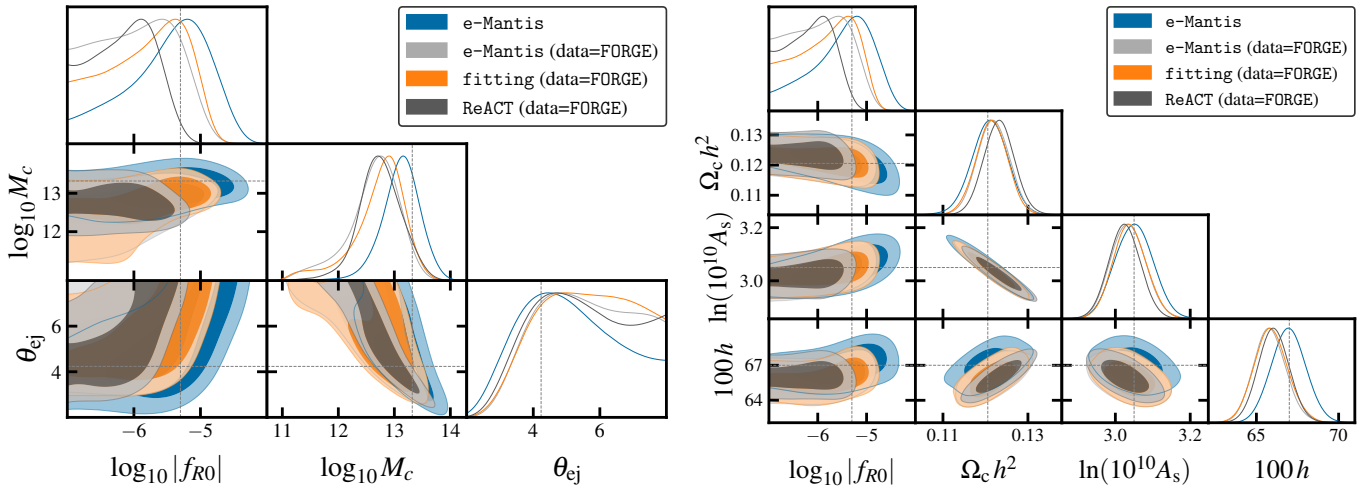

Fig. 9. Constraints on parameters from the WL pessimistic setting with baryons. As in Fig. 7, we show the case where `e-Mantis` is used to create both the data and fitting as well as the case where `FORGE` is used to create the data and it is fitted by `e-Mantis`, `ReACT` and the fitting formula. *Left panel*: Constraints on baryon parameters. *Right panel*: Constraints on cosmological parameters.

comparison due to various differences in the setting. Nonetheless, our constraint is comparable to the one presented in Schneider et al. (2020) [$\log_{10}|f_{R0}| < -5.3$] where a similar analysis was done by combining `BCemu` and the fitting formula. We cannot also make a comparison with Harnois-Déraps et al. (2023) as baryonic effects are not included in their analysis, but again the constraint is comparable [$\log_{10}|f_{R0}| < -5.24$]. On the other hand, our constraint is much weaker than the one found in Spurio Mancini & Bose (2023) that used `ReACT` with `BCemu`. We refer the readers to Spurio Mancini & Bose (2023) for possible explanations. In our analysis, `ReACT` tended to prefer a lower value of $\log_{10}|f_{R0}|$ when `FORGE` is used as data and this might be one of the reasons for this difference. Finally, a recent study by Tsedrik et al. (2024) used similar settings to forecast cosmic shear constraints on model-independent parametrisations of modified gravity theories with scale-independent linear growth. They found that a much better understanding of baryonic feedback is needed in order to detect a screening transition.

6. Theoretical error

Any theoretical prediction will in general come with an associated error and should be included in the likelihood if it is not subdominant to the statistical error. In this section, we discuss the implementation of theoretical errors in our pipeline. To be conservative, we will estimate the theoretical errors using `ReACT` and `FORGE` as these two give the most discrepant results. We will apply this method to `ReACT` for the pessimistic case of WL to check if we can remove the parameter biases that we observe for the data created by `FORGE` at the expense of enlarging error bars.

6.1. Adding uncorrelated theoretical errors to the likelihood

The implementation presented in this work is based on the work of Audren et al. (2013). We have adjusted the recipe to the full 3×2pt likelihood, although we only show the result for the WL analysis in this paper. Here we will discuss the idea behind the formulation. We calculate the angular power spectrum from the power spectrum using Eq. (1). If there is some uncertainty in the modelling of the nonlinear power spectrum, this propagates to the angular power spectrum. As Eq. (1) is a linear functional of the nonlinear matter power spectrum, $\mathcal{F}[P_{\delta\delta}]$, we can propagate the error on the power spectrum, $\Delta P_{\delta\delta}$, and find

$$\mathcal{F}[P_{\delta\delta} + \Delta P_{\delta\delta}] = \mathcal{F}[P_{\delta\delta}] + \mathcal{F}[\Delta P_{\delta\delta}] =: C_{ij}^{XY}(\ell) + E_{ij}^{XY}(\ell)\,, \quad (34)$$

where we have defined the angular power spectrum error $E_{ij}^{XY}(\ell)$. Following Audren et al. (2013), we define the relative error on the power spectrum

$$\Delta P_{\delta\delta}(k,z) := \alpha(k,z)\, P_{\delta\delta}(k,z)\,. \quad (35)$$

We take the most conservative approach and assume that the error is uncorrelated between different values of ℓ. To account for this uncorrelated theoretical error, we add a new nuisance parameter ε_ℓ for each multipole and define the shifted covariance as

$$\tilde{C}_{ij}^{XY}(\ell) := \hat{C}_{ij}^{XY}(\ell) + \varepsilon_\ell\, L^{1/2}\, E_{ij}^{XY}(\ell)\,. \quad (36)$$

The free parameters ε_ℓ have the following meaning: they quantify how many standard deviations the shifted covariance matrix is from the theoretically predicted one. We treat them as a random Gaussian variable with zero mean and a standard deviation of one. The normalisation factor $L^{1/2}$ where $L = \ell_{\rm max} - \ell_{\rm min} + 1$ will become clear soon. The log-likelihood function becomes

$$\tilde{\chi}^2(\varepsilon_\ell) = \sum_{\ell=\ell_{\rm min}}^{\ell_{\rm max}} \left[(2\ell+1) f_{\rm sky} \left(\frac{\tilde{d}_\ell^{\rm mix}(\varepsilon_\ell)}{\tilde{d}_\ell^{\rm th}(\varepsilon_\ell)} + \ln \frac{\tilde{d}_\ell^{\rm th}(\varepsilon_\ell)}{d_\ell^{\rm obs}} - N \right) + \varepsilon_\ell^2 \right]. \quad (37)$$

The quantities $\tilde{d}^{\rm th}$ and $\tilde{d}^{\rm mix}$ are constructed in the same way as $d^{\rm th}$ and $d^{\rm mix}$ [see Eqs. (15) to (17)] using the shifted covariance matrix $\tilde{C}_{ij}^{XY}(\ell)$. To include the theoretical error, we vary ε_ℓ and marginalise over them. To a very good approximation, this is equivalent to minimising the $\tilde{\chi}^2$ with respect to ε_ℓ at the level of the likelihood. Thus, we define our new χ^2 as the minimum of the $\tilde{\chi}^2$ with respect to ε_ℓ:

$$\chi^2 := \min_{\varepsilon_\ell \in \mathbb{R}^L} \tilde{\chi}^2. \quad (38)$$

