# Euclid preparation. XLI. Galaxy power spectrum modelling in real space

A Pezzotta, C Moretti, M Zennaro, A Moradinezhad Dizgah, M Crocce, E Sefusatti, I Ferrero, K Pardede, A Eggemeier, A Barreira, et al.

A&A, 687, A216 (2024)
https://doi.org/10.1051/0004-6361/202348939


Astronomy & Astrophysics

# *Euclid* preparation

## XLI. Galaxy power spectrum modelling in real space

Euclid Collaboration: A. Pezzotta[1], C. Moretti[2,3,4,5,6], M. Zennaro[7], A. Moradinezhad Dizgah[8], M. Crocce[9,10], E. Sefusatti[5,6,11], I. Ferrero[12], K. Pardede[2,13,6,11,5], A. Eggemeier[14], A. Barreira[15,16], R. E. Angulo[17,18], M. Marinucci[19,20], B. Camacho Quevedo[21,9], S. de la Torre[22], D. Alkhanishvili[14], M. Biagetti[23,6,2,5], M.-A. Breton[9,10,24], E. Castorina[25,26], G. D'Amico[19,27], V. Desjacques[20], M. Guidi[28,29], M. Kärcher[22,30], A. Oddo[2,6], M. Pellejero Ibanez[17,4], C. Porciani[14], A. Pugno[14], J. Salvalaggio[5,11,31,6], E. Sarpa[2,11], A. Veropalumbo[32,33], Z. Vlah[34,35,36], A. Amara[37], S. Andreon[32], N. Auricchio[38], M. Baldi[39,38,40], S. Bardelli[38], R. Bender[1,41], C. Bodendorf[1], D. Bonino[42], E. Branchini[43,33,32], M. Brescia[44,45,46], J. Brinchmann[47], S. Camera[48,49,42], V. Capobianco[42], C. Carbone[50], V. F. Cardone[51,52], J. Carretero[53,54], S. Casas[55], F. J. Castander[9,21], M. Castellano[51], S. Cavuoti[45,46], A. Cimatti[56], G. Congedo[4], C. J. Conselice[57], L. Conversi[58,59], Y. Copin[60], L. Corcione[42], F. Courbin[61], H. M. Courtois[62], A. Da Silva[63,64], H. Degaudenzi[65], A. M. Di Giorgio[66], J. Dinis[64,63], X. Dupac[59], S. Dusini[67], A. Ealet[60], M. Farina[66], S. Farrens[68], P. Fosalba[21,10], M. Frailis[5], E. Franceschi[38], S. Galeotta[5], B. Gillis[4], C. Giocoli[38,69], B. R. Granett[32], A. Grazian[70], F. Grupp[1,71], L. Guzzo[25,32,26], S. V. H. Haugan[12], F. Hormuth[72], A. Hornstrup[73,74], K. Jahnke[75], B. Joachimi[76], E. Keihänen[77], S. Kermiche[78], A. Kiessling[79], M. Kilbinger[80], T. Kitching[81], B. Kubik[60], M. Kunz[8], H. Kurki-Suonio[82,83], S. Ligori[42], P. B. Lilje[12], V. Lindholm[82,83], I. Lloro[84], E. Maiorano[38], O. Mansutti[5], O. Marggraf[14], K. Markovic[79], N. Martinet[22], F. Marulli[85,38,40], R. Massey[86], E. Medinaceli[38], Y. Mellier[87,88], M. Meneghetti[38,40], E. Merlin[51], G. Meylan[61], M. Moresco[85,38], L. Moscardini[85,38,40], E. Munari[5], S.-M. Niemi[89], C. Padilla[53], S. Paltani[65], F. Pasian[5], K. Pedersen[90], W. J. Percival[91,92,93], V. Pettorino[94], S. Pires[68], G. Polenta[95], J. E. Pollack[96,97], M. Poncet[98], L. A. Popa[99], L. Pozzetti[38], F. Raison[1], A. Renzi[100,67], J. Rhodes[79], G. Riccio[45], E. Romelli[5], M. Roncarelli[38], E. Rossetti[39], R. Saglia[41,1], D. Sapone[101], B. Sartoris[41,5], P. Schneider[14], T. Schrabback[102], A. Secroun[78], G. Seidel[75], M. Seiffert[79], S. Serrano[21,9,103], C. Sirignano[100,67], G. Sirri[40], L. Stanco[67], C. Surace[22], P. Tallada-Crespí[104,54], A. N. Taylor[4], I. Tereno[63,105], R. Toledo-Moreo[106], F. Torradeflot[54,104], I. Tutusaus[107], E. A. Valentijn[108], L. Valenziano[38,109], T. Vassallo[41,5], Y. Wang[110], J. Weller[41,1], G. Zamorani[38], J. Zoubian[78], E. Zucca[38], A. Biviano[5,6], E. Bozzo[65], C. Burigana[111,109], C. Colodro-Conde[112], D. Di Ferdinando[40], G. Mainetti[113], M. Martinelli[51,52], N. Mauri[56,40], Z. Sakr[114,107,115], V. Scottez[87,116], M. Tenti[40], M. Viel[6,5,2,11,3], M. Wiesmann[12], Y. Akrami[117,118], V. Allevato[45], S. Anselmi[67,100,24], C. Baccigalupi[2,5,11,6], M. Ballardini[119,120,38], F. Bernardeau[121,88], A. Blanchard[107], S. Borgani[31,6,5,11], S. Bruton[122], R. Cabanac[107], A. Cappi[38,123], C. S. Carvalho[105], G. Castignani[85,38], T. Castro[5,11,6,3], G. Cañas-Herrera[89,124], K. C. Chambers[125], S. Contarini[85,40,38], A. R. Cooray[126], J. Coupon[65], S. Davini[33], G. De Lucia[5], G. Desprez[127], S. Di Domizio[43,33], H. Dole[128], A. Díaz-Sánchez[129], J. A. Escartin Vigo[1], S. Escoffier[78], P. G. Ferreira[7], F. Finelli[38,109], L. Gabarra[7], K. Ganga[97], J. García-Bellido[117], F. Giacomini[40], G. Gozaliasl[130,82], A. Hall[4], S. Ilić[131,98,107], S. Joudaki[132,91,92], J. J. E. Kajava[133,134], V. Kansal[135,136,137], C. C. Kirkpatrick[77], L. Legrand[8], A. Loureiro[138,139], J. Macias-Perez[140], M. Magliocchetti[66], F. Mannucci[141], R. Maoli[142,51], C. J. A. P. Martins[143,47], S. Matthew[4], L. Maurin[128], R. B. Metcalf[85,38], M. Migliaccio[144,145], P. Monaco[31,5,11,6], G. Morgante[38], S. Nadathur[132], Nicholas A. Walton[146], L. Patrizii[40], V. Popa[99], D. Potter[147], A. Pourtsidou[4,148], M. Pöntinen[82], I. Risso[149], P.-F. Rocci[128], M. Sahlén[150], A. G. Sánchez[1], A. Schneider[147], M. Sereno[38,40], P. Simon[14], A. Spurio Mancini[81], J. Steinwagner[1], G. Testera[33], R. Teyssier[151], S. Toft[74,152], S. Tosi[43,33,32], A. Troja[100,67], M. Tucci[65], J. Valiviita[82,83], D. Vergani[38], G. Verza[153,154], and P. Vielzeuf[78]

*(Affiliations can be found after the references)*

Received 13 December 2023 / Accepted 15 March 2024

**ABSTRACT**

We investigate the accuracy of the perturbative galaxy bias expansion in view of the forthcoming analysis of the *Euclid* spectroscopic galaxy samples. We compare the performance of a Eulerian galaxy bias expansion using state-of-the-art prescriptions from the effective field theory of large-scale structure (EFTofLSS) with a hybrid approach based on Lagrangian perturbation theory and high-resolution simulations. These models are benchmarked against comoving snapshots of the flagship I $N$-body simulation at $z = (0.9, 1.2, 1.5, 1.8)$, which have been populated with H$\alpha$ galaxies leading to catalogues of millions of objects within a volume of about $58\,h^{-3}\,\mathrm{Gpc}^3$. Our analysis suggests that both models can be used to provide a robust inference of the parameters $(h, \omega_\mathrm{c})$ in the redshift range under consideration, with comparable constraining power. We additionally determine the range of validity of the EFTofLSS model in terms of scale cuts and model degrees of freedom. From these tests, it emerges that the standard third-order Eulerian bias expansion – which includes local and non-local bias parameters, a matter counter term, and a correction to the shot-noise contribution – can accurately describe the full shape of the real-space galaxy power spectrum up to the maximum wavenumber of $k_\mathrm{max} = 0.45\,h\,\mathrm{Mpc}^{-1}$, and with a measurement precision of well below the percentage level. Fixing either of the tidal bias parameters to physically motivated relations still leads to unbiased cosmological constraints, and helps in reducing the severity of projection effects due to the large dimensionality of the model. We finally show how we repeated our analysis assuming a volume that matches the expected footprint of *Euclid*, but without considering observational effects, such as purity and completeness, showing that we can get constraints on the combination $(h, \omega_\mathrm{c})$ that are consistent with the fiducial values to better than the 68% confidence interval over this range of scales and redshifts.

**Key words.** cosmological parameters – cosmology: theory – large-scale structure of Universe

## 1. Introduction

The large-scale distribution of galaxies is an extremely important source of cosmological information from the low-redshift Universe, complementing observations of the cosmic microwave background (CMB), such as those made using the *Wilkinson* Microwave Anisotropy Probe (WMAP; Hinshaw et al. 2013) and *Planck* (Planck Collaboration VI 2020). In the course of the past two decades, observations of the large-scale structure (LSS) from spectroscopic galaxy redshift surveys, such as the 2dF Galaxy Redshift Survey (2dFGRS; Colless et al. 2001), the 6dF Galaxy Survey (6dFGS; Jones et al. 2009), the VIMOS VLT Deep Survey (VVDS; Le Fèvre et al. 2013), the Sloan Digital Sky Survey (SDSS; York et al. 2000), the WiggleZ Dark Energy Survey (WiggleZ; Drinkwater et al. 2010), the VIMOS Public Extragalactic Redshift Survey (VIPERS; Guzzo et al. 2014), the Galaxy And Mass Assembly (GAMA; Driver et al. 2011), and the Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2012) and its extension (eBOSS; Dawson et al. 2016), have provided a wealth of information on how gravitational instability shapes the large-scale matter distribution and on the relation between matter and galaxy density perturbations. At the same time, such observations have stood as a testing ground for what has ultimately emerged as the standard cosmological model.

In the next decade, this picture is going to be significantly enriched by analyses performed by Stage-IV spectroscopic surveys, such as the Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration 2016) and *Euclid* (Laureijs et al. 2011), which are going to explore a still relatively uncharted epoch at $1 \lesssim z \lesssim 2$, when the Universe was only about half of its current age. In particular, *Euclid* is going to collect the redshift of millions of H$\alpha$-emitting galaxies across a total sky surface of $15\,000\,\mathrm{deg}^2$, therefore increasing the statistical constraining power on the cosmological parameters to an unprecedented level for spectroscopic analyses in the low-redshift Universe. It comes with no surprise that the increase in statistical significance of the observations must necessarily be accompanied by an equivalent increase in the accuracy of the theoretical recipes used to analyse the data in order to keep systematic errors in the theory at a fraction of the statistical error budget. This becomes even more relevant in terms of the range of validity of the considered models, whose reach must be properly benchmarked against realistic mock samples.

The standard cosmological probe for galaxy clustering is the galaxy two-point correlation function (2PCF), or its Fourier transform, the galaxy power spectrum. Both statistics quantify the excess probability of finding galaxy pairs at a given separation with respect to the case of a purely random (Poissonian) distribution. These observables have been extensively used by recent experiments to place constraints on the cosmological parameters, either focusing on specific features such as baryon acoustic oscillation (BAO) and redshift-space distortions (RSDs) in the so-called template-fitting approach (Peacock et al. 2001; Tegmark et al. 2006; Guzzo et al. 2008; Blake et al. 2011; Reid et al. 2012; Beutler et al. 2012, 2016; Contreras et al. 2013; Howlett et al. 2015; Okumura et al. 2016; Alam et al. 2017; Pezzotta et al. 2017; Gil-Marín et al. 2018; Hou et al. 2018; Wang et al. 2018; Zhao et al. 2018), or in full-shape analyses (Sánchez et al. 2014, 2016; Grieb et al. 2017; Ivanov et al. 2020; Tröster et al. 2020; d'Amico et al. 2020; Semenaite et al. 2022, 2023; Chen et al. 2022; Philcox & Ivanov 2022; Carrilho et al. 2023; Moretti et al. 2023).

For both kinds of approaches, the inference of cosmological information is made intrinsically more difficult by the presence of three separate effects that build on linear theory predictions. These are the non-linear gravitational evolution of the dark matter density field (Bernardeau et al. 2002; Baumann et al. 2012; Carrasco et al. 2012), the relationship between the galaxy $\delta_\mathrm{g}$ and the matter $\delta$ density fields, known as thr 'galaxy bias' (Kaiser 1984; Bardeen et al. 1986; see Desjacques et al. 2018 for a recent review), and finally the apparently anisotropic pattern in the distribution of galaxies due to the effect of their peculiar velocities on the observed redshift (RSDs, Kaiser 1987; Hamilton 1992; Fisher 1995; Scoccimarro et al. 1999b; Scoccimarro 2004; Taruya et al. 2010; Senatore & Zaldarriaga 2014; Perko et al. 2016). Each of these effects needs to be carefully modelled in order to infer accurate cosmological constraints from the full shape of the galaxy power spectrum/2PCF. This goal can be achieved in different ways: using numerical methods, such as $N$-body simulations (e.g. Kuhlen et al. 2012; Schneider et al. 2016; Springel et al. 2021), adopting analytical approaches based on a perturbative solution to the equations governing the evolution of the matter and galaxy density fields (e.g. Fry & Gaztanaga 1993; Bernardeau et al. 2002; McDonald & Roy 2009; Carrasco et al. 2012; Assassi et al. 2014; Senatore 2015; Desjacques et al. 2018), or resorting to hybrid methods that combine the previous two

methodologies (e.g. Euclid Collaboration 2021; Angulo et al. 2021; Zennaro et al. 2023; Pellejero-Ibañez et al. 2023).

In terms of galaxy bias, it has become standard practice to adopt a perturbative expansion of the galaxy density field using Eulerian coordinates, which can be expressed as a sum of partial derivatives of the gravitational and velocity divergence potentials, each one weighted by a corresponding free parameter to be fitted against measurements. When considering the one-loop galaxy power spectrum, the list includes the linear bias, $b_1$, expressed in terms of the dark matter density field in the large-scale limit, as $\delta_g = b_1\delta$ (Kaiser 1984; Bardeen et al. 1986; Cole & Kaiser 1989; Nusser & Davis 1994; Mo & White 1996; Sheth & Tormen 1999), plus the next-to-leading-order correction obtained from a spherically symmetric gravitational collapse via a power-law Taylor expansion $\delta_g = \sum_n b_n \delta^n/n!$ of the matter density field (Szalay 1988; Coles 1993; Fry & Gaztanaga 1993; Scoccimarro et al. 2001; Smith et al. 2007; Manera et al. 2010; Desjacques et al. 2010; Frusciante & Sheth 2012; Schmidt et al. 2013). This is further supplemented by the presence of non-local contributions generated by the cosmic tidal field (Bouchet et al. 1992; Catelan et al. 1998, 2000; McDonald & Roy 2009), which have been proven to be essential for a correct description of the clustering of dark matter halos (Manera & Gaztañaga 2011; Roth & Porciani 2011) and to secure consistency between the results from the analysis of two- and three-point correlation measurements (Pollack et al. 2012, 2014). These extra corrections were first detected in $N$-body simulations in Chan et al. (2012) and Baldauf & Seljak (2012), and have since then become a standard ingredient in the bias expansion. Additionally, the latter also takes into account the effects of short-range non-localities during the processes of galaxy formation, which lead to the presence of higher-than-second-order derivatives of the gravitational potential. At leading order in the power spectrum, higher derivatives appear with a term scaling as $\nabla^2\delta$ (Bardeen et al. 1986; Matsubara 1999; Desjacques 2008; Desjacques et al. 2010). Finally, the dependence on short-wavelength modes is included via an additional stochastic field $\varepsilon_g(\boldsymbol{x})$ (Dekel & Lahav 1999; Sheth & Lemson 1999; Taruya & Soda 1999; Matsubara 1999; Bonoli & Pen 2009; Hamaus et al. 2010; Schmidt 2016; Ginzburg et al. 2017)[1], which is responsible for a shot-noise contribution to the power spectrum. This correction deviates from the predictions of a purely Poissonian distribution, and at the same time can introduce a scale dependence due to the physical scale at which two objects can be mistaken for a single one, similarly to the exclusion effect for dark matter halos (Scherrer & Weinberg 1998; Sheth & Lemson 1999; Cooray & Sheth 2002; Smith et al. 2007; Baldauf et al. 2013, 2016).

The high dimensionality of the parameter space for the model described above can be reduced by employing a set of physically motivated relations expressing a few higher-order bias parameters in terms of lower-order ones. A typical assumption is the conserved evolution of tracers (coevolution), which, from a local-in-matter-density expansion at the moment of formation, leads to the well-known local Lagrangian relations (Chan et al. 2012; Baldauf & Seljak 2012; Eggemeier et al. 2019). The latter have been adopted in the literature as a fairly conservative trade-off between sampling the whole set of bias parameters and fixing some of the model degrees of freedom, most notably in the analysis of the BOSS DR12 data release (Sánchez et al. 2016; Grieb et al. 2017) to improve the statistical constraints on the cosmological parameters obtained from the anisotropic 2PCF and power spectrum.

The standard bias expansion has been the subject of several tests in the literature, together with a validation of the coevolution relations mentioned in the previous paragraph. As an example, Saito et al. (2014) checked the consistency between the bias parameters fitted from the halo power spectrum and bispectrum (the Fourier transform of the three-point function) using a sample of measurements in different mass bins and at different redshifts, revealing an agreement between the two sets of bias measurements up to $k \sim 0.1\,h\,\mathrm{Mpc}^{-1}$. The use of an irreducible bias basis, and also properly including a higher-derivative correction, was tested in Angulo et al. (2015), who showed that with this approach it is possible to extend the validity of the one-loop galaxy bias expansion up to $k \sim 0.3\,h\,\mathrm{Mpc}^{-1}$ even at $z = 0$. More recently, Eggemeier et al. (2020) analysed the accuracy of this expansion at fixed cosmology using simulated halo occupation distribution (HOD) catalogues built to mimic the clustering properties of the SDSS Main Galaxy Sample (Strauss et al. 2002) and of the BOSS CMASS and LOWZ samples (Eisenstein et al. 2011; Dawson et al. 2012; Reid et al. 2016). The authors focused on the necessity to introduce both a higher-derivative term and a scale-dependent correction to shot noise while analysing the auto galaxy and cross galaxy-matter power spectrum. Findings from this study indicate that the standard one-loop bias expansion can be broken on scales $k \sim 0.2\,h\,\mathrm{Mpc}^{-1}$ unless higher-order stochastic corrections are taken into consideration. Pezzotta et al. (2021) and Eggemeier et al. (2021) extended this analysis to include a determination of the cosmological parameters, and explored the additional constraining power coming from the one-loop galaxy bispectrum. These works show how fixing the quadratic tidal bias as a function of the linear bias provides accurate results up to $k \sim 0.35\,h\,\mathrm{Mpc}^{-1}$ for the galaxy power spectrum. A similar analysis was carried out by Oddo et al. (2021), who assessed the constraining power of the galaxy bispectrum on the cosmological parameters, displaying a consistency up to $k \sim 0.3\,h\,\mathrm{Mpc}^{-1}$ for the one-loop power spectrum and $k \sim 0.09\,h\,\mathrm{Mpc}^{-1}$ for the tree-level bispectrum. Equivalent analyses in redshift space (Markovič et al. 2019; Bose et al. 2020; de la Bella et al. 2020; Gualdi et al. 2021; Rizzo et al. 2023; Nicola et al. 2024), or in terms of field-level comparisons (Schmittfull et al. 2019), have also appeared in recent years, leading to compatible scenarios.