The normalisation factor $L^{1/2}$ can be explained as follows. If we were to measure $\hat{C}^{\rm obs}(\ell) = \hat{C}^{\rm th}(\ell) + E(\ell)$, the minimisation would

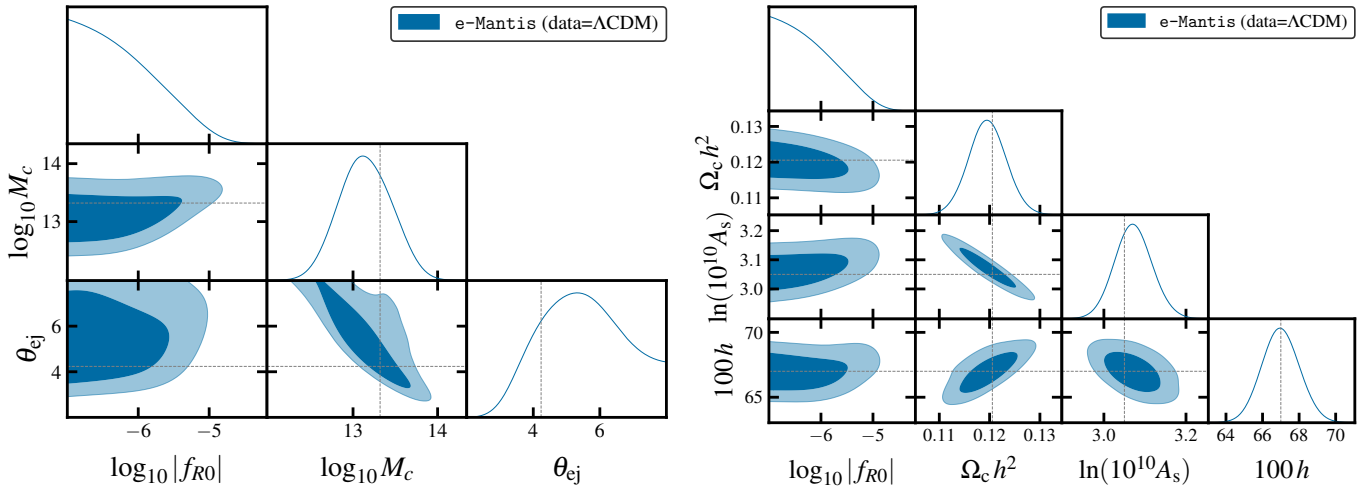

Fig. 10. Constraints on parameters from the WL pessimistic setting with baryons where the data are created in ΛCDM and it is fitted by `e-Mantis`. *Left panel*: Constraints on baryon parameters. *Right panel*: Constraints on cosmological parameters.

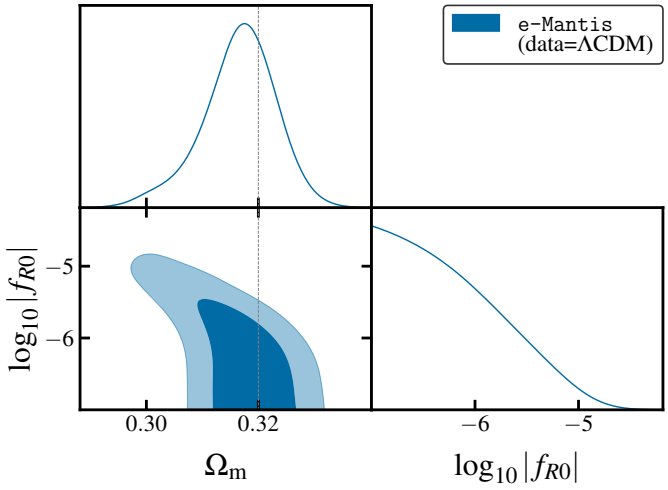

Fig. 11. The same as in Fig. 10, but we show the constraint on the derived quantity $\Omega_{\rm m}$.

find $\varepsilon_\ell = L^{-1/2}$. The resulting $\chi^2 = \sum_\ell \varepsilon_\ell^2 = \sum_\ell L^{-1} = 1$ would match our expectation that a one-sigma theoretical error for each ℓ results in an increase of the χ^2 by one.

The main ingredient of this formulation is the relative error function $\alpha(k, z)$ defined in Eq. (35). The construction of $\alpha(k, z)$ using `FORGE` and `ReACT` will be discussed in Appendix B.

6.2. Numerical implementation

The numerical computation of the theory error covariance $E_{ij}^{XY}(\ell)$ follows the prescription presented by Casas et al. (2024). For the minimisation, we can use the fact that all the free ε_ℓ are independent of each other, and we can do the minimisation for each multipole separately. We use Newton's method to find the minimum. For this, we have to compute the first and second derivatives of the likelihood with respect to ε_ℓ. For any single multipole, we find

$$\frac{\mathrm{d}\tilde{\chi}}{\mathrm{d}\varepsilon_\ell} = (2\,\ell + 1)\, f_{\rm sky} \left(\frac{\left(\tilde{d}_\ell^{\rm mix}\right)' + \left(\tilde{d}_\ell^{\rm th}\right)'}{\tilde{d}_\ell^{\rm th}} - \frac{\tilde{d}_\ell^{\rm mix}\,\left(\tilde{d}_\ell^{\rm th}\right)'}{\left(\tilde{d}_\ell^{\rm th}\right)^2} \right) + 2\,\varepsilon_\ell \,. \quad (39)$$

The derivatives of the determinants are computed using Jacobi's formula. This gives, for example

$$\left(\tilde{d}_\ell^{\rm th}\right)' = \det\left(\hat{C}_\ell^{\rm th}\right) L^{1/2}\, \mathrm{Tr}\left[\left(\hat{C}_\ell^{\rm th}\right)^{-1} E_\ell\right] , \quad (40)$$

and a similar expression for $\left(\tilde{d}_\ell^{\rm mix}\right)'$. We compute the inverse of the covariance numerically and obtain the second derivatives numerically from the first derivative by doing a double-sided three-point stencil. The minimisation would then need to be done for each multipole. For the pessimistic settings, this would correspond to a minimisation in a 1500-dimensional parameter space. To save time, we only do the minimisation on a logarithmically-spaced grid with 100 discrete values. The other values are obtained from an interpolating function. We checked that the effect of the interpolation does not change the results by more than 1%. The obtained ε_ℓ tend to vary continuously with ℓ, as they try to mimic the effects of changing other theory parameters to the observed power spectrum.

6.3. Application to `ReACT`

We first apply this method to the case where the data are created by `FORGE` and the parameter fitting is done by `ReACT`. Due to the significant bias in the recovered parameters for both $|f_{R0}|$ and cosmological parameters, we find that the inclusion of theoretical errors is not enough to mitigate the bias. Therefore, we additionally correct the prediction of `ReACT` by `FORGE` for the fiducial cosmology as

$$\Xi_{\texttt{ReACT-FORGE}} = \Xi_{\texttt{ReACT}} \left(\frac{\Xi_{\texttt{FORGE}}}{\Xi_{\texttt{ReACT}}} \right)_{\rm fiducial} . \quad (41)$$

The constraints on cosmological parameters are shown in the left panel of Fig. 12. In this case, as expected, we recover the input parameters in an unbiased way when the data are created by `FORGE`. We see that theoretical errors affect mainly $\log_{10}|f_{R0}|$ and h. In order to check a non-trivial case, next we create the data with `e-Mantis`. The result is shown in the right panel of Fig. 12. Without theoretical errors, the means of 1D marginalised constraints for $\log_{10}|f_{R0}|$, $\Omega_c h^2$ and $\ln(10^{10}A_s)$ are slightly biased compared with the input values. The inclusion of the theoretical errors largely resolves these biases not only by enlarging the error bars but also by making the means closer to the input values.

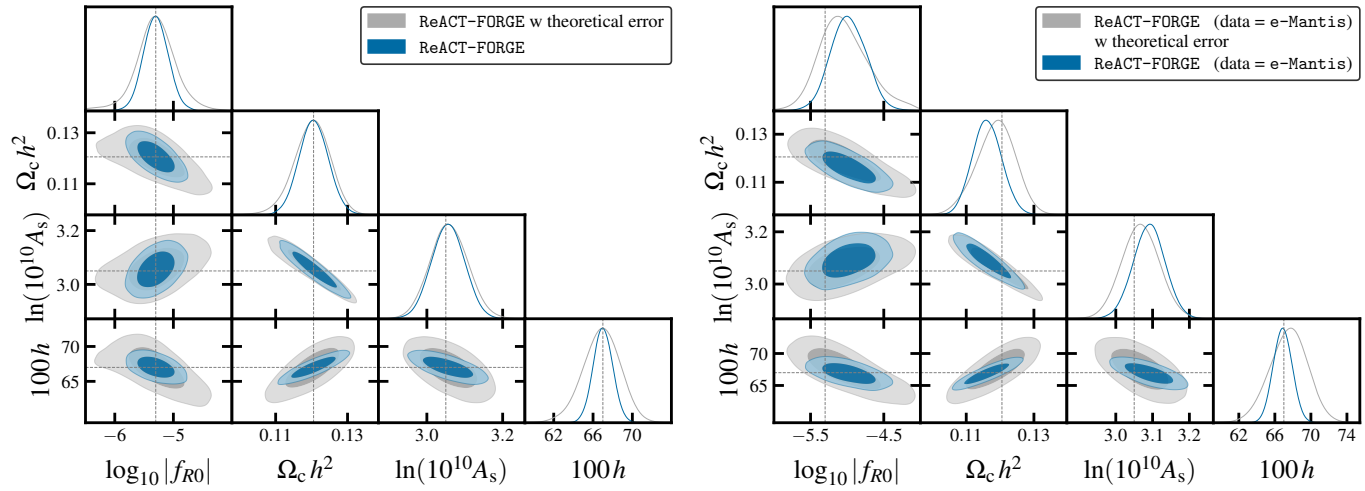

Fig. 12. The effect of adding theoretical errors. *Left panel*: Data are created by FORGE in the WL pessimistic setting and fitted by ReACT corrected by FORGE for fiducial cosmology, Eq. (41), with and without theoretical errors. *Right panel*: Data are created by e-Mantis in this case.

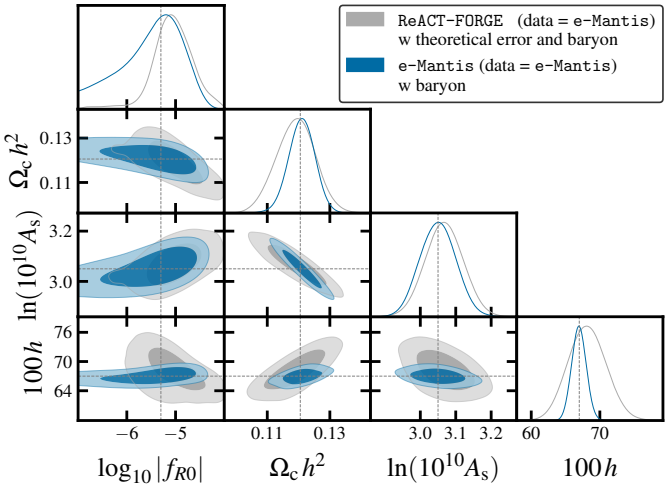

Fig. 13. The effect of adding theoretical errors in the presence of baryonic effects. Data are created by e-Mantis and fitted by ReACT corrected by FORGE for the fiducial cosmology with theoretical errors. For a comparison, we also show the result obtained by using e-Mantis to fit the same data.

We note that the inclusion of theoretical errors in this case is important to justify the rescaling of Ξ in Eq. (41). We also note that the theoretical error included here overestimates the errors significantly now that we corrected the prediction of ReACT by FORGE for the fiducial cosmology.

Finally, we study the impact of adding baryonic effects in the presence of theoretical errors. We again use e-Mantis as the data. The result is shown in Fig. 13. Due to the enlarged errors, constraints on parameters are not affected significantly by baryonic effects. Also, the means of the 1D marginalised constraint remain consistent with the input parameters. However, the inclusion of the theoretical errors changes the degeneracies between cosmological parameters, baryon parameters and $\log_{10}|f_{R0}|$. This leads to a tighter lower bound on $\log_{10}|f_{R0}|$ compared with the case where the same data are fitted by e-Mantis itself. This reinforces our conclusion that the 1D marginalised constraint on $\log_{10}|f_{R0}|$ is sensitive to degeneracies among parameters, and the difference in the theoretical predictions strongly affects the constraint.

The result shown here is just an illustration of the inclusion of theoretical errors and their effects on the parameter constraints. Our implementation is flexible and it can be applied to any theoretical error described by the relative error function $\alpha(k, z)$ defined in Eq. (35).

7. Conclusion

In this paper, we studied the effect of using different nonlinear predictions for the dark matter power spectrum on the parameter constraints in the Hu–Sawikci $f(R)$ gravity model obtained from *Euclid* primary photometric probes. We implemented four different models in the MontePython pipeline to predict angular power spectra for weak lensing (WL), photometric galaxy clustering (GC_{ph}) and their cross-correlation (XC_{ph}). Comparing with the N-body simulation data obtained in paper 1, we found that e-Mantis agreed very well with ECOSMOG that was used to run simulations to construct the emulator, while FORGE had larger errors compared with the MG-AREPO simulation. The agreement is better for one of the N-body simulations used to construct FORGE obtained in paper 2 run by MG-AREPO. This indicates that the difference between e-Mantis and FORGE is larger than the one in the baseline N-body simulations (i.e. ECOSMOG and MG-AREPO) mainly due to the way FORGE was constructed. In the *Euclid* reference cosmology, FORGE gives a larger Ξ, the ratio between the power spectrum in $f(R)$ and in ΛCDM, on all scales compared with e-Mantis and the fitting formula at the 2% level. ReACT underestimates Ξ compared with FORGE more than e-Mantis and the fitting formula at intermediate k.

We used the fiducial value of $|f_{R0}| = 5 \times 10^{-6}$ ($\log_{10}|f_{R0}| = -5.301$) and ran MCMC in the *Euclid* fiducial cosmology defined in EC20. For the fitting formula, the Fisher Matrix forecast and MCMC results generally agree well although the MCMC result is non-Gaussian. This is partly caused by a lack of cosmological parameter dependence of Ξ, which affects the degeneracy between $\log_{10}|f_{R0}|$ and cosmological parameters for large $\log_{10}|f_{R0}|$. e-Mantis gives more Gaussian constraints with smaller errors. When FORGE is used to create the data, the 1D mean of $\log_{10}|f_{R0}|$ is not strongly biased in the case of e-Mantis and the fitting formula and the 1D bias is at most 0.6σ. Even

for the 3×2pt analysis including all the probes and their cross-correlations, the 1D bias is 0.15σ for $\log_{10}|f_{R0}|$ in the case of `e-Mantis`.