On a partially different side, numerical simulations (see e.g. Kuhlen et al. 2012, for a review) have proven to be an optimal way to reproduce the evolution of the matter and galaxy density field deep into the non-linear regime, and their use in analyses of galaxy redshift surveys has therefore multiplied in recent years thanks to a large number of different suites, such as DEMNUni (Castorina et al. 2015), UNIT (Chuang et al. 2019), Quijote (Villaescusa-Navarro et al. 2020), Uchuu (Ishiyama et al. 2021), and AbacusSummit (Maksimova et al. 2021; Yuan et al. 2022) simulations. In quantitative terms, state-of-art $N$-body simulations can achieve an accuracy of better than 2% on the shape of the non-linear matter power spectrum down to scales of $k \sim 10\,h\,\mathrm{Mpc}^{-1}$ (Schneider et al. 2016; Springel et al. 2021; Angulo et al. 2021), but unfortunately their application as a tool to infer cosmological information from real observations is limited by their extreme computational cost. However, in recent years, different methods have been proposed with the goal of increasing their range of applicability. This ranges from methods

[1] In order for $\varepsilon_g(\boldsymbol{x})$ to be completely uncorrelated from large-scale fluctuations, the hypothesis of primordial Gaussianity must hold true. On the contrary, $\varepsilon_g(\boldsymbol{x})$ cannot be treated as a purely stochastic contribution.

meant to speed up their production (e.g. Monaco et al. 2002; Tassev et al. 2013; Izard et al. 2016) to ones designed to find an optimal interpolation strategy among a limited pool of high-resolution simulations (Heitmann et al. 2013; Liu et al. 2018; Nishimichi et al. 2019; DeRose et al. 2019; Giblin et al. 2019; Wibking et al. 2018; Winther et al. 2019; Euclid Collaboration 2019, 2021). Among this second category of methods, we highlight `baccoemu` (Angulo et al. 2021), an emulator for the non-linear matter power spectrum that was recently extended to also include biased tracers in real (Zennaro et al. 2023) and redshift space (Pellejero-Ibañez et al. 2023), assuming a hybrid Lagrangian bias model, with the individual terms of the expansion directly emulated from high-resolution simulations, and a cosmology-rescaling technique meant to reduce the number of simulations required to train the emulator.

In this paper, we compare different models and test different scale cuts and bias relations on a sample of synthetic galaxy catalogues tailored to reproduce – to the best of our knowledge – the clustering signal of the H$\alpha$ sample that will be targeted by *Euclid*. As we are interested in the relative performance of different theory models, we do not consider the presence of observational systematic uncertainties in this analysis. For example, effects such as purity and completeness of the sample will induce variations in the comoving number density considered in this work. At the same time, the sample purity, which is determined by the presence of line and noise interlopers, will also modify the overall shape of the measured $n$-point statistics. All of these effects are going to be investigated by a dedicated group in the Euclid Consortium, while a specific analysis on theory model selection with a more realistic analysis (including survey mask, selection effects, and combining multiple redshift bins) is going to be developed in a future paper (Euclid Collaboration: Moretti et al., in prep.).

Our goal is to test the range of validity of the one-loop galaxy bias expansion, which we quantify by means of three different performance metrics (Osato et al. 2019): the figure of bias, quantifying the accuracy of the model in terms of the recovery of the model parameters; the goodness of fit, measuring how well the best-fit model compares to the input data vectors; and the figure of merit, quantifying the statistical power of the model in constraining the cosmological parameters. These metrics are computed for each of the configurations we test, as a function of the maximum wave mode $k_{\rm max}$ included in the fit, exploring different bias relations meant to reduce the dimensionality of the parameter space. As we limit our attention to the real-space galaxy power spectrum alone, we focus on the recovery of the dimensionless Hubble parameter $h$, which is defined in terms of the Hubble constant $H_0$ as $H_0 = 100\,h\,{\rm km\,s^{-1}\,Mpc^{-1}}$, and of the cold dark matter density parameter $\omega_{\rm c} \equiv \Omega_{\rm c}h^2$, where $\Omega_{\rm c}$ is the corresponding fractional density parameter. At the same time, we avoid sampling the scalar amplitude $A_{\rm s}$, as this would lead to a strong degeneracy with the linear bias parameter $b_1$. This degeneracy can be partially broken if only considering the additional constraining power from higher-order statistics, cross-correlation with the matter density field, or in a multi-tracer analysis, or by considering the apparently anisotropic clustering amplitude when also including RSDs.

This work is the first installment in a series of *Euclid* preparation papers meant to validate the theoretical framework used to analyse the full shape of two- and three-point clustering measurements from the final data sample. Here, we focus on the real-space galaxy power spectrum, while the corresponding three-point equivalent for the real-space galaxy bispectrum is going to be presented in Euclid Collaboration: Eggemeier et al. (in prep.). Both of these papers describe tests conducted in real space, that is, using the true comoving positions of galaxies inside the box instead of the positions displaced because of RSDs. While this choice excludes one of the main observational probes of galaxy clustering, such analyses can provide an important testing ground for the model of galaxy bias. This includes calibration of optimal scale cuts for the model[2], as well as testing different ways to reduce the dimensionality of the parameter space, such as using the coevolution relations. In addition, real-space analyses can become relevant in the context of modelling $3\times2$ point statistics (photometric galaxy clustering, weak gravitational lensing, and galaxy–galaxy lensing), such as in the analysis performed by the Dark Energy Survey (see e.g. Pandey et al. 2022; Porredon et al. 2022, for cosmological inference that requires a proper modelling of photometric galaxy bias). On the other hand, future installments of this series will focus on the modelling of the redshift-space equivalents of the statistics adopted in these works.

This article is structured in the following way. In Sect. 2 we present the simulated galaxy samples and the power-spectrum measurements and covariances used throughout the paper. In Sect. 3 we describe the theoretical models that we employed for the analysis, while in Sect. 4 we describe the fitting procedure and the performance metrics used to quantify the goodness of the models for different configurations as a function of the maximum mode included in the fit. Finally, in Sect. 5 we present the results of the analysis, and we draw our conclusions in Sect. 6.

## 2. Data

### 2.1. Euclid simulations

In order to determine the performance of the selected theoretical models, we first need a set of simulated data samples spanning the same redshift range that will be observed by *Euclid*, and for which the input cosmology is known a priori.

In the following we make use of four comoving outputs, selected to cover the redshift range $0.9 < z < 1.8$ of the Flagship I simulation[3]. The latter has been carried out on the supercomputer Piz Daint, which is hosted by the Swiss National Supercomputing Center (CSCS), using the `PKDGRAV3` algorithm (Potter et al. 2017), and consists of a record-setting $N$-body run with two trillions dark matter particles moving under the effect of gravity within a box of size $L = 3780\,h^{-1}\,{\rm Mpc}$. The mass resolution of the simulation ($m_{\rm p} \sim 2.398\times10^9\,h^{-1}\,M_\odot$) allows us to marginally resolve halos with a typical mass $M_{\rm h}$ of few $10^{10}\,h^{-1}\,M_\odot$, which host the majority of the H$\alpha$ emission line galaxies that are going to be targeted by *Euclid*. The nominal flat ΛCDM cosmology adopted to run the simulation as stated in Potter et al. (2017) differs from the fiducial cosmology assumed in this paper in the value of the spectral index ($n_{\rm s} = 0.96$ vs. $n_{\rm s} = 0.97$). This choice has been motivated since,

[2] The reader should bear in mind that these scale cuts are determined only from the performance of galaxy bias. More realistic scale cuts, also including the effect of, for example, RSDs will be provided in one of the next entries of the series (Euclid Collaboration: Camacho et al., in prep.).

[3] The roman numeral 'I' is meant to differentiate the simulation adopted in this work from its more recent version, i.e. Flagship II. The latter has been upgraded with respect to its predecessor in a number of way, such as by displaying a 2.5 times larger mass resolution, accounting for relativistic effects, and including massive neutrinos. However, because of the unavailability of halo comoving snapshots at the time the analysis presented in this paper first started, we decided to employ the older version of the Flagship for this work.

**Table 1.** Fiducial parameters of the flat ΛCDM cosmological model of the Flagship I simulation.

| $h$ | $\omega_c$ | $\omega_b$ | $M_\nu$ [eV] | $\sigma_8$ | $n_s$ |
|---|---|---|---|---|---|
| 0.67 | 0.121203 | 0.0219961 | 0 | 0.83 | 0.97 |

**Notes.** From left to right, the list includes the value of the Hubble parameter $h$, the density parameter of cold dark matter $\omega_c$, and baryons $\omega_b$, the total neutrino mass $M_\nu$, the rms density fluctuations inside a sphere of radius $8\,h^{-1}$ Mpc, $\sigma_8$, and the scalar index $n_s$.

during the course of this study, we observed subtle yet significant inconsistencies between our models and the measurement in the Flagship I simulation. After contacting the team responsible for running the simulation, they confirmed that the nominal parameters of the simulation were previously wrongly communicated, and that they are in agreement with the ones we have identified. The latter are obtained by performing cosmological fits to high-resolution dark matter power spectrum measurements at various redshifts, and are summarized in Table 1. This procedure is detailed in Appendix A.

Each comoving snapshot has been populated with galaxies by firstly generating a friends-of-friends (FoF) halo catalogue with a linking length $b = 0.2$ and a minimum halo mass corresponding to ten dark matter particles[4], where the halo mass $M_h$ is defined as the sum of all the identified particles. Subsequently, halos have been populated with galaxies using a HOD algorithm to match the abundance and clustering of the H$\alpha$ samples implemented in the main Flagship I lightcone catalog[5]. The latter, in turn, is meant to reproduce the number density and clustering properties corresponding to two different H$\alpha$ profiles, labelled as Model 1 and Model 3 in Pozzetti et al. (2016). These samples are defined by different templates for the evolution of the luminosity function $\phi(L, z)$, from the use of a standard Schechter parametrization for Model 1 (Schechter 1976), to the direct fit to real observations for Model 3. The net effect in terms of number density is that the Model 1 sample has almost twice as many objects as Model 3, which is more conservative in the selection of H$\alpha$ emitters, as shown in Fig. 4 of Pozzetti et al. (2016)[6]. A more detailed description of the Flagship H$\alpha$ lightcone and the pipeline for its construction will be provided in Euclid Collaboration: Castander et al., (in prep).

In terms of the comoving snapshots, the HOD we implemented consists of a 8-parameter model, where the mean occupation numbers of central and satellite galaxies are defined as

$$\langle N_{\rm cen}\rangle(M_h) = \frac{1}{2} f_{\rm cen}^{\rm max}\left[1 + {\rm erf}\left(\frac{\log M_h - \log M_{\rm min}}{\sigma_{\log M}}\right)\right] \times\left[1 - \frac{1 - f_{\rm cen}^{\rm min}/f_{\rm cen}^{\rm max}}{1 + 10^{\frac{2}{k}(\log M_h - \log M_{\rm drop})}}\right], \quad (1)$$

[4] Despite this small number, we only select halos hosting H$\alpha$ emitters with a minimum mass corresponding to few tens of matter particles, based on the redshift of the considered snapshot.

[5] The particle lightcone has been built on the fly when running the $N$-body code at the CSCS, and features a full-sky distribution of dark matter particles in the redshift range $0 < z < 2.3$. This has been later populated with `Rockstar` halos (Behroozi et al. 2013) and galaxies using HOD and abundance matching techniques to reproduce the expected number density and luminosity profile of the H$\alpha$ models described in Pozzetti et al. (2016).

[6] We do not consider the additional Model 2 in this analysis since the total number density of this sample is of the same order of the one of the Model 1 sample, which already provides an optimistic number count.

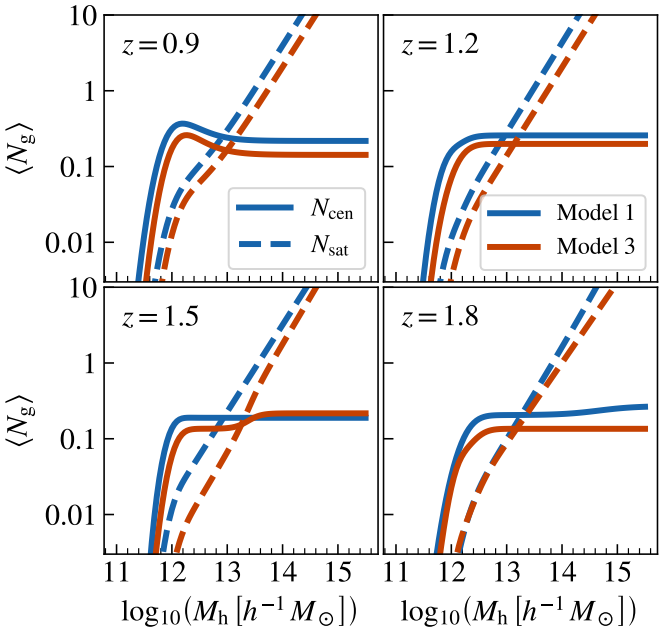


**Fig. 1.** Halo occupation dsitribution profiles of the eight H$\alpha$ samples employed in this analysis. Individual panels show the profiles at different redshift (increasing from the top left to the bottom right panel) for both Model 1 (blue) and 3 (orange). Solid and dashed lines identify the average number of central and satellite galaxies, respectively, in halos of mass $M_h$.

$$\langle N_{\rm sat}\rangle(M_h) = N_{\rm cen}\left(\frac{M_h}{M_1}\right)^{\alpha}. \quad (2)$$

Here, $M_{\rm min}$ is the typical minimum mass of halos hosting a central galaxy, $\sigma_{\log M}$ is the dispersion around $M_{\rm min}$, and $f_{\rm cen}^{\rm max}$ is the amplitude of the central galaxy occupation. We include mass-dependency of $\langle N_{\rm cen}\rangle$ above the transition scale $M_{\rm min}$ using three additional parameters, $M_{\rm drop}$, $k$, and $f_{\rm cen}^{\rm min}$. Finally, the mean occupancy of satellite galaxies is regulated by $M_1$, which is a simple normalisation factor, and $\alpha$, which corresponds to the slope of the power law distribution of satellites. In order to determine the distribution of galaxies inside halos, we employ a standard NFW profile (Navarro et al. 1996).

In Fig. 1 we show the fitted HOD profiles as a function of the host halo mass $M_h$ for the eight samples: four redshifts times two different models. In all panels, both centrals and satellites profiles are shown, marked respectively with continuous and dashed lines. The mean occupation of central galaxies $f_{\rm cen}^{\rm max}$ inside dark matter halos does not converge to one, even for the most massive halos selected by the halo finder. This is a consequence of having selected a subsample (H$\alpha$ in this case) from the whole population present in the ligthcone. The typical expection value for the occupancy of central galaxies for $M_h > M_{\rm min}$ slightly varies with the different samples, but is typically close to 0.2. This follows from the property of H$\alpha$ emitters to be relatively young, blue, and star-forming galaxies, whereas, in massive halos, environmental effects such as galaxy-galaxy interactions, ram pressure stripping, and AGN feedback can suppress star formation in galaxies, effectively reducing the likelihood of finding actively star-forming central galaxies.

The total number of galaxies for each sample, their number density, and the scale $k_{\rm sn}$ at which the Poissonian shot-noise $P_{\rm sn} \equiv 1/\bar{n}$ becomes the dominant contribution in the data vectors, are listed in Table 2. A warning to be made is that, to determine the parameters for the HOD models, we selected galaxies

**Table 2.** Specifications for the HOD galaxy samples used in this analysis.

| $z$ | HOD | $N_g$ | $\bar{n}$ $\left[h^3\,\mathrm{Mpc}^{-3}\right]$ | $k_{sn}$ $\left[h\,\mathrm{Mpc}^{-1}\right]$ | $\eta$ |
|---|---|---|---|---|---|
| 0.9 | 1 | 201 816 513 | 0.0037 | 0.64 | 6.67 |
| | 3 | 110 321 755 | 0.0020 | 0.51 | |
| 1.2 | 1 | 108 057 141 | 0.0020 | 0.56 | 5.88 |
| | 3 | 55 563 490 | 0.0010 | 0.39 | |
| 1.5 | 1 | 69 132 138 | 0.0013 | 0.45 | 5.26 |
| | 3 | 31 613 213 | 0.0006 | 0.26 | |
| 1.8 | 1 | 24 553 758 | 0.0005 | 0.26 | 3.33 |
| | 3 | 16 926 864 | 0.0003 | 0.22 | |

**Notes.** The table lists the total number of objects $N_g$, the mean comoving number density $\bar{n}$ of the sample, and the scale $k_{sn}$ at which Poissonian shot- noise becomes the leading contribution in the galaxy power spectrum. Following our convention on the normalization of the power spectrum, the latter is simply defined as the inverse of the mean number density. All the considered samples share the same volume, which coincides with the one of the Flagship I comoving outputs, i.e. $(3780\,h^{-1}\,\mathrm{Mpc})^3$. The last columns shows the volume factor $\eta$ between the full-box volume and the one of a *Euclid*-like shell, as defined in Eq. (4).

from the lightcone assuming a significantly faint H$\alpha$ flux limit, corresponding to $f_{H\alpha} = 2\times10^{-16}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$ (Scaramella et al. 2022), without assuming any realistic observational effect, such as target incompleteness, purity of the sample, and the impact of the angular footprint and radial selection function (Euclid Collaboration: Granett et al., in prep., Euclid Collaboration: Monaco et al., in prep.). This results in a sample with higher number density (see e.g. Bagley et al. 2020, for a recent forecast of H$\alpha$ emitters from HST), with measured galaxy power spectra that are less affected by shot-noise. At the same time, the lack of line and noise interlopers allows us to neglect any extra contribution (Euclid Collaboration: Risso et al., in prep., Euclid Collaboration: Lee et al., in prep.) to the model galaxy power spectrum presented in Sect. 3. As a consequence, given the high precision assumed to validate the theory models, we believe that our tests should provide a conservative estimate of their range of validity. We leave to future *Euclid* analyses a more dedicated study of the impact of observational systematics.

### 2.2. Measurements and covariances

For each of the samples described above, we measured the real-space galaxy power spectrum $P_{gg}(k)$ using the publicly available `PowerI4` code[7]. The latter provides the functionality to compute the power spectrum from a particle distribution within a regular cubic box, using a variety of particle assignment schemes. For this analysis, we made use of a fourth-order interpolation scheme, otherwise known as piecewise cubic spline (PCS; see Sefusatti et al. 2016, for the exact form of the kernel), coupled with an interlacing method to reduce the aliasing contribution at high wave modes $k$.

We measured the power spectrum in the range defined by $\left[k_F, k_{Nyq}\right]$, where $k_F = 2\pi/L \sim 0.0017\,h\,\mathrm{Mpc}^{-1}$ is the fundamental frequency in a box of linear size $L$, and $k_{Nyq} = \pi N_{grid}/L$ is the Nyquist frequency corresponding to a density grid of linear size $N_{grid}$. We choose a grid resolution of $N_{grid} = 1024$ for the three dimensions of the box, to obtain measurements of the power spectrum up to a maximum wave mode of $0.8\,h\,\mathrm{Mpc}^{-1}$, and we sampled the $k$ range using a linear binning with step $\Delta k = k_F$[8]. The top and central panels of Fig. 2 show the power spectrum measurements for the Model 1 and 3 HOD samples respectively, with the redshift evolution over the available simulation snapshots marked by different lines in each panel. Differently from the evolution of the matter power spectrum, the galaxy power spectrum features an increasingly lower amplitude at lower redshifts. This can be explained by a larger linear galaxy bias at high redshift that overcomes the growth of matter fluctuations.

Since only a single realization is available for each of the HOD samples, we estimate the error covariance matrices associated to the data vectors using an analytical prediction in the Gaussian approximation, as explained in Grieb et al. (2016). This implies that the variance $\sigma^2_P(k)$ associated to each $k$ bin is independent from the value of the galaxy power spectrum at different modes, and can be written as

$$\sigma^2_P(k) = \frac{2}{N_k} P^2_{gg}(k), \tag{3}$$

where $N_k$ identifies the number of independent wave modes falling in the bin $[k-\Delta k/2, k+\Delta k/2]$, while $P_{gg}(k)$ is the theoretical non-linear galaxy power spectrum including shot-noise contributions. The latter has been obtained from a preliminary fit of the full non-linear model to the data vector of each sample, assuming, in a first iteration, an approximate but reasonable evaluation of the covariance itself. We expect the Gaussian approximation to be sufficient to our goals. The only additional contribution due to the galaxy trispectrum, here neglected, while noticeable at the relevant scales (Scoccimarro et al. 1999a; Sefusatti et al. 2006; Blot et al. 2015, 2016; Bertolini et al. 2016; Wadekar & Scoccimarro 2020) does not lead to significant differences ($\lesssim$10%) on parameters constraints in the mildly non-linear regime (Blot et al. 2019; Wadekar et al. 2020). This is also supported by the goal of this analysis, which is testing the relative performance of different theory models rather than providing absolute values for the parameter uncertainties. However, we highlight how the Gaussian approximation is not expected to deliver completely realistic error bars (to the level of accuracy mentioned above), and that future analyses, both on simulated and real data will be integrated with a more complex model also including non-linear corrections (either analytical or from $N$-body simulations).

The bottom panel of Fig. 2 shows the standard deviation normalised by the corresponding galaxy power spectrum, where the latter is shot-noise-subtracted to highlight the different level of noise in our samples. For this reason, the eight cases exhibit a deviation from the linear relation in Eq. (3) (shown with a black dashed line) at small ($k \gtrsim 0.1\,h\,\mathrm{Mpc}^{-1}$) scales, where the shot-noise correction starts to become dominant with respect to the power spectrum signal. It should be noted how the relative error over this range of scales is well below the 1% level. Similarly,

[7] Available at https://github.com/sefusatti/PowerI4. For this purpose, we do not use the official *Euclid* code, LE3-PK-GC, since our only need is to measure the power spectrum from periodic boxes, without including also radial and angular selection effects that can be properly included using the official code.

[8] For a limited number of the configurations presented in later sections, we carried out consistency checks with a different linear binning, namely $\Delta k = 2k_F$ and $\Delta k = 3k_F$, showing how the final constraints do not depend significantly on this choice.

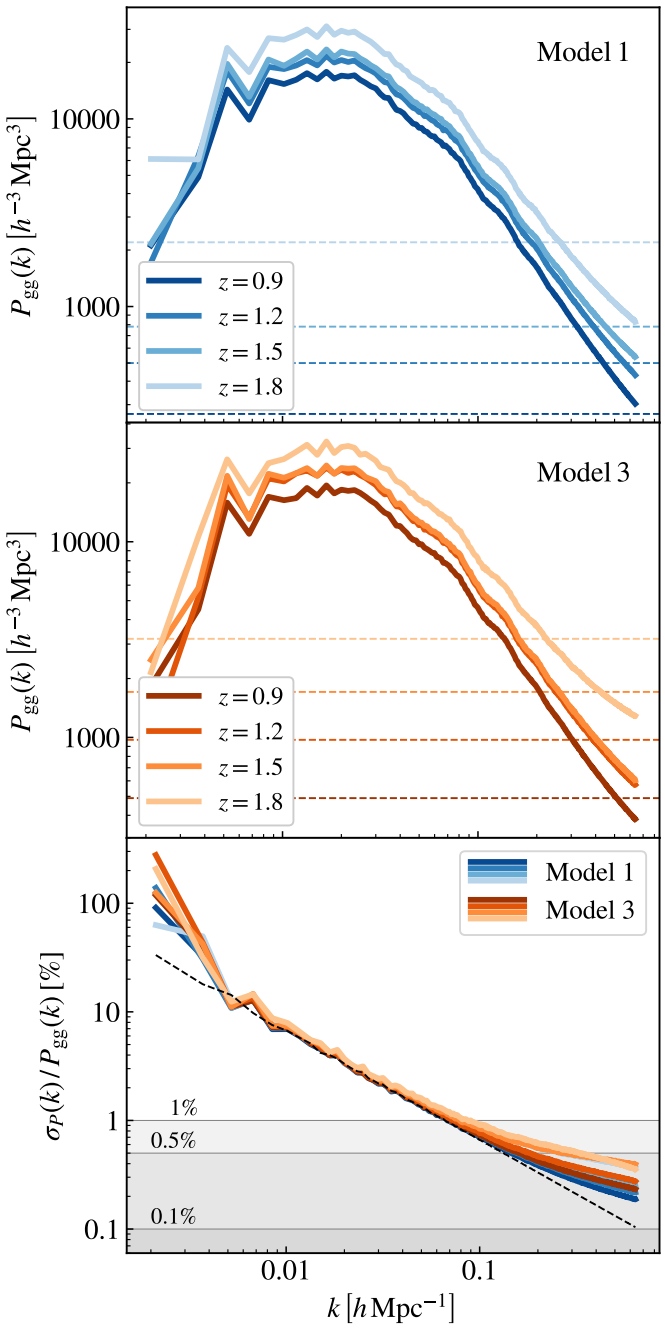


**Fig. 2.** Galaxy power spectrum measurements and uncertainties obtained from the Flagship I comoving snapshots. Top: measurements of the Model 1 HOD samples. The colour gradient identifies the different redshifts of the samples, as shown in the legend. Dashed horizontal lines correspond to the amplitude of the Poisson shot-noise term $P^{\rm sn}$ – obtained as the inverse of the number density specified in the last column of Table 2 – for the different redshifts. Centre: same but for the Model 3 HOD samples. Bottom: error-to-measurement ratios, assuming a Gaussian covariance matrix as in Eq. (3). The coloured solid lines are obtained using the Poisson noise-subtracted power spectra, while the dashed black line highlights the linear relationship from Eq. (3), i.e. $2/N_k$. Grey bands mark the 1%, 0.5%, and 0.1% limit.

at large scales ($k \lesssim 0.005\,h\,\mathrm{Mpc}^{-1}$) we find a departure from the linear relationship due to the small amplitude of $P_{gg}$, which can be clearly observed from the top and middle panel of Fig. 2. Finally, we note that this range of scales is also partially dominated by cosmic variance, due to the use of a single realisation of the Flagship simulation.

### 2.3. Volume rescaling

The main goal of our analysis is to carry out stringent tests to determine the range of validity of the standard one-loop galaxy bias model on the redshift range that will be explored by *Euclid*. For this we make use of a volume $V_{\rm box}$ corresponding to the full-box size of the Flagship comoving outputs, which is significantly larger than the volume that will be covered by *Euclid*. At the same time, we are interested in assessing the constraining power of the real-space galaxy power spectrum using a reference volume close to the one of an expected redshift bin of the full *Euclid* volume, $V_{\rm shell}$. With this purpose in mind, we define new covariance matrices for the different samples presented in the previous sections, with an overall amplitude rescaled by the ratio between the volume of the Flagship comoving outputs and that of the *Euclid*-like shells,

$$\eta = \frac{V_{\rm box}}{V_{\rm shell}}, \tag{4}$$

such that the rescaled covariance $C_{\rm shell}$ can be expressed[9] in terms of the original full-box covariance $C_{\rm box}$ as

$$C_{\rm shell} = \eta\, C_{\rm box}. \tag{5}$$

We follow Euclid Collaboration (2020) and assume four non-overlapping redshift shells, centered at $z = (0.9, 1.2, 1.5, 1.8)$, and with a depth of $\Delta z = (0.2, 0.2, 0.2, 0.3)$, respectively, over a total projected area of 15 000 square degrees. With these values, we derive volume factors $\eta$ for each of the considered redshift bins, shown in the last column of Table 2. We note that the mean values of the four redshift shells used in Euclid Collaboration (2020) do not match perfectly the redshifts of the four comoving snapshots used in this work. However, this is only marginally relevant, since we do not carry out a proper comparison to the Fisher forecasts obtained in that analysis. In fact, this will be a more suited aspect of investigation when considering the same observables, that is, the Legendre multipoles of the anisotropic galaxy power spectrum, and especially when considering more realistic number densities, as pointed out in Sect. 2.1.

A proper comparison between the results obtained using the full-box volume and the rescaled ones is presented in Sect. 5.5. In addition to the *Euclid*-like shells, we consider three additional volume rescalings, by dividing the range between $V_{\rm box}$ and $V_{\rm shell}$ into four evenly sized intervals. This leads to a total of five different sets of covariances, based on the volumes defined above.

## 3. Theoretical model

In this section we describe the theoretical framework of perturbation theory (PT), which is essential to understand the evolution of post-inflationary fluctuations in the matter density field $\delta$ into

[9] This rescaling is not valid in general, but can be performed when working under the assumption of a diagonal covariance matrix (see e.g. Eggemeier et al. 2020, for a similar rescaling to partially match the signal-to-noise ratio of different galaxy and halo samples). We note that the approach adopted in this analysis bears some limitations, since the amplitude of shot-noise is also rescaled, leading to slightly larger data uncertainties. However, this effect should partially account for the fact that we assume only a Gaussian recipe to predict the covariance matrix, thus underestimating the error.

the current large-scale distribution of galaxies via gravitational instability. This description is expected to be accurate only down to the mildly non-linear regime, where the amplitude of the density contrast $\delta$ is small enough to be perturbatively expanded. In the strong non-linear regime we expect this model to fail, as gravitational collapse leads to the formation of bound structures beyond the regime of validity of perturbative approaches.

For convenience, in the rest of this article we use the following notation for the integration over the infinite volume of a loop variable $\boldsymbol{q}$,

$$\int_{\boldsymbol{q}} \equiv \int \frac{\mathrm{d}^3 q}{(2\pi)^3}, \tag{6}$$

and adopt the following convention for the direct and inverse Fourier transform of the density contrast,

$$\delta(\boldsymbol{k}) \equiv (2\pi)^3 \int_{\boldsymbol{x}} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \delta(\boldsymbol{x}), \tag{7}$$

$$\delta(\boldsymbol{x}) \equiv \int_{\boldsymbol{k}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \delta(\boldsymbol{k}). \tag{8}$$

The three-dimensional Dirac function is represented with the standard notation $\delta_{\mathrm{D}}^{(3)}$. Finally, the power spectrum $P_{\mathrm{XX}}(k, z)$ of any component, matter or biased tracer, is defined as the auto-correlation of the corresponding density field $\delta_{\mathrm{X}}$, such that

$$\langle \delta_{\mathrm{X}}(\boldsymbol{k})\, \delta_{\mathrm{X}}(\boldsymbol{k}') \rangle \equiv (2\pi)^3 P_{\mathrm{XX}}(k)\, \delta_{\mathrm{D}}^{(3)}(\boldsymbol{k} + \boldsymbol{k}'), \tag{9}$$

where the presence of the Dirac function and the independence of the power spectrum from the orientation of the wave mode $\boldsymbol{k}$ reflect the underlying assumption of homogeneity and isotropy.

### 3.1. Eulerian framework and effective field theory

#### 3.1.1. Modelling of the non-linear matter power spectrum

We begin by summarising the most relevant outcomes of standard perturbation theory (SPT; see e.g. Bernardeau et al. 2002, for a review on the subject). Its main assumption is that on large scales the dynamics of dark matter can be approximated as that of a perfectly pressureless fluid, with negligible effects from particle shell-crossing in multi-streaming regions. Under the so-called Einstein–de Sitter (EdS) approximation, we can write the matter density contrast using a perturbative expansion,

$$\delta = \delta^{(1)} + \delta^{(2)} + \delta^{(3)} + \dots, \tag{10}$$

where at each order $n$ the individual contribution $\delta^{(n)}$ is a function of the linear density contrast $\delta_{\mathrm{L}}$[10],

$$\delta^{(n)}(\boldsymbol{k}) = \int_{\boldsymbol{q}_1 \cdots \boldsymbol{q}_n} \delta_{\mathrm{D}}^{(3)}(\boldsymbol{k} - \boldsymbol{q}_{1\dots n})\, F_n(\boldsymbol{q}_1, \dots, \boldsymbol{q}_n) \times \delta_{\mathrm{L}}(\boldsymbol{q}_1) \dots \delta_{\mathrm{L}}(\boldsymbol{q}_n). \tag{11}$$

Here $\boldsymbol{q}_{1\dots n} \equiv \boldsymbol{q}_1 + \dots + \boldsymbol{q}_n$, and $F_n$ is the $n$th order symmetrised PT kernel describing the non-linear interaction among fluctuations at different wave modes $\boldsymbol{q}_1, \dots, \boldsymbol{q}_n$. The vanishing argument of $\delta_{\mathrm{D}}^{(3)}$ reflects the translational invariance of the equations of motion in a spatially homogeneous universe.

Similarly, the non-linear matter power spectrum $P_{\mathrm{mm}}(k)$ can be expanded by combining Eqs. (9) and (10), leading to

$$P_{\mathrm{mm}}(k) = P_{\mathrm{L}}(k) + P^{\text{1-loop}}(k) + P^{\text{2-loop}}(k) + \dots, \tag{12}$$

where $P_{\mathrm{L}} \sim \langle \delta_{\mathrm{L}}^2 \rangle$ corresponds to the linear matter power spectrum, and at one-loop the only non-vanishing contributions are

$$\begin{aligned} P^{\text{1-loop}}(k) &= P_{22}(k) + P_{13}(k) \\ &= 2 \int_{\boldsymbol{q}} F_2^2(\boldsymbol{k} - \boldsymbol{q}, \boldsymbol{q})\, P_{\mathrm{L}}(|\boldsymbol{k} - \boldsymbol{q}|)\, P_{\mathrm{L}}(q) \\ &\quad + 6\, P_{\mathrm{L}}(k) \int_{\boldsymbol{q}} F_3(\boldsymbol{q}, -\boldsymbol{q}, \boldsymbol{k})\, P_{\mathrm{L}}(q). \end{aligned} \tag{13}$$

For the sake of completeness, we report the expanded expressions for the second- and third-order symmetrised kernels, $F_2(\boldsymbol{q}_1, \boldsymbol{q}_2)$ and $F_3(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3)$ in Appendix B.

The one-loop model in SPT, however, fails to accurately describe the non-linear damping of the acoustic oscillations due to bulk flow displacements (Eisenstein et al. 2007; Crocce & Scoccimarro 2008; Baldauf et al. 2015b). At first order, this effect can be reproduced in the theoretical model for $P_{\mathrm{mm}}(k)$ by a proper resummation of all infrared (IR) modes $q < k$, that is, of comoving separations larger than the one under consideration (see Crocce & Scoccimarro 2006, 2008 for a description of the BAO smearing in the context of renormalised perturbation theory).

A more standard procedure to include these corrections is based on the split of the linear power spectrum $P_{\mathrm{L}}$ as the sum of a smooth $P_{\mathrm{nw}}$ and wiggly $P_{\mathrm{w}}$ component (Seo et al. 2008; Baldauf et al. 2015b; Blas et al. 2016), that is

$$P_{\mathrm{L}}(k) = P_{\mathrm{nw}}(k) + P_{\mathrm{w}}(k). \tag{14}$$

At leading order, it is possible to estimate the amplitude of the damping factor making use of the Zeldovich approximation (Zel'dovich 1970). This leads to an expression for the leading-order, IR-resummed power spectrum,

$$P_{\mathrm{mm}}^{\text{IR-LO}}(k) = P_{\mathrm{nw}}(k) + \mathrm{e}^{-k^2\Sigma^2} P_{\mathrm{w}}(k), \tag{15}$$

where $\Sigma^2$, representing the variance of the relative displacement field (Eisenstein et al. 2007), is defined as

$$\Sigma^2 = \frac{1}{6\pi^2} \int_0^{k_{\mathrm{s}}} P_{\mathrm{nw}}(q) \left[ 1 - j_0\left(\frac{q}{k_{\mathrm{osc}}}\right) + 2\, j_2\left(\frac{q}{k_{\mathrm{osc}}}\right) \right] \mathrm{d}q. \tag{16}$$

Here $j_n$ is the $n$th order spherical Bessel function of the first kind, $k_{\mathrm{osc}} = 1/\ell_{\mathrm{osc}}$ is the wavelength corresponding to the BAO scale $\ell_{\mathrm{osc}} = 110\ h^{-1}\,\mathrm{Mpc}$[11], and $k_{\mathrm{s}} = 0.2\, h\,\mathrm{Mpc}^{-1}$ is the ultraviolet (UV) integration limit[12].

The next-to-leading order correction can be written by using the leading order term of Eq. (15) inside the expression for the one-loop corrections of Eq. (13). This leads to the final formulation for the non-linear IR-resummed power spectrum (Baldauf et al. 2015b; Blas et al. 2016),

$$\begin{aligned} P_{\mathrm{mm}}^{\text{IR-(LO+NLO)}}(k) &= P_{\mathrm{nw}}(k) + \left(1 + k^2\Sigma^2\right) \mathrm{e}^{-k^2\Sigma^2} P_{\mathrm{w}}(k) \\ &\quad + P^{\text{1-loop}}\left[P_{\mathrm{nw}} + \mathrm{e}^{-k^2\Sigma^2} P_{\mathrm{w}}\right](k), \end{aligned} \tag{17}$$

[10] In details, $\delta_{\mathrm{L}}$ represents the initial density contrast linearly extrapolated to the redshift under consideration.

[11] The value of $\ell_{\mathrm{osc}}$ should be varied as a function of the cosmological parameters. However, we cross-checked that for the relatively small parameter space explored in this analysis, its value do not deviate significantly from the one of a *Planck*-like cosmology.

[12] Despite the correct integration range being scale-dependent, as it accounts for all wave modes $q < k$, we fix the UV limit, similarly to what is done in Ivanov et al. (2020), as it can be shown that the integrand of Eq. (16) is not providing significant contributions at $q > 0.2\, h\,\mathrm{Mpc}^{-1}$.

where the square brackets of the last term mean that the evaluation of the one-loop correction is carried out using the leading order IR-resummed power spectrum in place of the linear one.

Another partial failure of the model, which is equally shared by any recipe based on perturbative methods, is that its range of validity is limited to quasi-linear scales, where the assumption of a pressureless fluid is still justified. However, on scales approaching the non-linear scale $k_{\rm NL}$[13], dark matter particles experience shell-crossing, effectively introducing a non-zero pressure, which under more realistic conditions is further enhanced by the presence of baryonic processes, such as galaxy formation, ISM cooling, and AGN and supernovae feedback. These effects can be described in terms of a non-trivial stress-energy tensor which, at leading order, results in an additional contribution to the matter power spectrum (Pueblas & Scoccimarro 2009; Carrasco et al. 2012; Baumann et al. 2012),

$$P_{\rm ctr}(k) = -2\,c_{\rm s}^2\,k^2\,P_{\rm mm}^{\rm IR\text{-}LO}(k), \tag{18}$$

usually denoted as counterterm in the EFTofLSS framework. Here, the parameter $c_{\rm s}$ can be interpreted as an effective speed of sound (Baumann et al. 2012; Carrasco et al. 2014; Baldauf et al. 2015a), reflecting the influence of short-wavelength perturbations, but accounts as well for the complex physics behind galaxy formation (when considering biased tracers of the matter density field).

Summarising, we can write the final expression for the model of the non-linear matter power spectrum as

$$P_{\rm mm}(k) = P_{\rm mm}^{\rm IR\text{-}(LO+NLO)}(k) + P_{\rm ctr}(k), \tag{19}$$

which contains one free parameter, $c_{\rm s}$, which must be treated as a nuisance parameter to be fitted against real or, in our case, simulated measurements.

#### 3.1.2. Modelling of the non-linear galaxy power spectrum

The general perturbative expansion of the galaxy density field $\delta_{\rm g}$ is based on the sum of all the individual operators that are a function of properties of the environment in which galaxies reside, such as the underlying matter density field and the large-scale tidal field. More precisely, this sum includes all those operators that are sourced by second derivatives of the gravitational potential $\Phi$ and the velocity potential $\Phi_v$ (see Desjacques et al. 2018, for a detailed review on the subject).

If we restrict our model to the one-loop prediction for the power spectrum, the relation between $\delta_{\rm g}$ and $\delta$ can be described considering only terms up to third order in the perturbations. In detail, this relation can be written as

$$\begin{aligned}\delta_{\rm g}(\boldsymbol{x}) = {} & b_1\,\delta(\boldsymbol{x}) + b_{\nabla^2\delta}\,\nabla^2\delta(\boldsymbol{x}) + \varepsilon_{\rm g}(\boldsymbol{x}) \\ & + \frac{b_2}{2}\,\delta^2(\boldsymbol{x}) + b_{\mathcal{G}_2}\,\mathcal{G}_2\left(\Phi_v\,|\,\boldsymbol{x}\right) + b_{\Gamma_3}\,\Gamma_3(\boldsymbol{x}) + \ldots,\end{aligned} \tag{20}$$

where each operator is multiplied by a free bias parameter that determines its overall amplitude[14]. The different terms in Eq. (20) can be summarised as follows.

(i) At leading order, the shot-noise-corrected galaxy density field can be expressed using a linear and local relation in $\delta$. This relation is characterised by a linear bias parameter, $b_1$, which simply rescales the underlying matter density contrast by a constant factor (Kaiser 1984).

(ii) The effect of short-range non-localities during the process of galaxy formation is characterised by the presence of higher derivatives of the gravitational potential (Bardeen et al. 1986; McDonald & Roy 2009; Desjacques et al. 2010). At leading order, the only non-zero term scales with the Laplacian of the matter density field, $\nabla^2\delta$, and has an amplitude regulated by the free parameter $b_{\nabla^2\delta}$. The formation of structures involves the collapse of matter from a finite region of space, which for dark matter halos is well approximated by their Lagrangian radius $R$. Since the estimation of the corresponding radius for a given galaxy sample can be cumbersome, here we absorb the value of $R$ inside the definition of $b_{\nabla^2\delta}$.

(iii) The impact of short-scale fluctuations on the galaxy density field at larger separations is determined by an additional stochastic field, $\varepsilon_{\rm g}$, which, under the assumption of Gaussian initial conditions, is completely uncorrelated from large-scale perturbations. If galaxies are randomly distributed, the stochastic contribution to the galaxy power spectrum is purely represented by the Poisson limit, $1/\bar{n}$, with $\bar{n}$ corresponding to the mean number density of the selected sample.

(iv) Moving to mildly non-linear scales, higher-order correlations of the density field appear (Coles 1993; Fry & Gaztanaga 1993), starting with a term proportional to $\delta^2$, characterised by a quadratic local bias $b_2$. This factor is expected from a spherically symmetric gravitational collapse, in which higher powers of $\delta$ become more relevant at progressively smaller scales. The third power of the matter density field is not included in Eq. (20) since its effect on the one-loop galaxy power spectrum is an extra contribution to the large-scale limit, which can be absorbed in the renormalization of the linear bias.