The impact of baryonic physics on WL was studied by using a baryonification emulator `BCemu`. For the optimistic setting, the $f(R)$ parameter and two main baryon parameters are well constrained despite the degeneracies among these parameters. However, the difference in the nonlinear dark matter prediction can be compensated by the adjustment of baryon parameters as well as cosmological parameters, and the 1D marginalised constraint on $\log_{10}|f_{R0}|$ is biased. This bias can be avoided in the pessimistic setting at the expense of weaker constraints. For the pessimistic setting, using the ΛCDM synthetic data for WL, we obtained the prior-independent bound of $\log_{10}|f_{R0}| < -5.6$ using `e-Mantis`.

`ReACT` shows a large bias in $\log_{10}|f_{R0}|$ as well as cosmological parameters when `FORGE` was used to create the data. This is because the prediction of `ReACT` is furthest away from `FORGE`. We implemented a method to include uncorrelated theoretical errors proposed in Audren et al. (2013) to address this issue. The method is based on the relative error function for the nonlinear dark matter power spectrum. We estimated this using the difference between `FORGE` and `ReACT`. We found that the inclusion of theoretical errors alone was not enough to mitigate the bias. We then corrected the prediction of `ReACT` with `FORGE` for the fiducial model. We applied this model to the data created by `e-Mantis`. We found that theoretical errors, in this case, helped reduce the bias not only by enlarging errors but also by making the means of 1D marginalised constraint closer to the input values. When we added baryonic effects with `BCemu`, errors were not significantly affected and the input values were still recovered. However, the lower bound on $\log_{10}|f_{R0}|$ is tighter than the one obtained by applying `e-Mantis` itself. This reinforces our conclusion that the 1D marginalised constraint on $\log_{10}|f_{R0}|$ is sensitive to the degeneracies among parameters.

Based on the result of this paper, we draw the following conclusions:

- It is important to check the agreement of different N-body codes that are used to create theoretical predictions. This is not only the code itself, but also accuracy settings. In paper 1, we found that the accuracy setting such as the refinement criteria in the adaptive mesh refinement method has a large effect on the power spectrum. With the controlled accuracy setting, it is possible to realise 1% agreements between different N-body codes in terms of $\Xi(k, z)$ in the Hu–Sawicki $f(R)$ gravity model.
- We then need to check the accuracy of emulators. We found that `FORGE` suffers from larger emulation errors and this leads to a larger difference between `FORGE` and `e-Mantis` compared with the difference in their baseline N-body code `MG-AREPO` and `ECOSMOG` in some cosmologies. Improving the emulation technique will be able to make the agreement better.
- Including baryonic effects can worsen the bias in the 1D marginalised constraint on $\log_{10}|f_{R0}|$. This is because the difference in the nonlinear dark matter power spectrum prediction can be compensated by the adjustment of baryon parameters, and the best-fit values are biased. In addition, the degeneracy between baryon parameters, cosmological parameters and $\log_{10}|f_{R0}|$ leads to a stronger volume effect. This bias can be avoided if there is a prior on baryon parameters from external data sets for example.
- To account for the uncertainty in the theoretical prediction for the nonlinear power spectrum, it is safer to include theoretical errors. In this paper, we used a conservative error estimation using `FORGE` and `ReACT`. An improvement in emulators will make the theoretical error smaller leading to better constraints. However, we still need to check whether or not the inclusion of baryons will change this conclusion.

The pipeline developed in this paper can be used to test the readiness of the nonlinear power spectrum prediction for the application to real data from *Euclid* in other extended cosmologies. For the Data Release 1 of *Euclid* and future data releases, we plan to improve the `FORGE` emulator and make the pipeline ready for the data analysis. We also plan to extend the analysis to the models studied in Frusciante et al. (2023).

Finally, we note that our forecasts have ignored observational systematics such as shear and redshift measurement biases. For Stage-IV surveys like *Euclid*, it is known that these need to be very well characterised in order to get accurate cosmological parameter estimates. This is a challenge that needs to be addressed for both ΛCDM and exotic cosmologies. A detailed study of this issue is beyond the scope of our paper, but we refer to Euclid Collaboration: Mellier et al. (2024) for details. In summary, the mean galaxy redshifts within the bins need to be known with an accuracy better than $\sim 0.002(1+z)$ if errors in cosmological parameters are not to be degraded. On the other hand, Spurio Mancini & Bose (2023) incldued a parameter that shifts the mean of the redshift distribution in each redshift bin and found that these parameters are not strongly degenerate with the $f(R)$ parameter.

Acknowledgements. The Euclid Consortium acknowledges the European Space Agency and a number of agencies and institutes that have supported the development of *Euclid*, in particular the Agenzia Spaziale Italiana, the Austrian Forschungsförderungsgesellschaft funded through BMK, the Belgian Science Policy, the Canadian Euclid Consortium, the Deutsches Zentrum für Luft- und Raumfahrt, the DTU Space and the Niels Bohr Institute in Denmark, the French Centre National d'Etudes Spatiales, the Fundação para a Ciência e a Tecnologia, the Hungarian Academy of Sciences, the Ministerio de Ciencia, Innovación y Universidades, the National Aeronautics and Space Administration, the National Astronomical Observatory of Japan, the Netherlandse Onderzoekschool Voor Astronomie, the Norwegian Space Agency, the Research Council of Finland, 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 (`www.euclid-ec.org`). K. K. is supported by STFC grant ST/W001225/1. B. B. is supported by a UK Research and Innovation Stephen Hawking Fellowship (EP/W005654/2). A. P. is a UKRI Future Leaders Fellow [grant MR/X005399/1]. P. C. is supported by grant RF/ERE/221061. M. C. acknowledges the financial support provided by the Alexander von Humboldt Foundation through the Humboldt Research Fellowship program, as well as support from the Max Planck Society and the Alexander von Humboldt Foundation in the framework of the Max Planck-Humboldt Research Award endowed by the Federal Ministry of Education and Research. F. P. acknowledges partial support from the INFN grant InDark and the Departments of Excellence grant L.232/2016 of the Italian Ministry of University and Research (MUR). FP also acknowledges the FCT project with ref. number PTDC/FIS-AST/0054/2021. GR's research was supported by an appointment to the NASA Postdoctoral Program administered by Oak Ridge Associated Universities under contract with NASA. GR was supported by JPL, which is run under contract by the California Institute of Technology for NASA (80NM0018D0004). Numerical computations were done on the Sciama High Performance Compute (HPC) cluster which is supported by the ICG, SEPNet and the University of Portsmouth. This work has made use of the Infinity Cluster hosted by Institut d'Astrophysique de Paris. We acknowledge open libraries support `IPython` (Perez & Granger 2007), `Matplotlib` (Hunter 2007), `Numpy` (Walt et al. 2011), and `SciPy` (Virtanen et al. 2020). For the purpose of open access, we have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising. Supporting research data are available on reasonable request from the corresponding author.

References

Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, arXiv:astro-ph/0609591

Angulo, R. E., Zennaro, M., Contreras, S., et al. 2021, MNRAS, 507, 5869
Arnold, C., Leo, M., & Li, B. 2019, Nature Astron., 3, 945
Arnold, C. & Li, B. 2019, MNRAS, 490, 2507
Arnold, C., Li, B., Giblin, B., et al. 2022, MNRAS, 515, 4161
Audren, B., Lesgourgues, J., Bird, S., Haehnelt, M. G., & Viel, M. 2013, JCAP, 01, 026
Baldi, M., Villaescusa-Navarro, F., Viel, M., et al. 2014, MNRAS, 440, 75
Bose, B., Carrilho, P., Marinucci, M., et al. 2023, arXiv:2311.13529
Bose, B., Cataneo, M., Tröster, T., et al. 2020, MNRAS, 498, 4650
Bose, B., Tsedrik, M., Kennedy, J., et al. 2023, MNRAS, 519, 4780
Bose, B., Wright, B. S., Cataneo, M., et al. 2021, MNRAS, 508, 2479
Bose, S., Li, B., Barreira, A., et al. 2017, JCAP, 02, 050
Brax, P., Casas, S., Desmond, H., & Elder, B. 2021, Universe, 8, 11
Brax, P., van de Bruck, C., Davis, A.-C., & Shaw, D. J. 2008, Phys. Rev. D, 78, 104021
Buchdahl, H. A. 1970, MNRAS, 150, 1
Burrage, C., March, B., & Naik, A. P. 2024, JCAP, 04, 004
Carrilho, P., Carrion, K., Bose, B., et al. 2022, MNRAS, 512, 3691
Carrion, K., Carrilho, P., Spurio Mancini, A., Pourtsidou, A., & Hidalgo, J. C. 2024, arXiv:2402.18562
Casas, S., Cardone, V. F., Sapone, D., et al. 2023, arXiv:2306.11053
Casas, S., Lesgourgues, J., Schöneberg, N., et al. 2024, A&A, 682, A90
Cataneo, M., Lombriser, L., Heymans, C., et al. 2019, MNRAS, 488, 2121
Cautun, M., Paillas, E., Cai, Y.-C., et al. 2018, MNRAS, 476, 3195
Clifton, T., Ferreira, P. G., Padilla, A., & Skordis, C. 2012, Phys. Rep., 513, 1
Desmond, H. & Ferreira, P. G. 2020, Phys. Rev. D, 102, 104060
Elbers, W., Frenk, C. S., Jenskins, A., et al. 2024, arXiv:2403.12967
Euclid Collaboration: Adamek et al. 2024, in prep
Euclid Collaboration: Blanchard, A., Camera, S., Carbone, C., et al. 2020, A&A, 642, A191
Euclid Collaboration: Knabenhans, M. et al. 2021, MNRAS, 505, 2840
Euclid Collaboration: Mellier, Y., Acevedo Barroso, J., Achúcarro, A., et al. 2024, A&A, accepted, arXiv:2405.13491
Euclid Collaboration: Racz et al. 2024, in prep
Euclid Collaboration: Sciotti, D., Gouyou Beauchamps, S., Cardone, V. F., et al. 2023, arXiv:2310.15731
Ferreira, P. G. 2019, Ann. Rev. Astron. Astrophys., 57, 335
Fiorini, B., Koyama, K., & Baker, T. 2023, JCAP, 12, 045
Frusciante, N. Pace, F., Cardone, V. F., et al. 2023, arXiv:2306.12368
Gelman, A. & Rubin, D. B. 1992, Statist. Sci., 7, 457
Giocoli, C., Baldi, M., & Moscardini, L. 2018, MNRAS, 481, 2813
Giri, S. K. & Schneider, A. 2021, JCAP, 12, 046
Gordon, C. & Trotta, R. 2007, MNRAS, 382, 1859
Grandis, S., Arico', G., Schneider, A., & Linke, L. 2024, MNRAS, 528, 4379
Guillet, T. & Teyssier, R. 2011, Journal of Computational Physics, 230, 4756
Harnois-Déraps, J., Hernandez-Aguayo, C., Cuesta-Lazaro, C., et al. 2023, MNRAS, 525, 6336
Holm, E. B., Nygaard, A., Dakin, J., et al. 2023, arXiv:2312.02972
Hu, W. & Sawicki, I. 2007, Phys. Rev., D76, 064004
Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90
Ishak, M. 2019, Living Rev. Rel., 22, 1
Joyce, A., Jain, B., Khoury, J., & Trodden, M. 2015, Phys. Rept., 568, 1
Khoury, J. & Weltman, A. 2004, Physical Review Letters, 93
Kou, R., Murray, C., & Bartlett, J. G. 2023, arXiv:2311.09936
Koyama, K. 2016, Rept. Prog. Phys., 79, 046902
Lewis, A. 2019, arXiv e-prints, arXiv:1910.13970
Li, B., Zhao, G.-B., Teyssier, R., & Koyama, K. 2012, JCAP, 01, 051
McKay, M. D., Beckman, R. J., & Conover, W. J. 1979, Technometrics, 21, 239
Mead, A., Brieden, S., Tröster, T., & Heymans, C. 2021, MNRAS, 502, 1401
Mead, A., Heymans, C., Lombriser, L., et al. 2016, MNRAS, 459, 1468
Mead, A., Peacock, J., Heymans, C., et al. 2015, MNRAS, 454, 1958
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
Perez, F. & Granger, B. E. 2007, Computing in Science Engineering, 9, 21
Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565
Piga, L., Marinucci, M., D'Amico, G., et al. 2023, JCAP, 04, 038
Pisanti, O., Mangano, G., Miele, G., & Mazzella, P. 2021, JCAP, 04, 020
Planck Collaboration: Aghanim, N., Akrami, Y., Ashdown, M., et al. 2020, A&A, 641, A6, [Erratum: Astron.Astrophys. 652, C4 (2021)]
Puchwein, E., Baldi, M., & Springel, V. 2013, MNRAS, 436, 348
Ramachandra, N., Valogiannis, G., Ishak, M., & Heitmann, K. 2021, Phys. Rev. D, 103, 123525
Rasmussen, C. E. & Williams, C. K. I. 2005, Gaussian Processes for Machine Learning (The MIT Press)
Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009
Schneider, A., Refregier, A., Grandis, S., et al. 2020, JCAP, 04, 020
Sotiriou, T. P. & Faraoni, V. 2010, Reviews of Modern Physics, 82, 451
Springel, V. 2010, MNRAS, 401, 791
Spurio Mancini, A. & Bose, B. 2023, arXiv:2305.06350
Spurio Mancini, A., Köhlinger, F., Joachimi, B., et al. 2019, MNRAS, 490, 2155
Spurio Mancini, A., Piras, D., Alsing, J., Joachimi, B., & Hobson, M. P. 2022, MNRAS, 511, 1771
Sáez-Casares, I., Rasera, Y., & Li, B. 2023, MNRAS, 527, 7242
Takahashi, R., Sato, M., Nishimichi, T., et al. 2012, ApJ, 761, 152
Teyssier, R. 2002, A&A, 385, 337
Tröster, T., Asgari, M., Blake, C., et al. 2021, A&A, 649, A88
Tsedrik, M., Bose, B., Carrilho, P., et al. 2024, arXiv:2404.11508
Valogiannis, G. & Bean, R. 2017, Phys. Rev. D, 95, 103515
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261
Walt, S. v. d., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science and Engg., 13, 22
Weinberger, R., Springel, V., & Pakmor, R. 2020, ApJS, 248, 32
Winther, H., Casas, S., Baldi, M., et al. 2019, Phys. Rev. D, 100, 123540
Winther, H. A., Koyama, K., Manera, M., Wright, B. S., & Zhao, G.-B. 2017, JCAP, 08, 006
Winther, H. A., Schmidt, F., Barreira, A., et al. 2015, MNRAS, 454, 4208