(v) Even starting with a purely local-in-matter-density bias expansion at the time of formation, non-linear evolution is responsible for the generation of large-scale tidal fields (Chan et al. 2012; Baldauf & Seljak 2012). At leading order, the corrections given by the tidal stress tensor are represented by a non-local quadratic bias, $b_{\mathcal{G}_2}$, and by the second-order Galileon operator, $\mathcal{G}_2$, defined as

$$\mathcal{G}_2\left(\Phi\,|\,\boldsymbol{x}\right) \equiv \left[\nabla_{ij}\,\Phi(\boldsymbol{x})\right]^2 - \left[\nabla^2\,\Phi(\boldsymbol{x})\right]^2. \tag{21}$$

In Fourier space, Eq. (21) can be written as

$$\mathcal{G}_2(\boldsymbol{k}) = \int_{\boldsymbol{q}} S\left(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}\right)\delta(\boldsymbol{q})\,\delta(\boldsymbol{k}-\boldsymbol{q}), \tag{22}$$

where

$$S\left(\boldsymbol{k}_1, \boldsymbol{k}_2\right) \equiv \frac{\left(\boldsymbol{k}_1\cdot\boldsymbol{k}_2\right)^2}{k_1^2\,k_2^2} - 1 \tag{23}$$

is the Fourier-space kernel corresponding to the second-order Galileon operator $\mathcal{G}_2$.

[13] This is typically defined as the scale at which the dimensionless matter power spectrum,

$$\Delta^2(k) \equiv \frac{k^3 P(k)}{2\pi^2},$$

becomes unity, that is, $\Delta^2(k_{\rm NL}) \equiv 1$.

[14] We note that this set of bias parameters needs to be renormalised before one can write the expression for the galaxy power spectrum (McDonald 2006; Assassi et al. 2014). This procedure is meant to remove the dependence on the cutoff scale used to define the galaxy density field, and to cancel the effect of higher-order bias parameters on large scales.

(vi) The next-to-leading-order correction to the tidal field can be obtained considering terms up to second-order in the potential of the displacement field (Chan et al. 2012). This contribution is represented by an additional non-local cubic bias, $b_{\Gamma_3}$, and by the operator

$$\Gamma_3(\boldsymbol{x}) \equiv \mathcal{G}_2\left(\Phi\,|\,\boldsymbol{x}\right) - \mathcal{G}_2\left(\Phi_v\,|\,\boldsymbol{x}\right), \tag{24}$$

whose net effect inside Eq. (20) is to include terms up to third order in perturbations of $\delta$.

All the terms giving a non-zero contribution to the one-loop galaxy power spectrum are listed in Eq. (20). The complete expression for $P_{\rm gg}$ then reads

$$P_{\rm gg}(k) = P_{\rm gg}^{\rm tree}(k) + P_{\rm gg}^{\rm 1\text{-}loop}(k) + P_{\rm gg}^{\rm ctr}(k) + P_{\rm gg}^{\rm noise}(k), \tag{25}$$

where the individual contributions can be written as

$$P_{\rm gg}^{\rm tree}(k) = b_1^2\, P_{\rm L}(k), \tag{26}$$

$$\begin{aligned} P_{\rm gg}^{\rm 1\text{-}loop}(k) &= P_{\rm gg,22}(k) + P_{\rm gg,13}(k) \\ &= 2\int_{\boldsymbol{q}} K_2^2\left(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}\right)\, P_{\rm L}\left(|\boldsymbol{k}-\boldsymbol{q}|\right)\, P_{\rm L}(q) \\ &\quad + 6\, b_1\, P_{\rm L}(k) \int_{\boldsymbol{q}} K_3\left(\boldsymbol{q}, -\boldsymbol{q}, \boldsymbol{k}\right)\, P_{\rm L}(q), \end{aligned} \tag{27}$$

$$\begin{aligned} P_{\rm gg}^{\rm ctr}(k) &= -2\, b_1 \left(b_1\, c_{\rm s}^2 + b_{\nabla^2\delta}\right) k^2\, P_{\rm L}(k) \\ &\equiv -2\, c_0\, k^2\, P_{\rm L}(k), \end{aligned} \tag{28}$$

$$P_{\rm gg}^{\rm noise}(k) = \frac{1}{\bar{n}}\left(1 + \alpha_{P,1} + \alpha_{P,2}\, k^2\right). \tag{29}$$

For the sake of completeness, a complete list of the individual one-loop corrections can be found in Appendix B. In the previous expressions, $K_2$ and $K_3$ are the symmetrised mode-coupling kernels for a generic biased tracer of the matter density field that follows the parametrization given in Eq. (20). In detail, they read

$$K_2(\boldsymbol{k}_1, \boldsymbol{k}_2) = b_1\, F_2(\boldsymbol{k}_1, \boldsymbol{k}_2) + \frac{1}{2} b_2 + b_{\mathcal{G}_2}\, S\left(\boldsymbol{k}_1, \boldsymbol{k}_2\right), \tag{30}$$

and

$$\begin{aligned} K_3(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) &= b_1\, F_3(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) + b_2\, F_2(\boldsymbol{k}_1, \boldsymbol{k}_2) \\ &\quad + 2\, b_{\mathcal{G}_2}\, S\left(\boldsymbol{k}_1, \boldsymbol{k}_{23}\right) F_2(\boldsymbol{k}_2, \boldsymbol{k}_3) \\ &\quad + 2\, b_{\Gamma_3}\, S\left(\boldsymbol{k}_1, \boldsymbol{k}_{23}\right) \left[F_2(\boldsymbol{k}_2, \boldsymbol{k}_3) - G_2(\boldsymbol{k}_2, \boldsymbol{k}_3)\right], \end{aligned} \tag{31}$$

where $G_2(\boldsymbol{k}_1, \boldsymbol{k}_2)$ is the standard one-loop kernel for the non-linear evolution of the velocity divergence field, and Eq. (31) has to be symmetrised with respect to its arguments $(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$.

Inside Eq. (29), $\alpha_{P,1}$ is a free nuisance parameter that accounts for deviations from a purely Poissonian shot-noise[15]. In addition, it is also required as a way to reabsorb the otherwise non-zero low-$k$ limit of one of the individual one-loop contributions, as explained in Appendix B. Similarly, $\alpha_{P,2}$ parametrises the next-to-leading order correction, which scale as $k^2$.

[15] This is expected since there is a physical separation under which two galaxies cannot simultaneously form, similarly to the exclusion effect for dark matter halos. The observed shot-noise can be either super- (signature of high-satellite star-forming galaxies) or sub-Poissonian (mostly typical of red central galaxies in massive halos), depending on the considered galaxy sample.

Since the leading-order higher-derivative correction is completely degenerate with the matter counterterm, as they are both proportional to the combination $k^2\, P_{\rm L}(k)$, we define a new more suited parameter,

$$c_0 \equiv b_1 \left(b_1\, c_{\rm s}^2 + b_{\nabla^2\delta}\right), \tag{32}$$

to avoid the presence of unnecessary degeneracies between the parameters of the model.

In the previous scheme we have deliberately omitted the resummation of infrared modes, but, similarly to the case discussed in Sect. 3.1.1, galaxy two-point clustering also has to be corrected for the effect of large-scale bulk motions. For this reason, we write the relations for the leading- and next-to-leading order IR-resummed galaxy power spectra (mimicking Eqs. (15) and (17)) as

$$P_{\rm gg}^{\rm IR-LO}(k) = b_1^2 \left[P_{\rm nw}(k) + \mathrm{e}^{-k^2\,\Sigma^2} P_{\rm w}(k)\right] + \frac{1}{\bar{n}}\left(1 + \alpha_{P,1}\right), \tag{33}$$

$$\begin{aligned} P_{\rm gg}^{\rm IR-(LO+NLO)}(k) &= b_1^2 \left[P_{\rm nw}(k) + \left(1 + k^2\Sigma^2\right)\mathrm{e}^{-k^2\Sigma^2} P_{\rm w}(k)\right] \\ &\quad + P_{\rm gg}^{\rm 1\text{-}loop}\left[P_{\rm nw} + \mathrm{e}^{-k^2\Sigma^2} P_{\rm w}\right](k) \\ &\quad + P_{\rm gg}^{\rm ctr}\left[P_{\rm nw} + \mathrm{e}^{-k^2\Sigma^2} P_{\rm w}\right](k) + P_{\rm gg}^{\rm noise}(k), \end{aligned} \tag{34}$$

where, once again, the square brackets of the second and third terms in Eq. (34) reflect how the evaluation of the one-loop and counterterm contributions is carried out sourcing the leading order IR-resummed matter power spectrum, $P_{\rm mm}^{\rm IR\text{-}LO}$, in place of the linear power spectrum, $P_{\rm L}(k)$.

### 3.1.3. Coevolution relations

A significant fraction of the bias parameters that have been introduced in this section enters in the expression for $P_{\rm gg}(k)$ only at higher-order, as clearly pointed out by the presence of only the linear bias $b_1$ in the expression for the leading-order galaxy power spectrum Eq. (33). This is significantly different from higher-order correlators of the galaxy density field, such as the galaxy bispectrum, for which both the local and non-local quadratic biases, $b_2$ and $b_{\mathcal{G}_2}$, appear also in the expression for the leading-order term, and can therefore be constrained with much better accuracy (Oddo et al. 2021; Eggemeier et al. 2021).

Given the poor constraining power of the galaxy power spectrum alone, it has become standard practice in real-data analyses to fix some of them to some physically motivated values or relations. This is important not only to obtain a larger constraining power for the remaining parameters, but also to ensure that none of them experiences strong degeneracies such as the one exhibited by the $(b_{\mathcal{G}_2}, b_{\Gamma_3})$ pair (see Appendix B). In this work, we test two different relations, which are briefly summarised in the next paragraphs.

As already explained in Sect. 3.1.2, even starting with a purely local-in-matter-density expression, $\delta_{\rm g}(\delta)$, at the time of formation, non-linear gravitational evolution is responsible for the generation of a large-scale tidal field (Fry 1996; Chan et al. 2012). This means that, even expressing the initial galaxy density field assuming only a spherically symmetric gravitational collapse – and thus with only local bias parameters $b_n \neq 0$ – tidal contributions appear at later times because of gravitational evolution, leading to the presence of non-negligible tidal biases. Assuming that the total number of objects is conserved in time, it is possible to find a relation between the late-time non-local

parameters and lower-order bias parameters, such that

$$b_{\mathcal{G}_2}^{\rm coev} = -\frac{2}{7}(b_1 - 1) + b_{\mathcal{G}_2}^{\mathcal{L}}, \tag{35}$$

$$b_{\Gamma_3}^{\rm coev} = -\frac{1}{6}(b_1 - 1) - \frac{5}{2} b_{\mathcal{G}_2} + b_{\Gamma_3}^{\mathcal{L}}, \tag{36}$$

where the bias parameters with a superscript $\mathcal{L}$ stand for the corresponding Lagrangian quantities, that is, at the time of formation. The previous relations are commonly referred to as coevolution, or local Lagrangian relations when setting to zero the Lagrangian bias, and have been extensively used in most real-data analyses to fix one or both non-local parameters (Feldman et al. 2001; Gil-Marín et al. 2015; Sánchez et al. 2016; Grieb et al. 2017). However, recent results (Lazeyras & Schmidt 2018; Abidi & Baldauf 2018) have indicated that measurements from numerical simulations seem to suggest lower values for $b_{\mathcal{G}_2}$ with respect to its local Lagrangian relation.

An alternative approach for fixing $b_{\mathcal{G}_2}$, found to be more accurate when compared to results from $N$-body simulations and derived using the excursion-set formalism, has been proposed by Sheth et al. (2013). In this case, it is possible to express $b_{\mathcal{G}_2}$ as a quadratic form in terms of the linear bias $b_1$, such that

$$b_{\mathcal{G}_2}^{\rm ex-set} = 0.524 - 0.547\, b_1 + 0.046\, b_1^2. \tag{37}$$

Such expressions are based on theoretical considerations on halo bias that only take into account the halo mass. As a consequence, they neglect potentially important effects, such as assembly bias (see e.g. Lazeyras et al. 2021; Barreira et al. 2021, for an analysis carried out, respectively, on dark matter halos and galaxies from hydrodinamical simulations). This means that their applicability to a real-data analysis must be carefully assessed (see Eggemeier et al. 2020; Pezzotta et al. 2021, for recent applications). Nonetheless, their use in this analysis is well justified, since we focus on HOD samples for which the assignment of a galaxy into a host halo is only determined by the mass of the latter.

In Sect. 5 we carry out tests to determine whether the previously defined relations can be employed to analyse clustering measurements adopting *Euclid* requirements.

### 3.2. Hybrid Lagrangian bias expansion model

In the previous sections, the relationship between the galaxy and the matter density field has been described through an Eulerian-based framework. However, this is not the only description of the galaxy power spectrum in the quasi-linear regime. Other approaches are possible, often based to various degrees on results from numerical simulations. We consider here the so-called 'hybrid Lagrangian' models. They draw from Lagrangian perturbation theory for the bias expression connecting galaxy and matter overdensities, but rely on simulations to capture the development of non-linearities when converting Lagrangian quantities to the observable Eulerian quantities.

The Lagrangian bias expansion describes the clustering of biased tracers in terms of a superposition of Lagrangian operators advected to Eulerian coordinates. It was first developed at one-loop in PT by Matsubara (2008), while Modi et al. (2020) proposed to combine the perturbative approach on bias with measurements of the advected operators from $N$-body simulations. This hybrid approach potentially allows us to push the bias expansion formalism to smaller scales with respect to purely perturbative approaches (see e.g. Hadzhiyska et al. 2021a; DeRose et al. 2023, for recent applications of the model to DES Y1 data and numerical simulations). Recent implementations include the works of Kokron et al. (2021) and Zennaro et al. (2023) in real space, and of Pellejero-Ibañez et al. (2023) in redshift space. In the present work we consider the implementation in the code `baccoemu` (Zennaro et al. 2023)[16]. It describes the Eulerian galaxy overdensity in terms of a second-order expansion of the Lagrangian galaxy density field $\delta_{\rm g}(\boldsymbol{q})$ where $\boldsymbol{q}$ is the Lagrangian position corresponding to the Eulerian position $\boldsymbol{x} = \boldsymbol{q} + \boldsymbol{\Psi}(\boldsymbol{q})$ with $\boldsymbol{\Psi}(\boldsymbol{q})$ being the displacement field-connecting initial and final positions. This means that the Eulerian overdensity is given by

$$1 + \delta_{\rm g}(\boldsymbol{x}) = \int_{\boldsymbol{q}} w(\boldsymbol{q})\, \delta_{\rm D}^{(3)}(\boldsymbol{x} - \boldsymbol{q} - \boldsymbol{\Psi}(\boldsymbol{q})), \tag{38}$$

where $w(\boldsymbol{q})$ expresses the weighting function that transforms the matter field into the galaxy field,

$$w(\boldsymbol{q}) = 1 + b_1^{\mathcal{L}}\, \delta(\boldsymbol{q}) + b_2^{\mathcal{L}} \left(\delta^2(\boldsymbol{q}) - \left\langle \delta^2 \right\rangle\right) + b_{s^2}^{\mathcal{L}} \left[s^2(\boldsymbol{q}) - \left\langle s^2 \right\rangle\right] + b_{\nabla^2\delta}^{\mathcal{L}}\, \nabla^2\delta(\boldsymbol{q}). \tag{39}$$

Here the total list of operators built on the matter density field $\delta(\boldsymbol{q})$ consists of $O = \{1, \delta, \delta^2, s^2, \nabla^2\delta\}$, and the individual entries correspond to the fully non-linear matter distribution, not weighted (1) and weighted ($\delta$) by the linear overdensity field, the squared linear overdensity field $\delta^2$, the squared traceless tidal field $s^2$ [17], and the Laplacian of the linear overdensity field $\nabla^2\delta$, respectively. Note that unlike the Eulerian bias basis presented before, the expansion in Eq. (39) does not include the next-to-leading-order correction to the tidal field, captured by the operator $\Gamma_3$[18]. This implies that the two bases are only equivalent under the assumption of coevolution for the Eulerian parameter $b_{\Gamma_3}$ (see Eq. 36).

The final model depends on four free parameters, the linear bias $b_1^{\mathcal{L}}$, the local quadratic bias $b_2^{\mathcal{L}}$, the tidal quadratic bias $b_{s^2}^{\mathcal{L}}$, and the higher-derivative bias $b_{\nabla^2\delta}^{\mathcal{L}}$, to which we add the extra free parameter $\alpha_{P,1}$ to account (at first order) for non-Poissonian shot-noise, in the same way as it is done in the Eulerian PT model. We use a different notation for the quadratic tidal bias, since the definition of the tidal field operator is slightly different from the one presented in Sect. 3.1. The same is true for the Laplacian bias, which in this case only models higher-derivative corrections, but could also (partially) absorb unmodelled non-local effects coming from higher orders, extra physics, such as baryonic effects, or the smoothing of the density field performed in Lagrangian space.

The galaxy power spectrum can then be expressed as

$$P_{\rm gg}(k) = \sum_{i,j} b_i^{\mathcal{L}}\, b_j^{\mathcal{L}}\, P_{ij}(k) + \frac{1 + \alpha_{P,1}}{\bar{n}_{\rm g}}, \tag{40}$$

[16] https://bacco.dipc.org/emulator.html

[17] In this expansion $s_{ij}(\boldsymbol{q}) = \partial_i\, \partial_j\, \Phi(\boldsymbol{q}) - \delta_{ij}^{\rm K}\, \delta(\boldsymbol{q})$, where $\delta_{ij}^{\rm K}$ is the Kronecker delta function. This definition matches the one of the second-order Galileon operator $\mathcal{G}_2$, as in Eq. (21), with the only difference being that the two operators are defined in Lagrangian and Eulerian space, respectively. This is different from the parametrisation adopted in e.g. Desjacques et al. (2018), where the quadratic tidal operator is defined as $K_{ij} = \partial_i\, \partial_j\, \Phi(\boldsymbol{k}) - \frac{1}{3}\delta_{ij}^{\rm K}\, \delta(\boldsymbol{k})$.

[18] We note that there is nothing preventing a complete expansion up to third order in $\delta$ even in Lagrangian space. While the presence of this contribution may be partially relevant in terms of field level or higher-order statistics, such as the galaxy bispectrum, accuracy checks carried out by the `baccoemu` team have led to the conclusion that neglecting the cubic operator is a robust assumption for the analysis of the galaxy power spectrum.

where $P_{ij}(k)$ are the 15 cross-spectra of the five previously defined advected operators. To compute the $P_{ij}$ terms, `baccoemu` has been trained with high-resolution $P_{ij}$ measurements from 800 combinations of cosmologies and redshifts, obtained applying the cosmology-rescaling technique to four main $N$-body simulations (Angulo & White 2010; Zennaro et al. 2019; Contreras et al. 2020).

As a final remark, we note that, even if it is possible to find a relation between the Lagrangian and Eulerian bias parameters, the two sets do not exactly correspond to the same physical quantities. This happens because in the purely perturbative Eulerian framework they properly represent the response of galaxy formation to large-scale perturbations, whereas in the hybrid Lagrangian one this physical meaning is lost due to the extrapolation of the individual operators to the non-linear regime (since the individual operators of the bias expansion are measured from $N$-body simulations and thus contain higher-order contributions).

## 4. Model selection and fitting procedure

In this section we describe the methodology used to determine the best combination between different models, scale cuts, and bias configurations. In addition we list the details of the fitting procedure and the priors of the selected parameter spaces.

### 4.1. Performance metrics

In the context of model selection, the most relevant aspects to take into consideration are the range of validity of a given model and the precision and accuracy of the constraints on the parameters of interest. The procedure that we adopt is based on the selection of the maximum wave mode $k_{\rm max}$ up to which the model is still capable of providing a good description of the data vectors, while still recovering the correct input parameters. This can be quantified by means of three different performance metrics (employed in e.g. Osato et al. 2019; Eggemeier et al. 2020, 2021; Pezzotta et al. 2021), which are described in the next subsections.

#### 4.1.1. Figure of bias

One of the main requirement that the theoretical model has to satisfy is that its fit to the data return unbiased model parameters. The parameters controlling bias, shot-noise, and counterterms can be effectively treated as free nuisance parameters, to be marginalised over after sampling the joint posterior distribution. The set of parameters of interest is therefore restricted to the cosmological parameters, in our case $\boldsymbol{\theta} \equiv \{h,\, \omega_{\rm c}\}$.

We quantify the unbiasedness of the model in the recovery of $\boldsymbol{\theta}$ in terms of the figure of bias (FoB) defined as

$$\mathrm{FoB}(\boldsymbol{\theta}) \equiv \left[ \left( \langle \boldsymbol{\theta} \rangle - \boldsymbol{\theta}_{\rm fid} \right)^{\intercal} S^{-1}(\boldsymbol{\theta}) \left( \langle \boldsymbol{\theta} \rangle - \boldsymbol{\theta}_{\rm fid} \right) \right]^{\frac{1}{2}}, \tag{41}$$

where $\langle \boldsymbol{\theta} \rangle$ and $\boldsymbol{\theta}_{\rm fid}$ represent the mean of the posterior distribution of the selected parameters and their fiducial values, respectively, and $S(\boldsymbol{\theta})$ is a square matrix containing the auto- and cross-covariance among all the entries of the vector $\boldsymbol{\theta}$[19]. The meaning of Eq. (41) is straightforward: we are quantifying the deviation of the posterior distribution from the fiducial values of the corresponding parameters, and expressing this information in terms of the intrinsic error of those parameters. In the case where $\boldsymbol{\theta}$ consists of only one parameter, the FoB simply expresses how far the posterior is from the fiducial value in units of the standard deviation of the parameter, with the 68% and 95% percentiles corresponding to values of FoB of 1 and 2, respectively. Note that when considering more than one parameter these values change, as they need to be computed by directly integrating a multivariate normal distribution with the corresponding number of dimensions. For $n = 2$, we evaluate that the new thresholds for the 68% and 95% percentiles are 1.52 and 2.49, respectively.