[1] Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK
[2] Institute for Theoretical Particle Physics and Cosmology (TTK), RWTH Aachen University, 52056 Aachen, Germany
[3] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
[4] Laboratoire Univers et Théorie, Observatoire de Paris, Université PSL, Université Paris Cité, CNRS, 92190 Meudon, France
[5] Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal
[6] Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Edifício C8, Campo Grande, PT1749-016 Lisboa, Portugal
[7] Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute (AIRUB), German Centre for Cosmological Lensing (GCCL), 44780 Bochum, Germany
[8] Universität Bonn, Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany
[9] INAF-Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, 40129 Bologna, Italy
[10] Istituto Nazionale di Fisica Nucleare, Sezione di Bologna, Via Irnerio 46, 40126 Bologna, Italy
[11] Dipartimento di Fisica, Università degli Studi di Torino, Via P. Giuria 1, 10125 Torino, Italy
[12] INFN-Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy
[13] INAF-Osservatorio Astrofisico di Torino, Via Osservatorio 20, 10025 Pino Torinese (TO), Italy
[14] Higgs Centre for Theoretical Physics, School of Physics and Astronomy, The University of Edinburgh, Edinburgh EH9 3FD, UK
[15] Institut universitaire de France (IUF), 1 rue Descartes, 75231 PARIS CEDEX 05, France
[16] Institut für Theoretische Physik, University of Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany
[17] Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, CNRS, UPS, CNES, 14 Av. Edouard Belin, 31400 Toulouse, France
[18] Université St Joseph; Faculty of Sciences, Beirut, Lebanon
[19] Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, 0315 Oslo, Norway
[20] Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK
[21] Department of Astrophysics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
[22] Dipartimento di Fisica e Astronomia, Università di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy
[23] INFN-Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
[24] Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain
[25] Institut de Ciencies de l'Espai (IEEC-CSIC), Campus UAB, Carrer de Can Magrans, s/n Cerdanyola del Vallés, 08193 Barcelona, Spain

[26] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, 91109, USA
[27] Institut de Physique Théorique, CEA, CNRS, Université Paris-Saclay 91191 Gif-sur-Yvette Cedex, France
[28] School of Mathematics and Physics, University of Surrey, Guildford, Surrey, GU2 7XH, UK
[29] INAF-Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano, Italy
[30] IFPU, Institute for Fundamental Physics of the Universe, via Beirut 2, 34151 Trieste, Italy
[31] INAF-Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, 34143 Trieste, Italy
[32] INFN, Sezione di Trieste, Via Valerio 2, 34127 Trieste TS, Italy
[33] SISSA, International School for Advanced Studies, Via Bonomea 265, 34136 Trieste TS, Italy
[34] Institut d'Astrophysique de Paris, UMR 7095, CNRS, and Sorbonne Université, 98 bis boulevard Arago, 75014 Paris, France
[35] Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 1, 85748 Garching, Germany
[36] Dipartimento di Fisica, Università di Genova, Via Dodecaneso 33, 16146, Genova, Italy
[37] INFN-Sezione di Genova, Via Dodecaneso 33, 16146, Genova, Italy
[38] Department of Physics "E. Pancini", University Federico II, Via Cinthia 6, 80126, Napoli, Italy
[39] INAF-Osservatorio Astronomico di Capodimonte, Via Moiariello 16, 80131 Napoli, Italy
[40] INFN section of Naples, Via Cinthia 6, 80126, Napoli, Italy
[41] Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, PT4150-762 Porto, Portugal
[42] Faculdade de Ciências da Universidade do Porto, Rua do Campo de Alegre, 4150-007 Porto, Portugal
[43] Aix-Marseille Université, CNRS, CNES, LAM, Marseille, France
[44] INAF-IASF Milano, Via Alfonso Corti 12, 20133 Milano, Italy
[45] Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Avenida Complutense 40, 28040 Madrid, Spain
[46] Port d'Informació Científica, Campus UAB, C. Albareda s/n, 08193 Bellaterra (Barcelona), Spain
[47] INAF-Osservatorio Astronomico di Roma, Via Frascati 33, 00078 Monteporzio Catone, Italy
[48] Dipartimento di Fisica e Astronomia "Augusto Righi" - Alma Mater Studiorum Università di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
[49] Instituto de Astrofísica de Canarias, Calle Vía Láctea s/n, 38204, San Cristóbal de La Laguna, Tenerife, Spain
[50] European Space Agency/ESRIN, Largo Galileo Galilei 1, 00044 Frascati, Roma, Italy
[51] ESAC/ESA, Camino Bajo del Castillo, s/n., Urb. Villafranca del Castillo, 28692 Villanueva de la Cañada, Madrid, Spain
[52] Université Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, UMR 5822, Villeurbanne, F-69100, France
[53] Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
[54] Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, 08028 Barcelona, Spain
[55] Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig de Lluís Companys 23, 08010 Barcelona, Spain
[56] UCB Lyon 1, CNRS/IN2P3, IUF, IP2I Lyon, 4 rue Enrico Fermi, 69622 Villeurbanne, France
[57] Department of Astronomy, University of Geneva, ch. d'Ecogia 16, 1290 Versoix, Switzerland
[58] Université Paris-Saclay, CNRS, Institut d'astrophysique spatiale, 91405, Orsay, France
[59] INFN-Padova, Via Marzolo 8, 35131 Padova, Italy
[60] INAF-Istituto di Astrofisica e Planetologia Spaziali, via del Fosso del Cavaliere, 100, 00100 Roma, Italy
[61] Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191, Gif-sur-Yvette, France
[62] Institut d'Estudis Espacials de Catalunya (IEEC), Edifici RDIT, Campus UPC, 08860 Castelldefels, Barcelona, Spain
[63] FRACTAL S.L.N.E., calle Tulipán 2, Portal 13 1A, 28231, Las Rozas de Madrid, Spain
[64] INAF-Osservatorio Astronomico di Padova, Via dell'Osservatorio 5, 35122 Padova, Italy
[65] Universitäts-Sternwarte München, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstrasse 1, 81679 München, Germany
[66] Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano, Via Celoria 16, 20133 Milano, Italy
[67] Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK
[68] Felix Hormuth Engineering, Goethestr. 17, 69181 Leimen, Germany
[69] Technical University of Denmark, Elektrovej 327, 2800 Kgs. Lyngby, Denmark
[70] Cosmic Dawn Center (DAWN), Denmark
[71] Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France
[72] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
[73] NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[74] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
[75] Department of Physics and Helsinki Institute of Physics, Gustaf Hällströmin katu 2, 00014 University of Helsinki, Finland
[76] Aix-Marseille Université, CNRS/IN2P3, CPPM, Marseille, France
[77] Université de Genève, Département de Physique Théorique and Centre for Astroparticle Physics, 24 quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland
[78] Department of Physics, P.O. Box 64, 00014 University of Helsinki, Finland
[79] Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, Helsinki, Finland
[80] NOVA optical infrared instrumentation group at ASTRON, Oude Hoogeveensedijk 4, 7991PD, Dwingeloo, The Netherlands
[81] Centre de Calcul de l'IN2P3/CNRS, 21 avenue Pierre de Coubertin 69627 Villeurbanne Cedex, France
[82] INFN-Sezione di Milano, Via Celoria 16, 20133 Milano, Italy
[83] INFN-Sezione di Roma, Piazzale Aldo Moro, 2 - c/o Dipartimento di Fisica, Edificio G. Marconi, 00185 Roma, Italy
[84] Dipartimento di Fisica e Astronomia "Augusto Righi" - Alma Mater Studiorum Università di Bologna, via Piero Gobetti 93/2, 40129 Bologna, Italy
[85] Department of Physics, Institute for Computational Cosmology, Durham University, South Road, DH1 3LE, UK
[86] Université Paris Cité, CNRS, Astroparticule et Cosmologie, 75013 Paris, France
[87] University of Applied Sciences and Arts of Northwestern Switzerland, School of Engineering, 5210 Windisch, Switzerland
[88] Institut d'Astrophysique de Paris, 98bis Boulevard Arago, 75014, Paris, France
[89] Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona), Spain
[90] European Space Agency/ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
[91] DARK, Niels Bohr Institute, University of Copenhagen, Jagtvej 155, 2200 Copenhagen, Denmark
[92] Waterloo Centre for Astrophysics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
[93] Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
[94] Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada
[95] Space Science Data Center, Italian Space Agency, via del Politecnico snc, 00133 Roma, Italy
[96] Centre National d'Etudes Spatiales – Centre spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France

[97] Institute of Space Science, Str. Atomistilor, nr. 409 Măgurele, Ilfov, 077125, Romania
[98] Dipartimento di Fisica e Astronomia "G. Galilei", Università di Padova, Via Marzolo 8, 35131 Padova, Italy
[99] Departamento de Física, FCFM, Universidad de Chile, Blanco Encalada 2008, Santiago, Chile
[100] Universität Innsbruck, Institut für Astro- und Teilchenphysik, Technikerstr. 25/8, 6020 Innsbruck, Austria
[101] Satlantis, University Science Park, Sede Bld 48940, Leioa-Bilbao, Spain
[102] Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Tapada da Ajuda, 1349-018 Lisboa, Portugal
[103] Universidad Politécnica de Cartagena, Departamento de Electrónica y Tecnología de Computadoras, Plaza del Hospital 1, 30202 Cartagena, Spain
[104] INFN-Bologna, Via Irnerio 46, 40126 Bologna, Italy
[105] Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands
[106] Dipartimento di Fisica, Università degli studi di Genova, and INFN-Sezione di Genova, via Dodecaneso 33, 16146, Genova, Italy
[107] Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA
[108] INAF, Istituto di Radioastronomia, Via Piero Gobetti 101, 40129 Bologna, Italy
[109] Astronomical Observatory of the Autonomous Region of the Aosta Valley (OAVdA), Loc. Lignan 39, I-11020, Nus (Aosta Valley), Italy
[110] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[111] School of Physics and Astronomy, Cardiff University, The Parade, Cardiff, CF24 3AA, UK
[112] Junia, EPA department, 41 Bd Vauban, 59800 Lille, France
[113] ICSC - Centro Nazionale di Ricerca in High Performance Computing, Big Data e Quantum Computing, Via Magnanelli 2, Bologna, Italy
[114] Instituto de Física Teórica UAM-CSIC, Campus de Cantoblanco, 28049 Madrid, Spain
[115] CERCA/ISO, Department of Physics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA
[116] Departamento de Física Fundamental. Universidad de Salamanca. Plaza de la Merced s/n. 37008 Salamanca, Spain
[117] Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, Via Giuseppe Saragat 1, 44122 Ferrara, Italy
[118] Istituto Nazionale di Fisica Nucleare, Sezione di Ferrara, Via Giuseppe Saragat 1, 44122 Ferrara, Italy
[119] Center for Data-Driven Discovery, Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan
[120] Ludwig-Maximilians-University, Schellingstrasse 4, 80799 Munich, Germany
[121] Max-Planck-Institut für Physik, Boltzmannstr. 8, 85748 Garching, Germany
[122] Minnesota Institute for Astrophysics, University of Minnesota, 116 Church St SE, Minneapolis, MN 55455, USA
[123] Institute Lorentz, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands
[124] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice cedex 4, France
[125] Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA
[126] Department of Physics & Astronomy, University of California Irvine, Irvine CA 92697, USA
[127] Department of Astronomy & Physics and Institute for Computational Astrophysics, Saint Mary's University, 923 Robie Street, Halifax, Nova Scotia, B3H 3C3, Canada
[128] Departamento Física Aplicada, Universidad Politécnica de Cartagena, Campus Muralla del Mar, 30202 Cartagena, Murcia, Spain
[129] Instituto de Astrofísica de Canarias (IAC); Departamento de Astrofísica, Universidad de La Laguna (ULL), 38200, La Laguna, Tenerife, Spain
[130] Department of Physics, Oxford University, Keble Road, Oxford OX1 3RH, UK
[131] CEA Saclay, DFR/IRFU, Service d'Astrophysique, Bat. 709, 91191 Gif-sur-Yvette, France
[132] Department of Computer Science, Aalto University, PO Box 15400, Espoo, FI-00 076, Finland
[133] Instituto de Astrofísica de Canarias, c/ Via Lactea s/n, La Laguna E-38200, Spain. Departamento de Astrofísica de la Universidad de La Laguna, Avda. Francisco Sanchez, La Laguna, E-38200, Spain
[134] Univ. Grenoble Alpes, CNRS, Grenoble INP, LPSC-IN2P3, 53, Avenue des Martyrs, 38000, Grenoble, France
[135] Department of Physics and Astronomy, Vesilinnantie 5, 20014 University of Turku, Finland
[136] Serco for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
[137] ARC Centre of Excellence for Dark Matter Particle Physics, Melbourne, Australia
[138] Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia
[139] School of Physics and Astronomy, Queen Mary University of London, Mile End Road, London E1 4NS, UK
[140] Department of Physics and Astronomy, University of the Western Cape, Bellville, Cape Town, 7535, South Africa
[141] ICTP South American Institute for Fundamental Research, Instituto de Física Teórica, Universidade Estadual Paulista, São Paulo, Brazil
[142] IRFU, CEA, Université Paris-Saclay 91191 Gif-sur-Yvette Cedex, France
[143] Oskar Klein Centre for Cosmoparticle Physics, Department of Physics, Stockholm University, Stockholm, SE-106 91, Sweden
[144] Astrophysics Group, Blackett Laboratory, Imperial College London, London SW7 2AZ, UK
[145] INAF-Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125, Firenze, Italy
[146] Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 2, 00185 Roma, Italy
[147] Aurora Technology for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
[148] Centro de Astrofísica da Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal
[149] HE Space for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
[150] Dipartimento di Fisica - Sezione di Astronomia, Università di Trieste, Via Tiepolo 11, 34131 Trieste, Italy
[151] Theoretical astrophysics, Department of Physics and Astronomy, Uppsala University, Box 515, 751 20 Uppsala, Sweden
[152] Department of Physics, Royal Holloway, University of London, TW20 0EX, UK
[153] Department of Astrophysical Sciences, Peyton Hall, Princeton University, Princeton, NJ 08544, USA
[154] Cosmic Dawn Center (DAWN)
[155] Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 Copenhagen, Denmark
[156] Center for Cosmology and Particle Physics, Department of Physics, New York University, New York, NY 10003, USA
[157] Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, 10010, New York, NY, USA

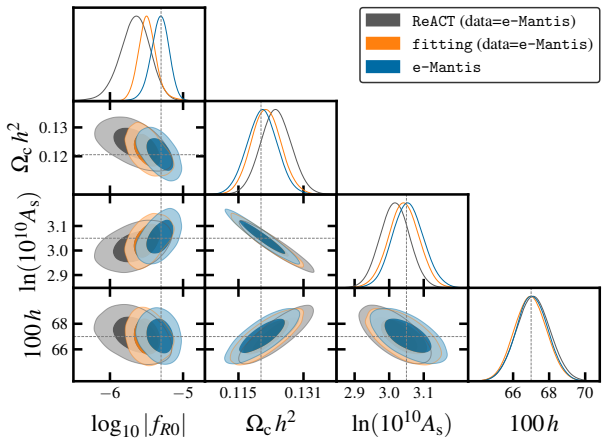

Fig. A.1. Bias due to different nonlinear modelling. The synthetic data are created by `e-Mantis` and fitted by two different models in WL pessimistic case without baryons.