#### 4.1.2. Goodness of fit

The goodness of fit quantifies the consistency of the theoretical model $P^{\rm th}$ with the input data vector $P^{\rm data}$. We consider the standard $\chi^2$ test, corresponding to

$$\chi^2(\boldsymbol{\theta}) = \sum_{i=1}^{N_{\rm bins}} \sum_{j=i}^{N_{\rm bins}} \left[ P_i^{\rm th}(\boldsymbol{\theta}) - P_i^{\rm data} \right] C_{ij}^{-1}(\boldsymbol{\theta}) \left[ P_j^{\rm th}(\boldsymbol{\theta}) - P_j^{\rm data} \right]. \tag{42}$$

This results in a distribution of $\chi^2$ values across the sampled parameter space. Instead of picking the $\chi^2$ corresponding to the maximum-likelihood position, whose estimation from the sampled posterior distribution is subject to noise[20], we compute the posterior-averaged value, $\left\langle \chi^2 \right\rangle$, from a weighted average over all sampled parameter combinations, which is instead a more stable quantity (see Appendix E for a comparison between the two approaches). The posterior-averaged $\chi^2$ is then compared to the predictions from the 68% and 95% percentiles of the $\chi^2$ distribution with the corresponding number of degrees of freedom. The latter is simply defined as $N_{\rm dof} = N_{\rm bins} - N_{\rm pars}$, where $N_{\rm bins}$ is the total number of independent wave mode bins up to the selected $k_{\rm max}$, and $N_{\rm pars}$ is the total number of free parameters of the model.

#### 4.1.3. Figure of merit

Finally, each configuration of the model – that is a given scale cut and bias assumptions – is inspected to determine its statistical power in constraining the parameters $\boldsymbol{\theta}$. For this purpose, similarly to what is done for the figure of bias, we define a figure of merit (FoM) for a given set of model parameters $\boldsymbol{\theta}$ as (Wang 2008)

$$\mathrm{FoM}(\boldsymbol{\theta}) = \left[ \det\left( S(\boldsymbol{\theta}) \right) \right]^{-1/2}, \tag{43}$$

where $S(\boldsymbol{\theta})$ is once again the covariance matrix of the parameters $\boldsymbol{\theta}$, and $\det(S)$ its determinant. The meaning of this quantity can be more clearly understood assuming a flat posterior distribution with null correlation between the entries of $\boldsymbol{\theta}$. In this case, $\det(S)$ represents the volume of the hyper-rectangle over which the posterior distribution of $\boldsymbol{\theta}$ is distributed. Similarly, for non-zero parameter correlations, $\det(S)$ represents the hyper-volume contained in the hyper-surface defined by the covariance matrix $S$. Therefore, a high value of the FoM corresponds to a more statistically significant constraint of the model parameters.

[19] This means that, for the case we are considering, where $\boldsymbol{\theta} = \{h,\, \omega_{\rm c}\}$, $S$ is a $2 \times 2$ matrix containing the variance of $h$ and $\omega_{\rm c}$ on its diagonal, and the cross-covariance between them on the off-diagonal entries.

[20] An optimal research of the maximum-likelihood position could be carried out employing a $\chi^2$ minimiser, or in the context of a more frequentist approach based on a profile likelihood, which we do not perform in this work.

In order to visualise how much can be gained by pushing the model to higher $k_{\rm max}$ values, in the next section we plot the FoM of each individual configuration normalised to that of a reference case, corresponding to the configuration with the EFTofLSS model at $k_{\rm max} = 0.1\,h\,{\rm Mpc}^{-1}$ with all nuisance parameters sampled as free parameters.

### 4.2. Fitting procedure

In order to properly sample the posterior distribution we need to compute the galaxy power spectrum and the likelihood for a large number of points in parameter space. To achieve convergence while keeping the number of evaluations as low as possible, an efficient sampling algorithm is needed.

All the results presented in this work have been obtained using a nested sampling approach (Skilling 2006), which differs from a standard Metropolis–Hastings (Metropolis et al. 1953; Hastings 1970) Markov-chain sampler in a number of ways. The main difference is that, using nested sampling, the whole hyper-dimensional parameter space is explored within the specified priors by means of a given number of live points, which are subsequently modified to track the posterior distribution of the parameters according to the value of the evidence. In this analysis we make use of the public code `PyMultiNest` (Buchner et al. 2014) with a total number of 1800 live points, after having checked that the output posterior distribution has properly converged with this number. Further details, together with a comparison of different samplers, are presented in Appendix F.

We adopt the approach of a full-shape analysis. This means that we directly sample the cosmological parameter space, with a model galaxy power spectrum that is generated at each step. A single evaluation of the theory models presented in Sect. 3 can take up to few seconds, since it combines a call to the Boltzmann solver to obtain linear theory predictions, and a call to the routines responsible for computing the non-linear corrections. Since the typical number of model evaluations for a single Markov chain can reach order of $O(10^6)$, the final running time necessary to obtain a converged posterior distribution can take up to several days.

In order to speed up the model evaluation, we make use of the publicly available `COMET` package (Eggemeier et al. 2022)[21] to emulate the EFTofLSS model, providing an evaluation of the full one-loop prediction in about $O(10\,{\rm ms})$. The code has been validated against a set of 1500 theory data vectors in a range of redshifts that covers the one we explore in this analysis, showing an averaged 0.1% systematic error for the final $P_{\rm gg}(k)$ model, and it is therefore suited to be used for this analysis[22]. The evaluation of the hybrid Lagrangian-bias-based model is instead carried out using the public emulator `baccoemu`, as mentioned in Sect. 3.2.

In all the cases, we assume a Gaussian likelihood function defined as

$$-2\ln\mathcal{L}(\boldsymbol{\theta}) = \sum_{i=1}^{N_{\rm bins}}\sum_{j=i}^{N_{\rm bins}}\left[P_i^{\rm th}(\boldsymbol{\theta}) - P_i^{\rm data}\right] C_{ij}^{-1}(\boldsymbol{\theta})\left[P_j^{\rm th}(\boldsymbol{\theta}) - P_j^{\rm data}\right], \tag{44}$$

which is computed at each point in parameter space explored by the sampler, and whose value is used to determine whether to assign to the current point one of the live points. The final output, which is saved to external files ready to be post-processed, consists of a list of points in parameter space together with the corresponding value of the log-likelihood.

[21] https://pypi.org/project/comet-emu/

[22] Further validation tests have been carried out against other codes owned by the authors of this paper, as shown in Appendix D.

### 4.3. Parameter priors

Our parameter space consists of both cosmological and nuisance parameters. Sampled cosmological parameters comprise the Hubble parameter $h$ and the cold dark matter density parameter $\omega_{\rm c}$. The latter can be constrained only through the full-shape of the galaxy power spectrum, especially via the position of the matter-radiation equality $k_{\rm eq}$, since geometric distance information is lost due to the fact that we conduct our analysis using real-space coordinates. For the same reason, $h$ can be artificially constrained because we fix all the other parameters affecting the amplitude of the matter power spectrum (Sánchez et al. 2022)[23]. We keep fixed the baryon density parameter $\omega_{\rm b}$ and the scalar spectral index $n_{\rm s}$, since galaxy clustering measurements on their own are not able to constrain them with the same level of precision of CMB data. At the same time, since in real space the primordial scalar amplitude $A_{\rm s}$ is strongly degenerate with the linear bias parameter $b_1$, at least on sufficiently large scales[24], we keep $A_{\rm s}$ fixed to its fiducial value, along with the rest of the cosmological parameters, to the values shown in Table 1.

The nuisance parameters are split into two sets, depending on the considered model. The parameters of the Eulerian bias expansion are composed of a mixture of bias parameters, $\{b_1, b_2, b_{\mathcal{G}_2}, b_{\Gamma_3}\}$, counterterms, $\{c_0\}$, and shot-noise parameters, $\{\alpha_{P,1}, \alpha_{P,2}\}$. All of them enter in the final expression for the galaxy power spectrum as shown in Sect. 3.1. When testing the bias relations presented in Sect. 3.1.3, the parameters subject to the bias relations are not sampled over, but computed at each step in the chain as a function of the lower-order bias parameters. The scale-dependent noise parameter $\alpha_{P,2}$ is kept fixed to 0 for the majority of the runs we carry out, except for the ones presented in Sect. 5.3, where we explicitly test the constraining power of the EFTofLSS model on this parameter in the range of redshifts that we are considering.

For the hybrid Lagrangian model we sample over a different set of bias parameters, $\left\{b_1^{\mathcal{L}}, b_2^{\mathcal{L}}, b_{s^2}^{\mathcal{L}}, b_{\nabla^2\delta}^{\mathcal{L}}\right\}$, and shot-noise, $\{\alpha_{P,1}\}$. In this case we do not consider relations among bias parameters, but every run will assume the full set.

When not mentioned otherwise, we adopt a completely agnostic approach, setting an uninformative flat prior for all the parameters, as shown in Table 3. The size of the prior for the two cosmological parameters and for most of the nuisance parameters has been selected to prevent the posterior distribution from becoming dominated by the imposed prior.

## 5. Results

In this section we present the results obtained by fitting the data samples presented in Sect. 2 with the two theoretical models described in Sect. 3. We start off with a comparison between

[23] In practice, expressing the galaxy power spectrum in $h\,{\rm Mpc}^{-1}$ units makes possible to constrain evolution parameters $(h, A_{\rm s}, w_0, w_{\rm a}, \omega_{\rm K}, \ldots)$ even when they are varied together in the same fit of the galaxy power spectrum data vector. This happens because with this set of units it is possible to break the degeneracy experienced by the evolution parameters that is otherwise present when expressing the data vector in ${\rm Mpc}^{-1}$ units (Sánchez et al. 2022).

[24] While the linear galaxy power spectrum depends on the combination $b_1^2 A_{\rm s}$, the non-linear corrections depend on a different combination of the linear bias and the scalar amplitude, so that they can in principle break the degeneracy. However, since loop corrections are subdominant with respect to the amplitude of the linear galaxy power spectrum, we find that a strong degeneracy is still present, even when including mildly non-linear scales in the fits.

the performance of these models in Sect. 5.1, and leave to later sections a more detailed description of the model selection carried out for the EFTofLSS model in terms of scale cuts and bias relations. For compactness, we call this model simply the EFT model.

### 5.1. Performance of Eulerian and hybrid Lagrangian bias expansion

In this section we carry out a comparison between the Eulerian expansion and the hybrid Lagrangian bias approach implemented in baccoemu in terms of the three performance metrics previously defined in Sect. 4.1. For this goal, we focus on fitting the galaxy power spectra of the Model 3 HOD sample at the redshifts of the four comoving snapshots, using both the rescaled *Euclid*-like covariances, and the ones from the full simulation box. For each case, we run multiple chains to assess the stability of the results as a function of the maximum wave mode $k_{\rm max}$. The latter is selected in the range $[0.10, 0.45]\ h\,{\rm Mpc}^{-1}$ using a linear spacing of $\Delta k = 0.05\,h\,{\rm Mpc}^{-1}$, for a total of eight different cases.

We select two different configurations of the EFT model. The first one corresponds to the case in which all the nuisance parameters $\{b_1, b_2, b_{\mathcal{G}_2}, b_{\Gamma_3}, c_0, \alpha_{P,1}\}$ are free to vary, with the only exception of the scale-dependent shot-noise parameter $\alpha_{P,2}$ which we set to zero. In the second one we additionally fix the cubic tidal bias $b_{\Gamma_3}$ to its coevolution relation (Eq. (36)). The latter case is chosen in order to provide an alternative model based on the Eulerian expansion of Sect. 3.1, with the same assumptions on galaxy bias as in baccoemu (see discussion in Sect. 3.2). In addition, this is one of the best configurations when considering the performance metrics on the combination $\{h, \omega_{\rm c}\}$, as we properly validate in Sect. 5.4.

As for baccoemu, we leave all bias parameters $\left\{b_1^{\mathcal{L}}, b_2^{\mathcal{L}}, b_{s^2}^{\mathcal{L}}, b_{\nabla^2\delta}^{\mathcal{L}}\right\}$ free to vary, with the addition of the parameter controlling the amplitude of the non-Poissonian stochastic noise, $\alpha_{P,1}$. Since baccoemu is an emulator based on $N$-body simulations, it is affected by two sources of noise: first, the emulation error, that is the noise introduced by the accuracy of the trained neural network itself; second, the training set error, that is the inaccuracies already present in the data used for training. The former is a scale-dependent quantity, which becomes progressively larger at small scales and caps at a maximum 0.5% of the galaxy power spectrum signal at $k \sim 0.7\,h\,{\rm Mpc}^{-1}$ for ΛCDM cosmologies well within the allowed parameter space; it can get to the order of $O(1\%)$ of the power spectrum signal for cosmologies closer to the limits of the emulator parameter space (Zennaro et al. 2023). On the other hand, the intrinsic error of the training set is induced by the cosmology-rescaling technique employed during its construction; it once again depends on scale, and is subpercent in the case of ΛCDM cosmologies, but could reach percent levels when also massive neutrinos and dynamical dark energy are considered (Contreras et al. 2020; Zennaro et al. 2023). To account for these combined effects, we consider two cases for the chains run with baccoemu. In the first one we employ the same covariance matrix used to analyse the data galaxy power spectra as in the EFT chains, while in the second one we add in quadrature a theory error corresponding to 0.5% of the galaxy power spectrum signal[25].

**Table 3.** Model parameters, split into cosmological and nuisance ones, with the latter further divided into the two bias models described in Sect. 3.

| | Parameter | Prior |
|---|---|---|
| | Cosmology | |
| | $h$ | $\mathcal{U}[0.55, 0.85]$ |
| | $\omega_{\rm c}$ | $\mathcal{U}[0.08, 0.16]$ |
| | Eulerian bias expansion | |
| Bias | $b_1$ | $\mathcal{U}[0.25, 4]$ |
| | $b_2$ | $\mathcal{U}[-10, 10]$ |
| | $b_{\mathcal{G}_2}$ | $\mathcal{U}[-4, 4]$ or fixed to Eq. (37) |
| | $b_{\Gamma_3}$ | $\mathcal{U}[-8, 8]$ or fixed to Eq. (36) |
| Counterterm | $c_0\,[({\rm Mpc}/h)^2]$ | $\mathcal{U}[-100, 100]$ |
| Shot-noise | $\alpha_{P,1}$ | $\mathcal{U}[-1, 2]$ |
| | $\alpha_{P,2}\,[({\rm Mpc}/h)^2]$ | $\mathcal{U}[-5, 5]$ or fixed to 0 |
| | Hybrid Lagrangian bias expansion | |
| Bias | $b_1^{\mathcal{L}}$ | $\mathcal{U}[-1, 3]$ |
| | $b_2^{\mathcal{L}}$ | $\mathcal{U}[-3, 3]$ |
| | $b_{s^2}^{\mathcal{L}}$ | $\mathcal{U}[-10, 10]$ |
| | $b_{\nabla^2\delta}^{\mathcal{L}}\,[({\rm Mpc}/h)^2]$ | $\mathcal{U}[-10, 10]$ |
| Shot-noise | $\alpha_{P,1}$ | $\mathcal{U}[-1, 1]$ |

**Notes.** The nuisance parameters consist of bias parameters, EFTofLSS counterterm, and shot-noise terms. For each parameter, the imposed prior is specified in the last column of the table. The letter $\mathcal{U}$ stands for a uniform distribution, with edges identified by the first and second element of the pair, respectively.

In Fig. 3 we show the performance metrics (FoB, FoM) in the case of the realistic *Euclid*-like volume. Note that we do not show the averaged reduced $\chi^2$ in this case, since the rescaled covariance matrix does not describe the fluctuations in the data vector, and therefore the collection of $\sum(P^{\rm th} - P^{\rm data})^2/\sigma^2$ values deviates from a $\chi^2$ distribution. At all redshifts the fits obtained with both models display a FoB within the 68% confidence interval up to $k_{\rm max} = 0.45\,h\,{\rm Mpc}^{-1}$, with only a partial preference for the EFT framework when considering the value of the FoB, which is anyway consistent to $1\sigma$ for all the cases.

In the figure, the FoM is normalised by the value obtained with the EFT configuration at $k_{\rm max} = 0.1\,h\,{\rm Mpc}^{-1}$, to show relative gains. As expected, the FoM of the EFT model is larger with a fixed $b_{\Gamma_3}$, because of the smaller number of free parameters. Similarly, we note that the combined constraining power on $(h, \omega_{\rm c})$ of baccoemu is degraded when including theory errors in the data covariance matrix.

When comparing the two different models, we note that baccoemu without including theory errors (orange line) reaches its maximum FoM value, comparable to the maximum value achieved by the Eulerian model (blue line), already at a lower $k_{\rm max}$, of about $0.25$–$0.3\,h\,{\rm Mpc}^{-1}$. This is a consequence of the extra parameter, $b_{\Gamma_3}$, present in the Eulerian bias model. As further evidence, the Eulerian bias model display higher FoM values when considering a fixed $b_{\Gamma_3}$ (light-blue line), in particular on scales $k_{\rm max} < 0.2\,h\,{\rm Mpc}^{-1}$. Above this threshold, we note that the EFT configuration features a slightly larger FoM than the

[25] The assumption of choosing an extra contribution of 0.5% of the power spectrum is well justified by the fact that we are only exploring a ΛCDM parameter space that is completely contained within the prior range of the emulator.

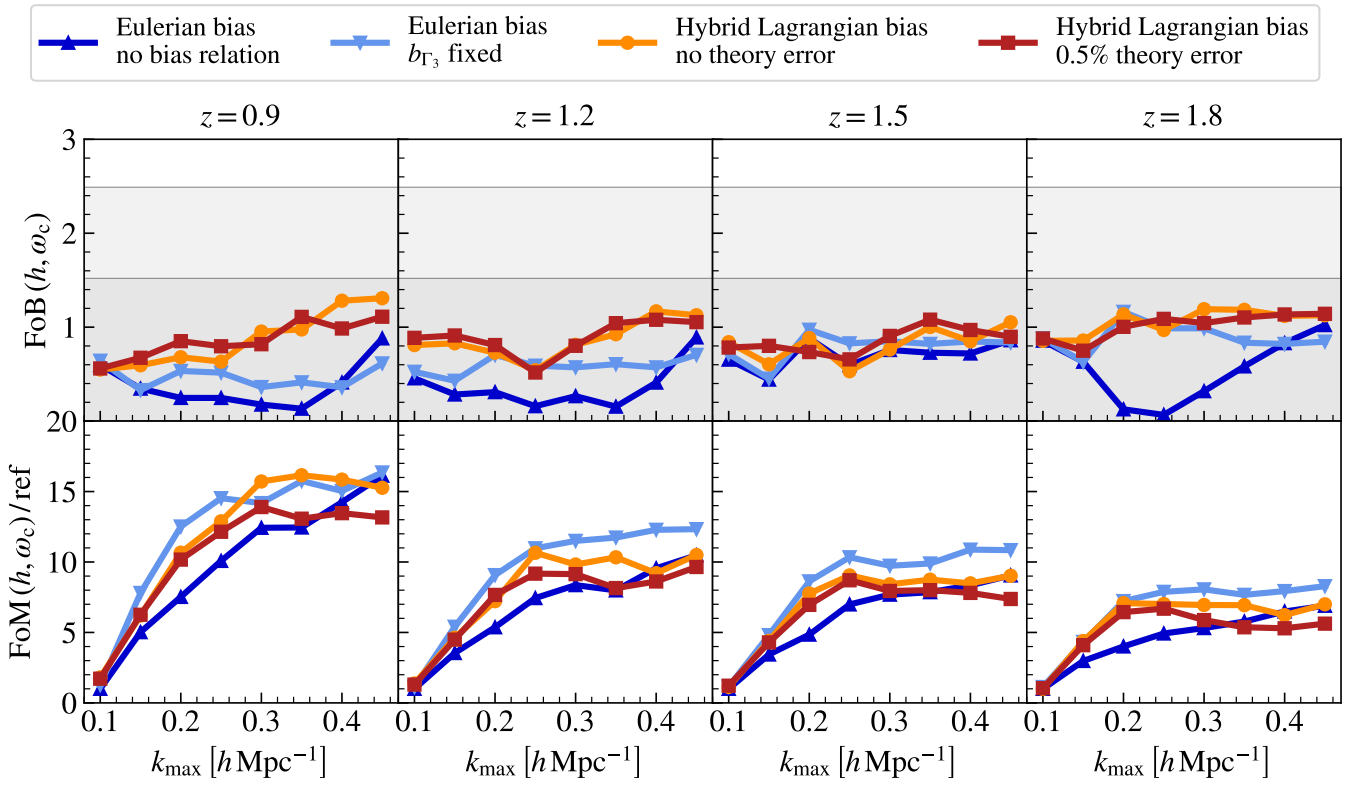


**Fig. 3.** Performance metrics (FoB in the top row and FoM in the bottom row) extracted from the Model 3 HOD samples as a function of the maximum wave mode $k_{max}$ of the fit, assuming the rescaled covariance matrices matching the four *Euclid* spectroscopic redshift bins described in Sect. 2.3. Different curves correspond to different models, as described in the legend. The FoM panels are normalised in units of the reference FoM, corresponding to the one of the EFT model with all parameters free at $k_{max} = 0.1\,h\,\mathrm{Mpc}^{-1}$. The grey bands in the FoB panels represent the 68% and 95% percentiles of the corresponding FoB distribution, as explained in Sect. 4.1.1.

one of the hybrid model, with the only exception of the $z = 0.9$ snapshot, for which the two curves have a similar amplitude at all scales (light blue vs. orange).

Except for the main EFT case, including smaller scales does not seem to increase the FoM beyond a scale of about $0.3\,h\,\mathrm{Mpc}^{-1}$. Since on these scales the theory error associated to `baccoemu` is of similar magnitude as the data covariance, we note that, including the extra 0.5% contribution (red line), the FoM starts flattening at a slightly lower $k_{max} \sim 0.25\,h\,\mathrm{Mpc}^{-1}$. This is mostly noticeable for the $z = 0.9$ snapshot, for which shot-noise becomes the dominant contribution at a much larger $k_{max}$.

A plot similar to the one in Fig. 3 is shown in Fig. 4, this time considering the covariance matrix corresponding to the full simulation volume of about $54\,h^{-3}\,\mathrm{Gpc}^3$. Since now the covariance matrix correctly represents the statistical fluctuations in the data vectors, we additionally show the $\chi^2$ averaged over the chain and normalised to the numbers of degrees of freedom. In this case, it is clear that not accounting for the theory error of `baccoemu` can lead to a bias in the cosmological parameters, most notably at low redshift. On the contrary, including the reference 0.5% theory error is enough to recover unbiased results, with the sole exception of the case at $k_{max} = 0.45\,h\,\mathrm{Mpc}^{-1}$ and $z = 0.9$.

The EFT model also returns unbiased measurements, with some spurious configurations outside the $1\sigma$ confidence interval for low $k_{max}$ values at $z = 1.5$. The main reason for this effect is likely imputable to the presence of projection effects when marginalising the posterior distribution in the $\{h, \omega_c\}$ plane, as we explain later in Sect. 5.4. The averaged $\chi^2$ behaves in a consistent way between the two models, displaying an amplitude that is constantly lower than the 95th percentile of the corresponding $\chi^2$ distribution for both sets of curves, with the only exception of the largest $k_{max}$ values of the $z = 1.8$ snapshot.

In terms of goodness of fit, we note that, due to our choice of reporting the posterior-averaged $\chi^2$ value instead of the maximum-likelihood value, the normalised $\chi^2$ can start off with values larger than 1 at low $k_{max}$ values. This is mostly caused by the non-gaussianity of the sampled posterior distribution when the data vectors cannot properly constrain the whole set of sampled parameters. In this case the averaged posterior can increase the value of the $\chi^2$ and making it appear artificially larger. In Appendix E we include an example using the maximum-likelihood $\chi^2$, showing how in this case the goodness of fit typically assumes values consistent with 1 on those scales.

While the constraining power of `baccoemu` is in this case limited by the theory error being of similar order as the statistical error of the synthetic data considered, it is highly competitive with the Eulerian approach on scales that are free from this limitation, at $k_{max} \lesssim 0.2\,h\,\mathrm{Mpc}^{-1}$. On the one hand, the full-volume test considered here leads to very conservative results: the errors associated with the full volume of the Flagship simulation are roughly a factor 2 smaller than the scaled errors considered in this work, and these scaled errors for Model 3, in turn, are roughly another factor 2 smaller than the errors expected assuming the volumes and number densities for typical redshift bins of the spectroscopic sample described in the forecasts of Euclid Collaboration (2020). In addition, all covariance matrices are computed in the Gaussian approximation, which might underestimate the amplitude of the errors at small scales. On the other hand, these results provide a motivation to

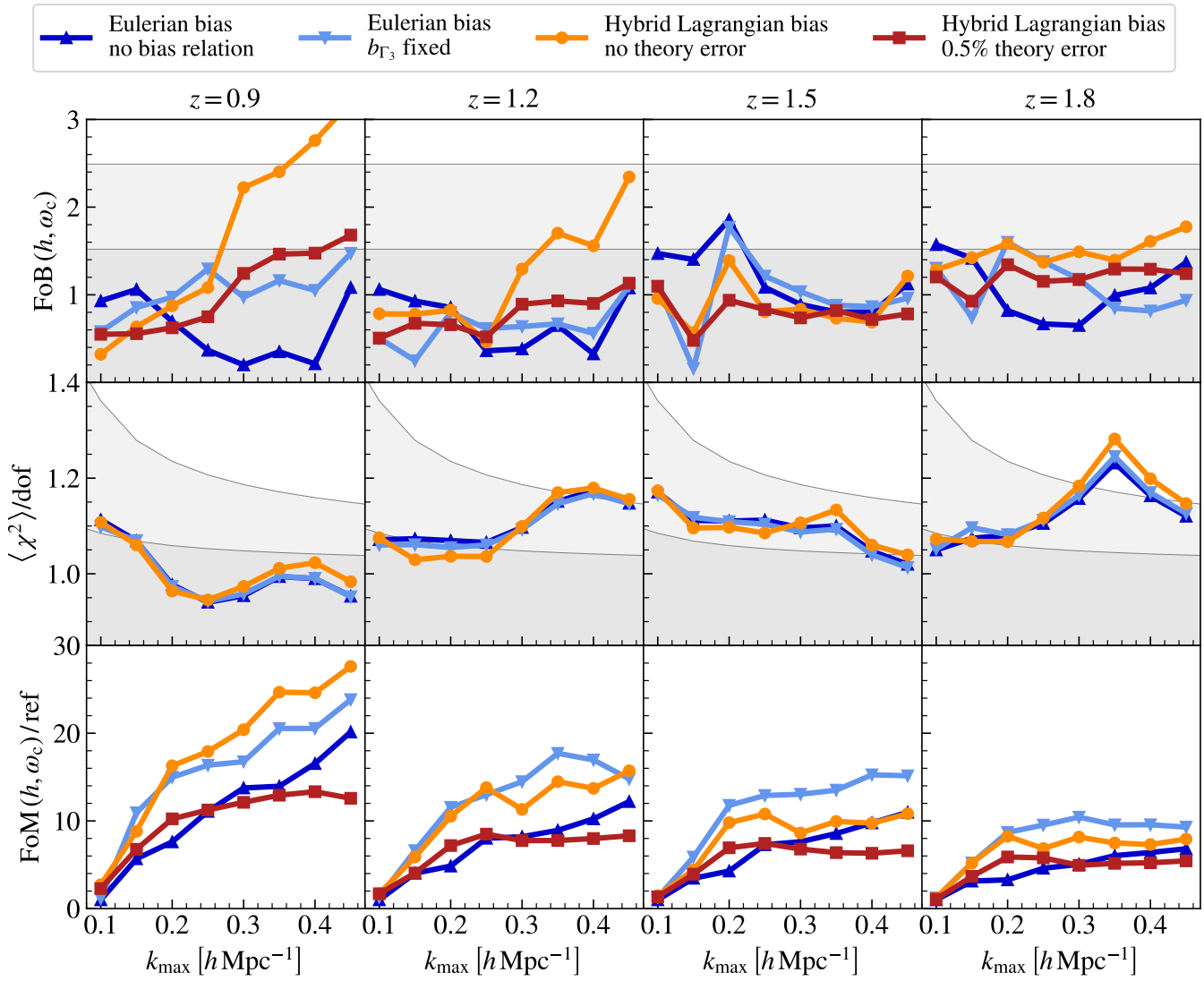


**Fig. 4.** Same as Fig. 3 but assuming a covariance matrix corresponding to the full simulation volume. The additional middle row corresponds to the averaged $\chi^2$ normalised by the total number of degrees of freedom. Similarly to the FoB panels, the grey shaded bands in the $\chi^2$ panels mark the 68th and 95th percentile of the corresponding $\chi^2$ distribution with the corresponding number of degrees of freedom.

further reduce the noise associated with emulators – for example through larger training sets, and employing Zeldovich control variates (Chartier et al. 2021; Kokron et al. 2022). This is key for the design of the next generation of emulators. In fact, even if the configuration without theory errors shows the limitation of currently available codes, the corresponding FoM curve highlights the potential gain achievable with a more accurate version of the emulator.

### 5.2. Testing the EFT model: Fixed cosmology

In the rest of this section we focus on testing the range of validity of the EFT model using different scale cuts, bias relations, and reference volumes. This test is limited only to the EFT model because, as shown in the previous subsection, we cannot run fits using `baccoemu` without accounting for the extra contribution from theory errors in the covariance matrix, especially when considering the extremely large precision of the full-box covariance. In order to assess the level of accuracy of the EFT model and determine its range of validity, we first carry out fits at fixed cosmology assuming the full volume of the simulation box. In this way, we focus exclusively on the performance of one-loop galaxy bias prediction with highly precise measurements, testing which scale cuts and bias relations lead to the best agreement between the theory model and the input data vectors.

The validation of the model includes an accuracy test consisting in recovering fiducial values for the linear bias $b_1$ and the shot-noise parameter $\alpha_{P,1}$ determined from the large-scale limit of the ratio between the measurements of the galaxy and of the matter power spectrum. This will reduce the effect of cosmic variance on the linear bias estimate (the cross galaxy-matter power spectrum is unfortunately not available). At large scales, we can assume a simple two-parameter, linear model given by

$$P_{\rm gg}(k) = b_1^2\, P_{\rm mm}(k) + \frac{1+\alpha_{P,1}}{\bar{n}}, \tag{45}$$

to be fit on scales $k_{\rm max} < 0.08\,h\,{\rm Mpc}^{-1}$. As a reference, the marginalised mean posterior values of both $b_1$ and $\alpha_{P,1}$ are listed in Table 4.

In Fig. 5 we show the marginalised constraints obtained fitting the full model of Eq. (34) to the eight data vectors, against the fiducial values of $\{b_1, \alpha_{P,1}\}$ obtained from the large-scale

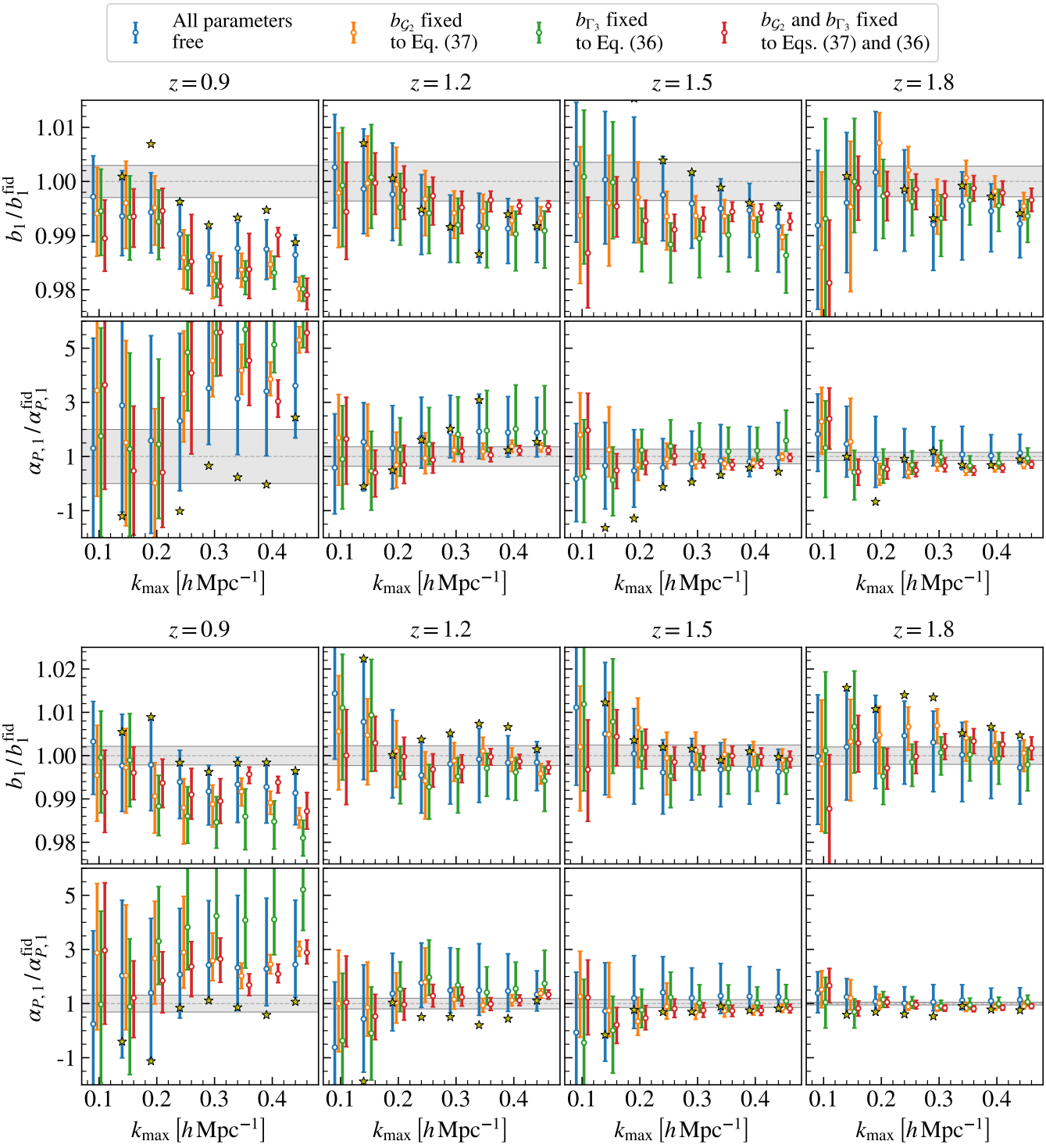


**Fig. 5.** Comparison between the marginalised constraints on the linear bias parameter $b_1$ and the shot-noise parameter $\alpha_{P,1}$ obtained at fixed cosmology, and the fiducial values listed in Table 4 obtained using only the large scale-limit of Eq. (45). The first two and last two rows show results for the Model 1 and Model 3 HOD samples, respectively. In both cases, the upper panels show constraints on the linear bias $b_1$, while the bottom ones show constraints on the constant shot-noise parameter $\alpha_{P,1}$. Different colours correspond to different assumptions on the total number of free bias parameters, as shown in the legend. Star symbols highlight the position of the maximum-likelihood for the case with all bias parameters free to vary. Dashed grey lines and shaded bands mark the fiducial value and $1\sigma$ confidence interval from Table 4.

limit as in Eq. (45). We test four different model configurations, which differ by the total number of bias parameters that are kept fixed to the relations presented in Sect. 3.1.3.

(i) All nuisance parameters are left free to vary while sampling the posterior distribution, for a total of six free parameters – linear bias $b_1$, local quadratic bias $b_2$, non-local quadratic bias $b_{\mathcal{G}_2}$, non-local cubic bias $b_{\Gamma_3}$, matter counterterm and higher-derivative bias $c_0$, and constant shot-noise parameter $\alpha_{P,1}$.

(ii) $b_{\mathcal{G}_2}$ is fixed to the excursion-set-based relation defined in Eq. (37), for a total number of five parameters.

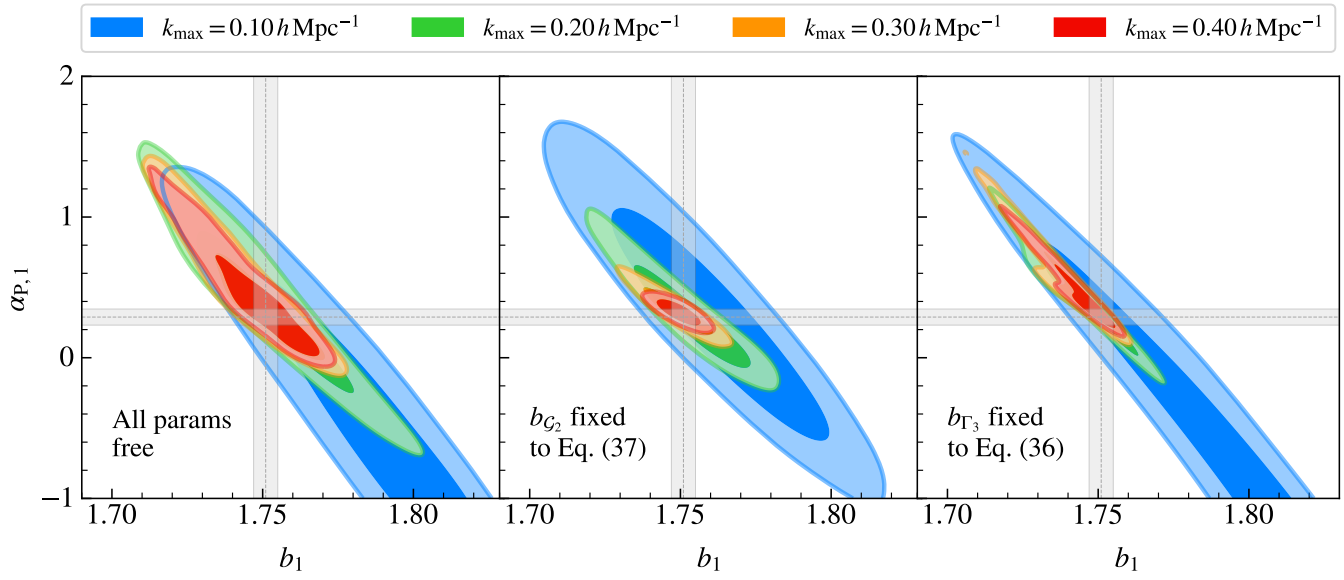


**Fig. 6.** Marginalised 2D constraints on the pair $(b_1, \alpha_{P,1})$ for the case corresponding to the Model 3 HOD sample at $z = 1.2$. The three different panels correspond to the case in which all the parameters are left free (left panel), $b_{\mathcal{G}_2}$ is fixed to the excursion-set relation as in Eq. (37) (middle panel), and $b_{\Gamma_3}$ is fixed to the coevolution relation as in Eq. (36) (right panel). Different colours represent different $k_{max}$ values, as listed in the legend. Fiducial 68% confidence intervals for both $b_1$ and $\alpha_{P,1}$ are shown with grey bands.

**Table 4.** Marginalised mean values of the linear bias $b_1$ and the shot-noise parameter $\alpha_{P,1}$ measured using the two-parameter model for the ratio $P_{gg}/P_{mm}$ presented in Eq. (45).

| Redshift | HOD model | $b_1$ | $\alpha_{P,1}\left[\frac{1}{\bar{n}}\right]$ |
|---|---|---|---|
| $z = 0.9$ | 1 | $1.350 \pm 0.004$ | $0.220 \pm 0.220$ |
| | 3 | $1.395 \pm 0.003$ | $0.253 \pm 0.079$ |
| $z = 1.2$ | 1 | $1.661 \pm 0.006$ | $0.424 \pm 0.152$ |
| | 3 | $1.751 \pm 0.004$ | $0.289 \pm 0.057$ |
| $z = 1.5$ | 1 | $1.977 \pm 0.007$ | $0.386 \pm 0.104$ |
| | 3 | $2.030 \pm 0.005$ | $0.219 \pm 0.032$ |
| $z = 1.8$ | 1 | $2.474 \pm 0.007$ | $0.257 \pm 0.039$ |
| | 3 | $2.486 \pm 0.005$ | $0.346 \pm 0.018$ |

**Notes.** Fits are carried out only considering scales up to $k_{max} = 0.08\,h\,Mpc^{-1}$.

(iii) $b_{\Gamma_3}$ is fixed to the coevolution relation defined in Eq. (36), for a total number of five parameters.
(iv) Both $b_{\mathcal{G}_2}$ and $b_{\Gamma_3}$ are fixed to the relations assumed in (ii) and (iii), respectively, for a total number of four parameters.

In all these cases, we keep the scale-dependent shot-noise parameter $\alpha_{P,2}$ fixed to zero (we test the validity of this assumption in Sect. 5.3).

Overall, the configuration with the largest number of free parameters – case (i) in the previous list, shown with blue points and error bars in Fig. 5 – is capable of capturing the correct amplitude of both $b_1$ and $\alpha_{P,1}$ for the majority of the tested $k_{max}$ values and redshifts, showing a mild running of the one-dimensional marginalised values that becomes relevant only for the lowest redshift snapshot we consider, on scales $k_{max} > 0.2\,h\,Mpc^{-1}$. The same effect is partially present for the $z = 1.2$ snapshot, although less significant: as a matter of fact, the marginalised constraints are consistent with their fiducial values at better than $2\sigma$. Rather than only considering the mean posterior distribution, it is instructive to also plot the maximum-likelihood point in the parameter space under consideration. We estimate this quantity using as a proxy the point in the sampled posterior distribution that maximises the likelihood, even though the latter is partially affected by a certain degree of stochasticity. In this case (star symbols in Fig. 5) we observe a shift towards the fiducial values, even if not for all configurations. A discrepancy between the maximum-likelihood point and the mean of the marginalised posterior is a clear hint at the presence of projection effects, also known as prior volume effects, due to the high dimensionality of the parameter space and to non-linear degeneracies among the model parameters.