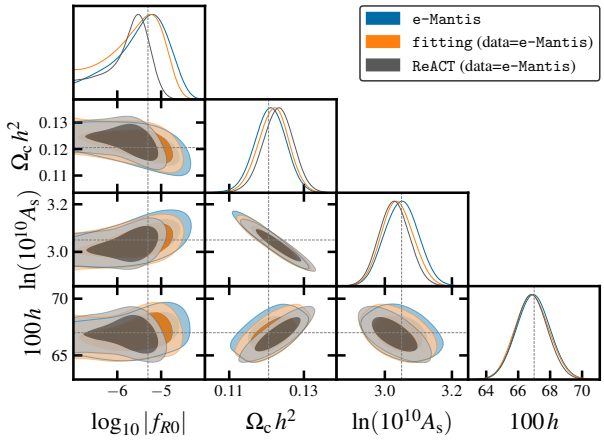

Fig. A.2. Bias due to different nonlinear modelling. The synthetic data are created by `e-Mantis` and fitted by two different models in WL pessimistic case with baryons.

Table A.1. The mean, standard deviation and 68.3% lower and upper limit of $\log_{10}|f_{R0}|$ for the WL pessimistic setting.

WL pessimistic					
	Mean	S.d.	Lower	Upper	B_{1D}
Data = `e-Mantis`					
fitting	−5.482	0.135	−5.615	−5.376	1.464
ReACT	−5.663	0.216	−5.851	−5.440	3.080

Appendix A: `e-Mantis` fiducial data

In this Appendix, we show the result of the cases where we use `e-Mantis` as fiducial data for the pessimistic setting of WL with and without baryons. Figure A.1 and Table A.1 show the results without baryons while the Figure A.2 and Table A.2 show the results with baryons. We observe that the agreements of the three models are generally better as expected although without baryons, 1D bias is still at the 3σ level for `ReACT` due to smaller errors of `e-Mantis`. On the other hand, in the case with baryons, 1D bias is reduced to less than 1σ.

Table A.2. The mean, standard deviation and 68.3% lower and upper limit of $\log_{10}|f_{R0}|$ for the WL pessimistic setting with baryons.

WL pessimistic with baryons					
	Mean	S.d.	Lower	Upper	B_{1D}
Data = `e-Mantis`					
fitting	−5.626	0.637	−6.085	−4.785	0.337
ReACT	−5.781	0.495	−6.053	−5.147	0.700

Appendix B: Construction of the relative error

To construct the relative error function, $\alpha(k,z)$, we can use the fact that $\Xi(k,z)$ is constructed as the ratio to the same ΛCDM nonlinear power spectrum calculated with the Halofit 'Takahashi' prescription in all models. This means that any deviations of the ΛCDM power spectrum from N-body simulations is also absorbed in Ξ. We can thus write the relative error as

$$\alpha(k,z) = \frac{\Delta P_{\delta\delta}(k,z)}{P_{\delta\delta}(k,z)} = \frac{\Delta\Xi(k,z)}{\Xi(k,z)}\,. \tag{B.1}$$

To construct $\Delta P_{\delta\delta}$ we follow the basic idea that all predictions for Ξ are equally accurate in calculating the modified-gravity power spectrum. The true power spectrum is thus only known up to the spread of the predictions. To be conservative with our forecast, we choose to calculate the error from the difference between the `FORGE` predictions and the `ReACT` predictions, which is the largest among these predictions. This is shown in Fig. B.1. The `ReACT` prescription underpredicts the power spectrum at intermediate scales and overpredicts slightly at smaller scales. This is most likely due to `ReACT` handling the one- and two-halo power spectra separately. `HMcode2020` predictions for the pseudo-cosmology non-linear power spectrum also contribute to these inaccuracies as we can see by comparing the left and right panels of Figure 4 in Cataneo et al. (2019). To simplify the exact difference between these two codes, we argue that the theoretical error should plateau at the maximum deviation. This is because at nonlinear scales, the modes of different wave numbers are no longer independent of each other. By plateauing, we essentially say that we can no longer be more precise in the computation of the power spectrum on smaller scales.

Finally, we construct α using only the difference of the predictions at a fiducial of $|f_{R0}| = 5\times10^{-6}$ in order not to bias our results. Our final fit has the form

$$\alpha(k,z) = A(z)\,\frac{x^2+x}{x^2+x+1}\,, \tag{B.2}$$

$$A(z) = \frac{A_1}{\exp\left(\frac{z-A_2}{A_3}\right)+1} + A_4\,, \tag{B.3}$$

$$k_p(z) = B_1\,\exp\left[\tanh\left(\frac{z-B_2}{B_3}\right)\right]\,, \tag{B.4}$$

where we use $x := k/k_p$. We separately fit the amplitude function $A(z)$ to the maximum deviation and the plateau wavenumber $k_p(z)$ to the wavenumber of maximum deviation. The best-fit values can be found in Table B.1 and are shown as the shaded region in Fig. B.1. We see that the zero line is within the 68.3% confidence bounds most of the time. This is in accordance with our construction that the two emulators should differ by the theoretical error.

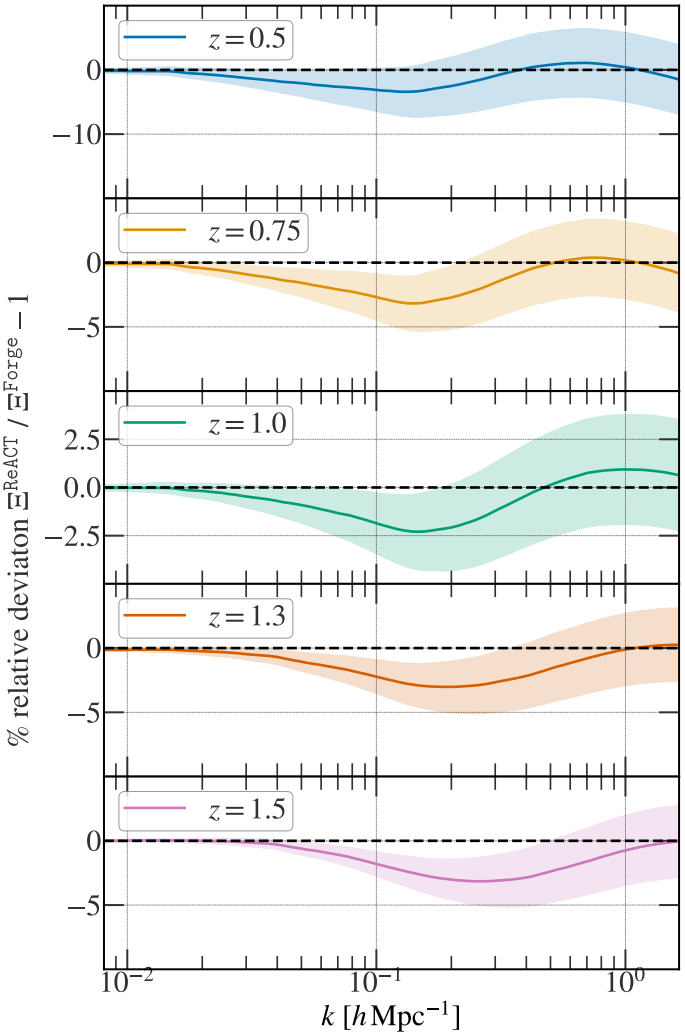

Fig. B.1. Relative difference between the power spectra from FORGE and ReACT. We show the relative deviation as the lines and the 68.3% confidence level as the shaded area. The boosts are all computed at the fiducial cosmology.

Table B.1. Best fit values of the relative error function.

Parameter	Best fit
A_1	3.56×10^{-2}
A_2	0.562
A_3	6.1×10^{-2}
A_4	2.90×10^{-2}
B_1	0.202
B_2	1.56
B_3	0.55