Fixing either $b_{\mathcal{G}_2}$ – case (ii), orange points and error bars – or $b_{\Gamma_3}$ – case (iii), green points and error bars – does not significantly help in terms of accuracy of the marginalised constraints, with systematic deviations that can still become larger than the $1\sigma$ confidence interval. However we find that, while fixing $b_{\Gamma_3}$ typically results in similar constraining power on both $b_1$ and $\alpha_{P,1}$, imposing a relation on $b_{\mathcal{G}_2}$ leads to definitely tighter posteriors. This is the result of breaking the strong degeneracy between the two non-local bias parameters, $b_{\mathcal{G}_2}$ and $b_{\Gamma_3}$, and at the same time the one between the quadratic biases, $b_2$ and $b_{\mathcal{G}_2}$, leaving the remaining parameters to be more tightly constrained (a clear example of these is displayed in the right panel of Fig. F.1). The same clearly happens when combining the two previous relations – case (iv), red points and error bars – since with this setup we completely break the degeneracies in the considered parameter space. However, in this case we observe a deviation from the fiducial values of $b_1$ which can reach more than $2\sigma$ for some of the configurations, in particular at low redshift, hinting at a departure from the assumption of conserved evolution.

The effect of the strong $b_2$-$b_{\mathcal{G}_2}$-$b_{\Gamma_3}$ degeneracy can be observed in a more direct way by inspection of the 2d marginalised constraints in the $b_1$-$\alpha_{P,1}$ subspace. In Fig. 6 we show such posterior distributions, taking as a reference the Model 3 sample at $z = 1.2$. The different panels correspond

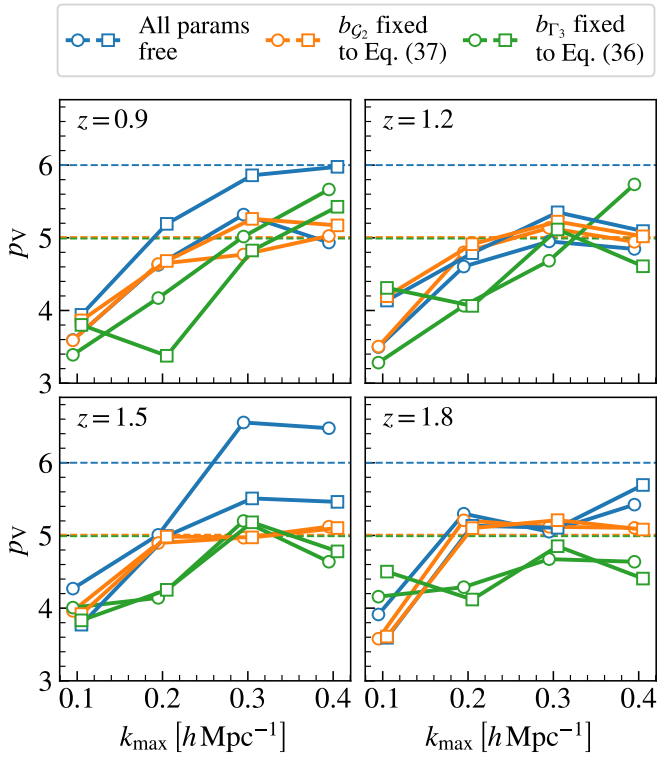


**Fig. 7.** Number of effective parameters that can be properly constrained by the input data vectors of the Model 1 (circles) and 3 (squares) HOD samples, respectively, as a function of the maximum wave mode included in the analysis and for different configurations of the bias model, as listed in the legend. Different panels correspond to different redshifts, as shown in the corresponding top left corner.

to different bias relations, from left to right: the case with all the parameters free to vary, with $b_{\mathcal{G}_2}$ fixed to the excursion-set-based relation, and with $b_{\Gamma_3}$ fixed to the coevolution relation. A first consideration to make is that there is a non-trivial degeneracy between the two parameters, for which projection effects might bias the 1d constraints without necessarily meaning that the hyper-dimensional posterior distribution does not cover the fiducial values of the parameters. Secondly, we can observe how the case with fixed $b_{\mathcal{G}_2}$ gives the tightest constraints for both parameters, with an increase in the merit of the constraints that is directly related to the maximum scale adopted in the fit, up to $k_{\rm max} = 0.4\,h\,{\rm Mpc}^{-1}$. Once more, this trend can be easily explained by the effective breaking of the degeneracy among the higher-order bias parameters.

For a limited number of cases, even when fixing one or more degrees of freedom, we find that the final posterior distribution can still appear multi-modal, leading to enlarged constraints when marginalising over the remaining parameters. For this reason, some of the chains where both tidal bias parameters $b_{\mathcal{G}_2}$ and $b_{\Gamma_3}$ are fixed feature marginalised constraints that are larger than the ones with one additional degree of freedom. This is clearly noticeable for the largest $k_{\rm max}$ bin of the Model 3 sample at $z = 0.9$.

The self-consistency of the different models in terms of the number of model parameters can be assessed using an additional statistics. In this context, we are interested in determining the total number of parameters that can be effectively constrained by the data vectors. A commonly employed statistics is represented by the $p_{\rm V}$ value, defined as (Gelman et al. 2014)

$$p_{\rm V} = \frac{1}{2}\left\langle \left(\chi^2 - \left\langle \chi^2 \right\rangle\right)^2 \right\rangle, \tag{46}$$

that is, the variance of the corresponding $\chi^2$ distribution. This number indirectly tracks the presence of degeneracies among the model parameters, and only converges to the total number of free parameters for a normal distribution. In order for a theory model to effectively constrain a given number of parameters, the $p_V$ value is expected to reach that same value, and can therefore be used as a proxy for the self-consistency of different model configurations. In Fig. 7 we show this value as measured from both sets of HOD samples and for different values of $k_{\rm max}$. In practice, we observe that the model with all parameters free never reaches the expected value of $p_V = 6$, even for the largest value of $k_{\rm max}$, with the exception of a couple of configurations. This shows that a six-parameter model is most likely resulting in overfitting. On the contrary, fixing one of the two tidal biases makes the $p_V$ reach the expected limit above some $k_{\rm max}$, with a transition that typically happens sooner for the case with fixed $b_{\mathcal{G}_2}$. This reinforces the conclusion that this configuration is preferred with respect to the others under consideration.

We note that these results may be partially affected by the presence of cosmic variance in the data vectors. For this reason, in Appendix G we explicitly assess the impact of this extra contribution, using both a smooth and a noisy realization of the data, generated using the theory code. This test shows that most of the residuals observed in Fig. 5 can be explained by sample variance affecting our data vectors.

### 5.3. Constraints on scale-dependent shot noise

So far, the stochastic field $\varepsilon_{\rm g}$ entering the expression for the galaxy density field in Eq. (20) has been assumed responsibile only for a constant offset from Poissonian predictions, via the parameter $\alpha_{P,1}$. An immediate check on the performance of the one-loop galaxy bias expansion can be carried out by further extending the model parameter space to also include the next-to-leading order correction to the stochastic field $\varepsilon_{\rm g}$. As already mentioned in Sect. 3, this leads to the presence of an additional $k^2$-dependent term in the galaxy power spectrum, whose amplitude is regulated by the extra parameter $\alpha_{P,2}$.

Figure 8 shows the marginalised one-dimensional constraints on $\alpha_{P,2}$, for both HOD samples, Model 1 on the left and Model 3 on the right, respectively. Since the large-scale limit of the galaxy power spectrum does not have enough constraining power on $\alpha_{P,2}$, we only consider values of $k_{\rm max}$ above $0.35\,h\,{\rm Mpc}^{-1}$ [26]. We never observe a statistically significant detection of the $\alpha_{P,2}$ parameter, with the majority of the marginalised constraints being consistent with $\alpha_{P,2} = 0$ well within the $2\sigma$ confidence interval. The only configurations for which this does not happen are the ones at high redshifts, specifically when considering high values of $k_{\rm max}$, since these are the configurations for which the parameter $\alpha_{P,2}$ is constrained with the highest precision. Performing the same test with one of the tidal bias fixed[27] does not lead to significantly different conclusions. This seems to suggest that the $\alpha_{P,2}$ parameter might have a more important role over a

[26] For some of the samples, this range of scales is already dominated by the Poissonian shot-noise contribution, as can be observed from the top and middle panel of Fig. 2.

[27] We choose to fix $b_{\mathcal{G}_2}$, motivated by the results of Sect. 5.2 that suggested this is the configuration less affected by degeneracies between the model parameters.

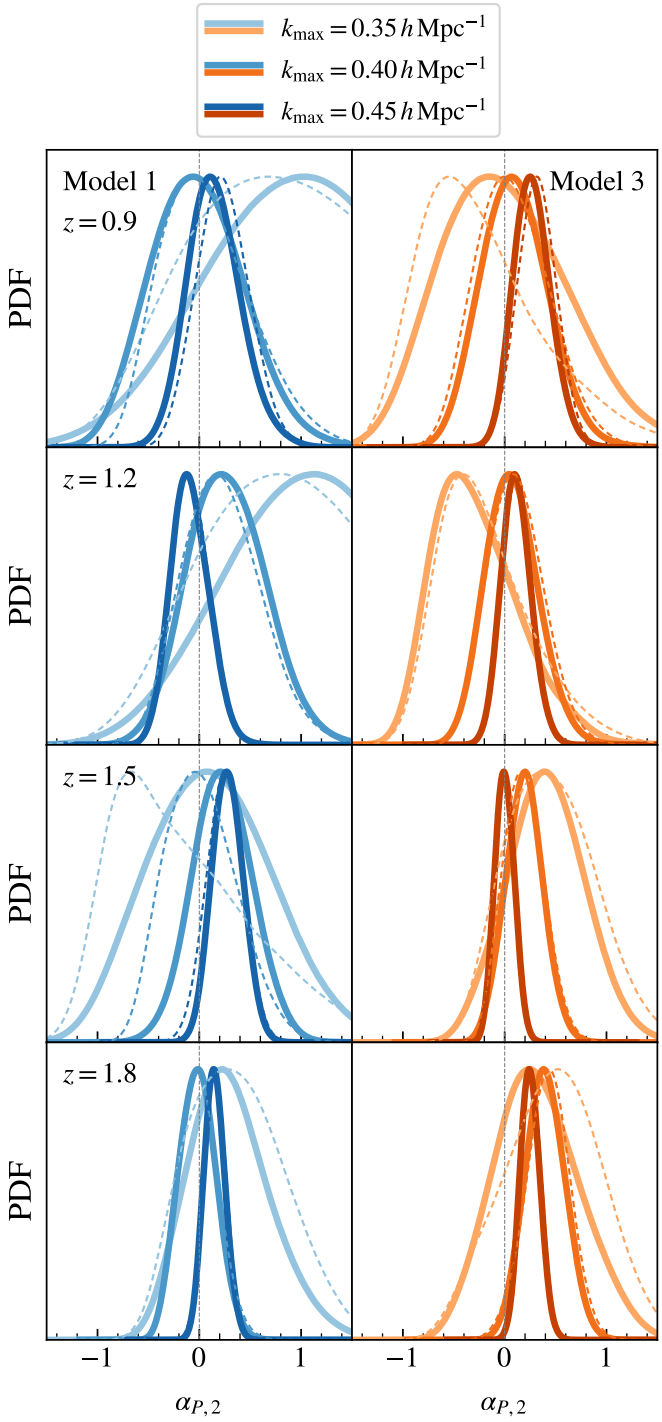


**Fig. 8.** Marginalised 1D constraints on the scale-dependent shot-noise parameter $\alpha_{P,2}$ in the fits with fixed cosmological parameters. Different rows correspond to different redshifts (top to bottom, low to high redshift), while different columns correspond to different HOD samples. Within each panel, the colour gradient marks different values of $k_{\rm max}$, as detailed in the legend. Solid/dashed lines correspond to the configurations with all the nuisance parameters free to vary, and with $b_{\mathcal{G}_2}$ fixed to the excursion-set relation as in Eq. (37), respectively.

range of redshifts where the shot-noise correction is more relevant.

We postpone further tests to a next installment of this series of papers, since a more careful check should be carried out adopting samples with different number densities – as this value determines the range of scales where the transition from signal to noise takes place – in particular considering values that would represent in a more reliable way the expected H$\alpha$ galaxy distribution detected by *Euclid*. At the same time, an important test should be carried out using the redshift-space galaxy power spectrum (Euclid Collaboration: Camacho et al., in prep.), for which extra $k^2$-dependent noise corrections are required, as a function of the orientation with respect to the line of sight (Philcox & Ivanov 2022; Carrilho et al. 2023; Moretti et al. 2023). Finally, we note that the findings of this analysis are in line with the conclusions from Pezzotta et al. (2021), for which, in terms of constraints on the cosmological parameters $\{h, \omega_{\rm c}\}$, a clear detection of scale-dependent stochastic parameters happens only when considering the combined information from the galaxy-galaxy and galaxy-matter power spectra.

### 5.4. Testing the EFT model: results on cosmological parameters

After having investigated the performance of the EFT model, we now turn our attention to the study of how cosmological constraints can be affected by different choices of model configuration. Specifically, we assume the same parameter space already used in Sect. 5.1, which also includes the Hubble parameter $h$ and the cold dark matter density parameter $\omega_{\rm c}$, while keeping the rest of the cosmological parameters fixed to their fiducial values, as listed in Table 1. A standard full-shape analysis of the redshift-space galaxy power spectrum would typically also include the scalar amplitude of the power spectrum, $A_{\rm s}$, since the anisotropies introduced by peculiar velocities make possible to break the strong $A_{\rm s}$-$b_1$ degeneracy that is otherwise present when considering real-space coordinates[28]. However, since this analysis revolves around the real-space galaxy power spectrum, we ought to choose a more conservative approach in order to obtain as least degenerate constraints as possible on the rest of the cosmological parameters. The sampling of $A_{\rm s}$ will be performed in the rest of the papers of this series, when considering the additional information content of the galaxy bispectrum (Euclid Collaboration: Eggemeier et al., in prep.) and RSDs (Euclid Collaboration: Camacho et al., in prep., Euclid Collaboration: Pardede et al., in prep.).

Figures 9 and 10 show the three performance metrics defined in Sect. 4.1 for the Model 1 and Model 3 HOD samples respectively, assuming the full-box volume of the Flagship I simulation. In both cases, the FoB and FoM panels refer to the combination between the two cosmological parameters we are sampling over.

#### 5.4.1. Figure of bias

In terms of FoB, we observe a consistent trend across each model configuration, indicating an unbiased combined measurement of the cosmological parameters even well within the mildly nonlinear regime, at $k_{\rm max} \gtrsim 0.3\,h\,{\rm Mpc}^{-1}$. The only exception is represented by the configuration in which both tidal bias parameters, $b_{\mathcal{G}_2}$ and $b_{\Gamma_3}$, are simultaneously kept fixed to the excursion-set relation (Eq. (37)) and to the coevolution relation (Eq. (36)), respectively. This outcome is unsurprising, as we are reducing by two the total number of degrees of freedom of the model. Notably, the bias on the cosmological parameters gets larger

[28] Specifically, the $A_{\rm s}$-$b_1$ degeneracy can be broken thanks to the different impact that these two parameters have on the amplitude of the leading-order power spectrum multipoles, $P^{(\ell)} \propto A_{\rm s}\, b_1^{2-\ell/2}$.

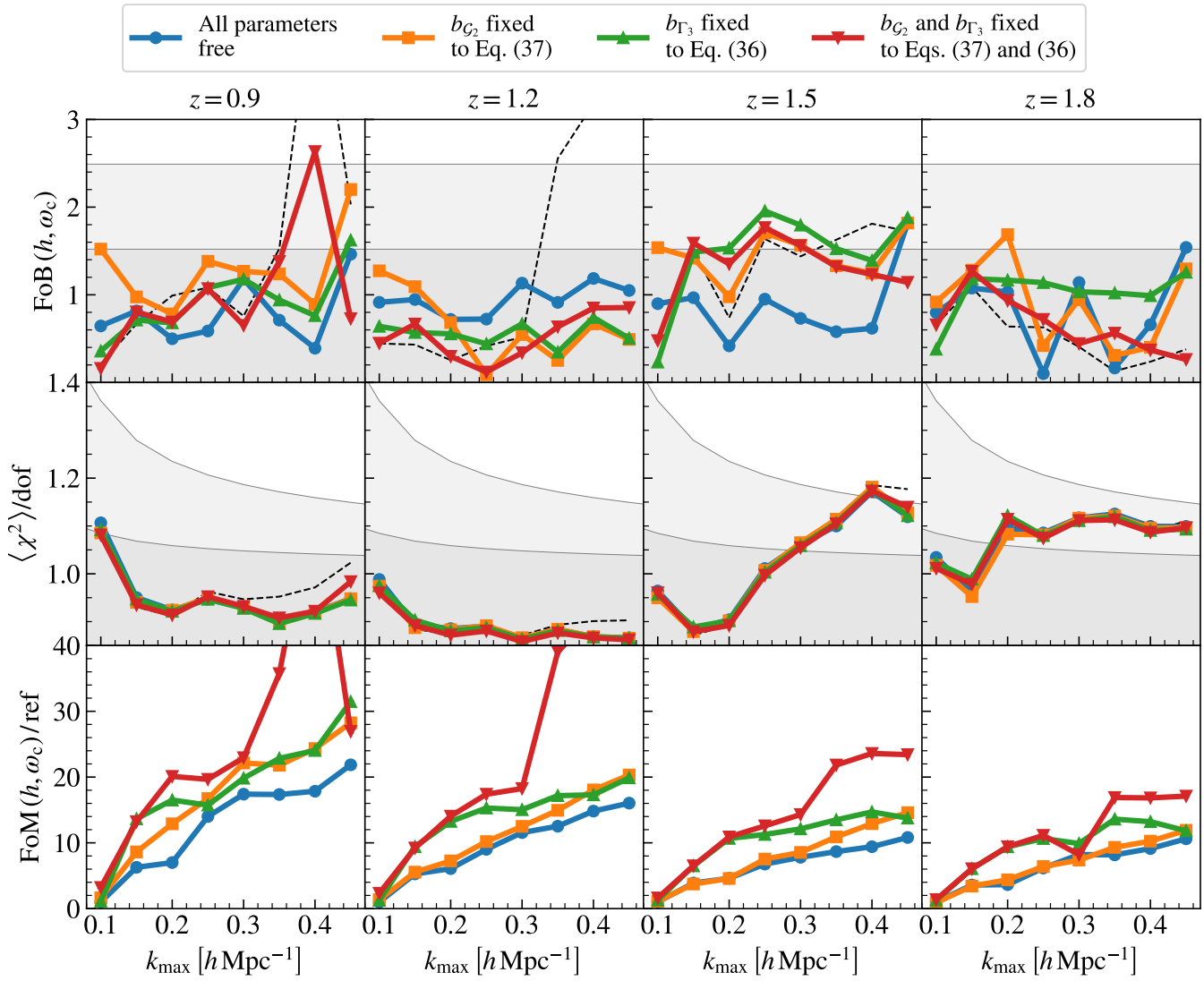


**Fig. 9.** Performance metrics of the Model 1 HOD samples for the various configurations defined in Sect. 5.2 as a function of the maximum wave mode $k_{max}$, and for the four different redshifts of the samples. The metrics shown are figure of bias (top), goodness of fit (middle), and figure of merit (bottom). Different colours correspond to different model configurations, as listed in the legend. The black dashed line shows as a reference the case in which both tidal bias parameters, $b_{\mathcal{G}_2}$ and $b_{\Gamma_3}$, are set to 0. The grey bands in the FoB and $\chi^2$ panels represent the 68% and 95% percentiles of the corresponding distributions. The FoM panels show the figure of merit normalised to the one of the standard run – with all bias parameters free to vary – at $k_{max} = 0.1\,h\,\mathrm{Mpc}^{-1}$.

at lower redshift, hinting to a departure from the coevolution relations as non-linear gravitational effects become more pronounced. Typically, this deviation occurs at scales of approximately $k_{max} = 0.3\,h\,\mathrm{Mpc}^{-1}$. However, we observe that this configuration behaves surprisingly well for high-redshift snapshots, exhibiting a FoB well within the 68% confidence interval.

In terms of overall stability of the results, we observe a deviation at low $k_{max}$ values for the Model 3 sample. This is more strongly affecting high-redshift snapshots, for which the value of the FoB at $k_{max} \sim 0.1\,h\,\mathrm{Mpc}^{-1}$ already exceeds the corresponding 68% confidence level, and only gets below the threshold when including additional signals from smaller scales. This effect is primarily attributed to the presence of projection effects, owing to the large dimensionality of the selected parameter space. Specifically, we find that all samples display a non-negligible correlation between the cold dark matter density parameter, $\omega_c$, and the EFT counterterm $c_0$, resulting in a systematic shift of $\omega_c$ for the lowest values of $k_{max}$, where there is insufficient constraining power to accurately constrain $c_0$. Fixing the cubic tidal bias $b_{\Gamma_3}$ to the coevolution relation typically helps to restore the cosmological parameters to their fiducial positions. This happens due the further degeneracy between $b_{\Gamma_3}$ and $c_0$ over the mildly non-linear regime. The relative importance of the $\omega_c$-$c_0$ degeneracy gets amplified only when considering the snapshots at the highest redshifts. As a partial confirmation of this trend, Pezzotta et al. (2021) did not report either any low-$k$ systematic effect when analysing mock galaxies meant to reproduce the clustering properties of the BOSS – CMASS and LOWZ – and SDSS MGS samples, since, in that case, the considered redshift range was much lower ($0.1 \lesssim z \lesssim 0.6$) than the one analysed in this work. As a further cross-check, the same effect is partially present when combining the full shape of the galaxy power spectrum and bispectrum in a joint analysis (Euclid Collaboration: Eggemeier et al., in prep.), albeit with a lower significance, due to the additional constraining power of higher-order statistics.

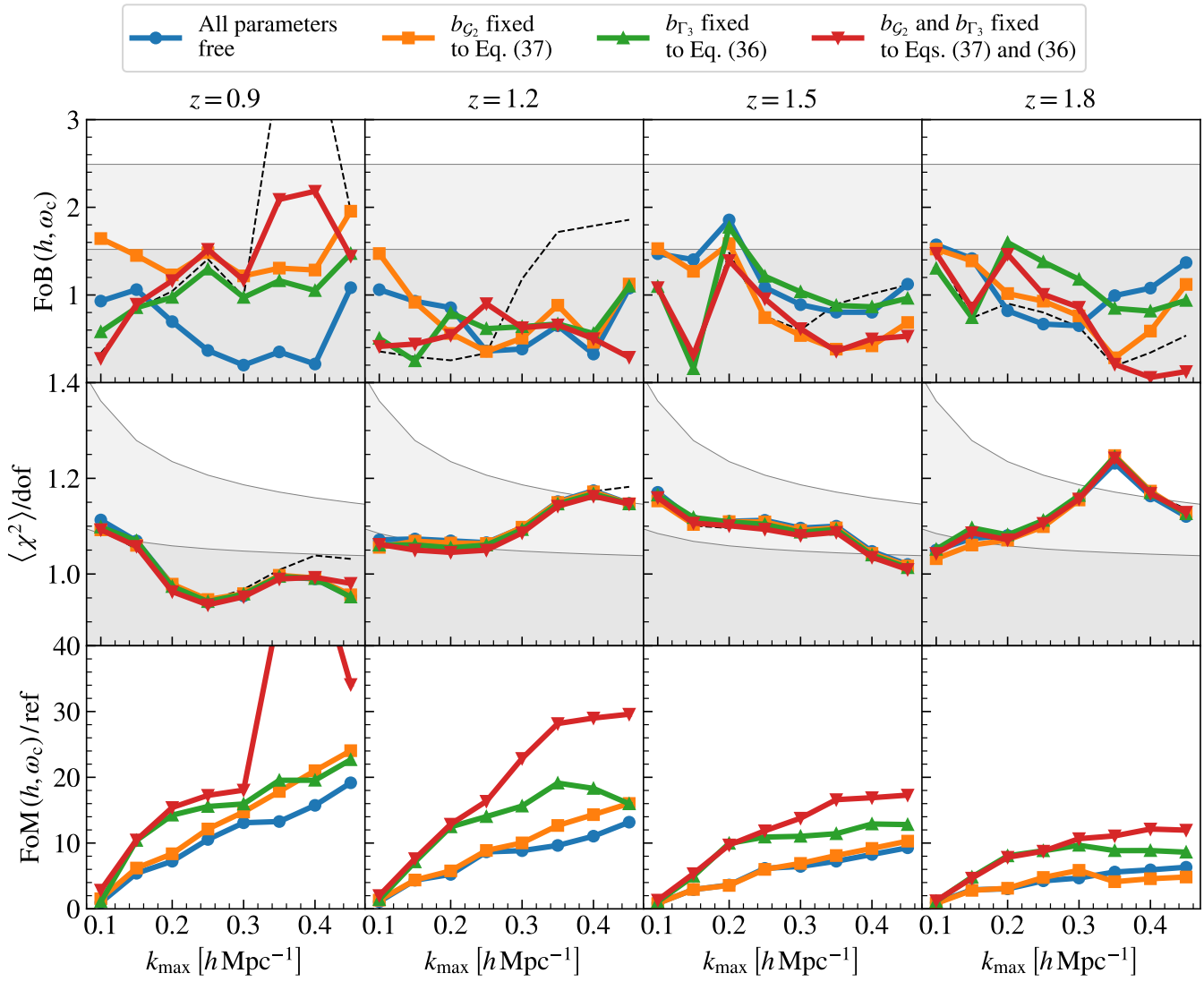


**Fig. 10.** Same as in Fig. 9, but for the Model 3 HOD samples.

#### 5.4.2. Goodness of fit

Similarly to the case at fixed cosmology, we find that the goodness of fit for the different models is consistent among the various model configurations, with only a small departure of the case where both tidal biases, $b_{\mathcal{G}_2}$ and $b_{\Gamma_3}$, are kept fixed to their corresponding relations and for the largest $k_{\rm max}$ values we consider in this analysis. This is visible in the case of the Model 1 sample at $z = 0.9$, for which there is an increase in the averaged $\chi^2$ value at $k_{\rm max} \gtrsim 0.3\,h\,{\rm Mpc}^{-1}$, which also corresponds to the transition of the FoB to values above the 68% percentile value. Otherwise, we find that the different model configurations provide a systematically consistent goodness of fit, with a reduced average $\chi^2$ value that is typically well within the 95% percentile of the corresponding $\chi^2$ distribution.

When considering the two different HOD models, the $\chi^2$ for the Model 3 HOD samples is consistently larger than for the Model 1 case (see the middle panel of Fig. 10). The most significant deviation is affecting the high-redshift snapshots, for which the average $\chi^2$ spuriously gets larger than the 95% confidence interval for some of the selected $k_{\rm max}$ values. In practice, this deviation is still consistent to better than the $3\sigma$ confidence interval. In addition, we remind that here we are using a single noisy realization, meaning sample variance could partially be driving some of the constraints. Moreover, we are analysing the data vectors with an extremely high level of precision, due to choice of using the full volume of the simulation and to the high number density of the HOD samples. As a further evidence for the goodness of our fits, in Fig. 11 we show the residuals between the maximum-likelihood theory vectors obtained at different values of $k_{\rm max}$ against the input data vectors, assuming the most relaxed model configuration. For all the samples that are under examination, we find that the broadband of the input galaxy power spectrum is perfectly recovered, and that the worst performance (in terms of goodness of fit) is only imputable to the scatter of the noisy data vector around the best fit – with a significance that is larger for some of the samples, such as for the Model 3 sample at $z = 1.8$.

Finally, in Figs. 9 and 10 we show with a dashed black line the FoB and averaged $\chi^2$ of the case where the values of the tidal biases are set to 0. In this case, we observe a departure of the goodness of fit from the other configurations, in a redshift-dependent way, which is also accompanied by a breaking of the model in terms of FoB. This shows that the use of coevolution relations can drastically improve the performance of the model, with respect to simply set the non-local bias parameters to zero.

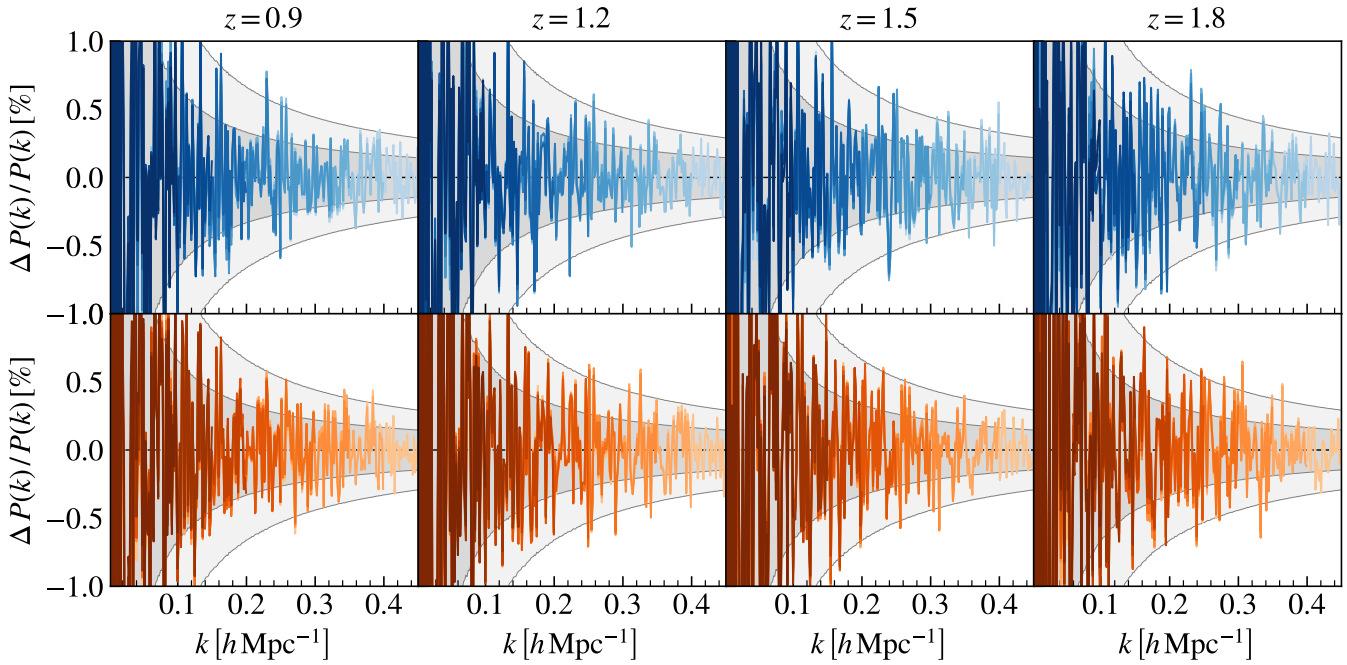


**Fig. 11.** Residuals of the maximum-likelihood best fits against the input galaxy power spectrum data vectors, assuming the case with all bias parameters free to vary. Different columns correspond to different redshifts, as shown on top of the corresponding column, while different rows mark either the Model 1 (top row) or Model 3 (bottom row) HOD sample. Different colour shades mark the best fits obtained at different $k_{\rm max}$ values, from $0.1\,h\,\mathrm{Mpc}^{-1}$ up to $0.45\,h\,\mathrm{Mpc}^{-1}$.

#### 5.4.3. Figure of merit

As expected, the FoM monotonically increases when including additional information from more non-linear scales, with a relative gain with respect to the most relaxed configuration (all bias parameters free to vary, at $k_{\rm max} = 0.1\,h\,\mathrm{Mpc}^{-1}$) that becomes larger moving towards lower redshifts. In fact, extending the fitting range to the maximal value of $k_{\rm max} = 0.45\,h\,\mathrm{Mpc}^{-1}$, we find that the trend for the FoM of the different bias configurations is approximately 10, 15, 20 and 30 times larger than the reference at $z = 1.8, 1.5, 1.2$ and 0.9, respectively. The only exception is represented by the high-redshift snapshots of the Model 3 sample, for which the total galaxy power spectrum on mildly non-linear scales becomes dominated by the shot-noise correction earlier than for the rest of the samples, and for which we observe that the FoM reaches a plateau at $k_{\rm max} \gtrsim 0.3\,h\,\mathrm{Mpc}^{-1}$. As expected, the case with two less degrees of freedom consistently gains more constraining power on the cosmological parameters, leading to much tighter constraints in particular at the largest $k_{\rm max}$ value we probe. However, we note that these gains are directly correlated with the breaking of the model in terms of FoB at $z = 0.9$[29], and might therefore lead to biased cosmological constraints if used in a real-data analysis. Nevertheless, for most of the tested cases, combining the two relations still leads to acceptable results up to the maximal scale we are considering.

[29] The breaking of this particular configuration, with both $b_{\mathcal{G}_2}$ and $b_{\Gamma_3}$ fixed to the bias relations, also exhibits a FoM at $k_{\rm max} = 0.45\,h\,\mathrm{Mpc}^{-1}$ that is lower than the one obtained at lower $k_{\rm max}$. We do not explore this configuration, since the model cannot be used with this configuration, but we argue that a more careful investigation of this effect should be carried out using a larger set of simulations, to reduce the importance of cosmic variance, which might partially drive these effects.

#### 5.4.4. Summary

Overall, we find that fixing only the quadratic tidal bias $b_{\mathcal{G}_2}$ leads to the most stable results, with a FoB that is typically – except for some spurious scale cut – well within the 68% percentile of the corresponding distribution, and with a FoM which is systematically larger than in the case where all the parameters are free to vary. The performance of the case with a fixed cubic tidal bias is also consistent, but with the caveat that the underlying bias parameters experience a strong degeneracy among themselves, as shown in Sect. 5.2. However, we find that this case typically achieves a FoM larger than the one with fixed quadratic tidal bias, with the latter catching up only at large enough values of $k_{\rm max}$. Also, in a range of scales up to $k_{\rm max} \sim 0.3\,h\,\mathrm{Mpc}^{-1}$, the case with relations applied to both tidal biases matches almost identically the case with fixed cubic bias, highlighting again how this parameter has a much larger impact when constraining the cosmological parameters considered in this analysis.

Nevertheless, as already mentioned in Sec. 3.1.3, we argue that the applicability of the excursion set and coevolution relations may partially fail when adopting them with real data, given the lack of any realistic dependency on other quantities different from the halo mass, such as assembly bias (Croton et al. 2007; Barreira et al. 2021; Hadzhiyska et al. 2021b; Lazeyras et al. 2023) or the scatter in the $\langle N \rangle (M_{\rm h})$ relation (Behroozi et al. 2010; Zehavi et al. 2011). For this reason, extended studies on more realistic simulated samples will be needed to properly benchmark these relations when applying them to real *Euclid* observations.

As a final remark, in Nicola et al. (2024) the authors carried out a comparison similar to the one presented in this paper, by forecasting the impact of different bias models in the analysis of the *Rubin* Observatory Legacy Survey of Space and Time (LSST; Ivezić et al. 2019) photometric observations. Differently from this work, their analysis showed that the use of minimal bias models with 1/2 less degrees of freedom could bias the recovery

of cosmological parameters at more than $3\sigma$ also when restricting the analysis to relatively large scales ($k_{\rm max} \sim 0.15\,h\,{\rm Mpc}^{-1}$). We argue that this apparent inconsistency is mainly imputable to the different observables selected in the two analyses. While in this work we make use of the real-space galaxy power spectrum at some fixed redshifts $z \gtrsim 1$, their analysis is focused on the $3\times2$-point data combination in tomographic bins that are for the most part at lower redshifts ($0 < z < 1$), for which the assumption of a negligible tidal or higher-derivative bias is no longer valid. As a further evidence, the dashed black lines in Figs. 9 and 10 show how setting both tidal biases to 0 makes the Eulerian bias model break sooner when considering the snapshots at lower redshifts, especially the one at $z = 0.9$.

### 5.5. Dependence on sample volume

As anticipated in Sect. 2.3, in addition to performing a model selection using extremely high-precision measurements – with a Gaussian covariance matrix derived assuming the full box volume of the Flagship I simulation – in this section we test the performance of the Eulerian bias expansion for smaller values of the galaxy sample volume, in order to provide more realistic forecasts for the analysis of the *Euclid* spectroscopic data. Once again, we remind the reader that for these tests we consider HOD catalogs with a number density larger than the one expected for the real galaxy samples, and that more realistic mocks will be used in forthcoming analyses that will also consider observational systematic effects, such as target purity and incompleteness, and observational effects such as the radial and angular selection function.

In this section, we consider the same data vectors obtained from measurements of the Flagship simulation snapshots already used in the previous ones. The dependence on the volume is explored by rescaling the corresponding covariance. We consider four different volumes corresponding to possible *Euclid*-like shells, as explained in Sect. 2.3. For each comoving snapshot, at $z = (0.9, 1.2, 1.5, 1.8)$, with reference volume $V_{\rm box}$, we define the volume $V_{\rm shell}$ of a spectroscopic bin corresponding to a total angular surface of 15 000 square degrees – and with a depth of $\Delta z = (0.2, 0.2, 0.2, 0.3)$ – following the choices made in Euclid Collaboration (2020). The three additional volume rescalings are obtained by selecting the values that divide the interval $[\eta, 1]$ into four equi-partitioned subintervals, where $\eta = V_{\rm box}/V_{\rm shell}$.

In Fig. 12 we show the trends of the FoB and FoM for the previously defined samples, with Model 1 and 3 in the top and bottom two rows, respectively. Thick solid lines correspond to the case with $b_{\Gamma_3}$ fixed to the coevolution relation, which we selected as one of the best performing model among the ones that we have tested in Sect. 5.4. On the contrary, thin dashed lines correspond to the case with all the bias parameters free to vary. Different colours identify the different covariance matrices used in the fits, from the already shown full-box case – in light blue/orange – to the case corresponding to the *Euclid*-like shells – in dark blue/orange. As expected, we find that in both cases the FoB for the two cosmological parameters $(h, \omega_{\rm c})$ becomes progressively smaller, reflecting the increasing amplitude of the covariance matrices used in the fits of the input data vectors. At the same time, we observe how this trend is tightly correlated to a decrease in the FoM, with the severity of the drop being almost proportional to the factor between the original and the rescaled volume, $\eta$.

It comes with no surprise that the reference bias model is well-performing up to the highest value of $k_{\rm max}$ we consider, even under realistic assumptions. Additionally, for the case with fixed $b_{\Gamma_3}$, we can observe how the precision on the cosmological constraints reaches a plateau above a typical threshold that corresponds to the transition scale between the regimes dominated by the signal and by the shot-noise contribution, respectively. For this reason, it is possible to gain additional constraining power by pushing the analysis to high values of $k_{\rm max}$ at $z = 0.9$, for which the number density of the sample is significantly larger than the one of the high-redshift samples. In contrast, when all the parameters are free to vary, the trend for the FoM curves is to gain additional constraints from smaller scales, even above the scale of transition, possibly pointing to a further breaking of parameter degeneracies that are no longer present when fixing the value of $b_{\Gamma_3}$ to the coevolution relation.

As a final consideration, we note that the Eulerian bias expansion is performing significantly well for all the considered rescalings of the covariance matrix. The constraining power of the EFT model in terms of the combination $\{h, \omega_{\rm c}\}$ can be enhanced employing one of the coevolution relations described in Sect. 3.1.3 without the appearance of systematic errors, even when considering samples with a number density significantly larger than the one expected from the real *Euclid* data. This analysis, limited to real space, motivates further tests including higher-order statistics, such as the galaxy bispectrum, and taking into account redshift-space distortions. These topics will be properly explored in the next entries of this series.

## 6. Conclusions

In this paper we carried out an analysis meant to assess the performance of state-of-art models for one-loop galaxy bias over a redshift range that is well representative of the spectroscopic galaxy sample that will be one of the main targets of *Euclid*. We employed a set of four FoF halo catalogues from comoving snapshots of the Flagship I simulation at $z = (0.9, 1.2, 1.5, 1.8)$, which were subsequently populated with H$\alpha$ galaxies using HOD prescriptions based on the Model 1 and 3 from Pozzetti et al. (2016). Each snapshot features an outstanding volume of $(3780\,h^{-1}\,{\rm Mpc})^3$ and a high comoving number density (from $\sim10^{-4}\,h^3\,{\rm Mpc}^{-3}$ to $\sim10^{-3}\,h^3\,{\rm Mpc}^{-3}$), which corresponds to a flux limit of $f_{\rm H\alpha} = 2\times10^{-16}\,{\rm erg\,cm^{-2}\,s^{-1}}$. These snapshots can therefore be used to assess the accuracy of the perturbative bias expansion at a high level of precision.

We tested two galaxy bias models for the full shape of the real-space galaxy power spectrum. The first one adopts a Eulerian bias expansion, and is based on the recently developed EFTofLSS modelling, in which the impact of small-scale physics, as well as the integration of ultraviolet modes in SPT, can be captured by a set of counter terms, which reduce to a single one when considering real-space coordinates. The final parameter space consists of two cosmological parameters, the Hubble parameter $h$, and the cold dark matter density parameter $\omega_{\rm c}$, plus a set of six nuisance parameters, consisting of the linear bias $b_1$, the quadratic bias $b_2$, the tidal quadratic and cubic biases, $b_{\mathcal{G}_2}$ and $b_{\Gamma_3}$, the matter counter term and higher-derivative bias $c_0$, and two extra parameters representing deviations from Poissonian shot noise: a constant offset, $\alpha_{P,1}$, and a scale-dependent term, $\alpha_{P,2}$. The second model adopts a similar one-loop expansion of the galaxy power spectrum, but this time using Lagrangian coordinates, and is based on the emulation of the individual terms of the expansion starting from a limited set of high-resolution $N$-body simulations. This is achieved thanks to the cosmology-rescaling technique presented in Angulo et al. (2021). In addition to the two cosmological parameters, this

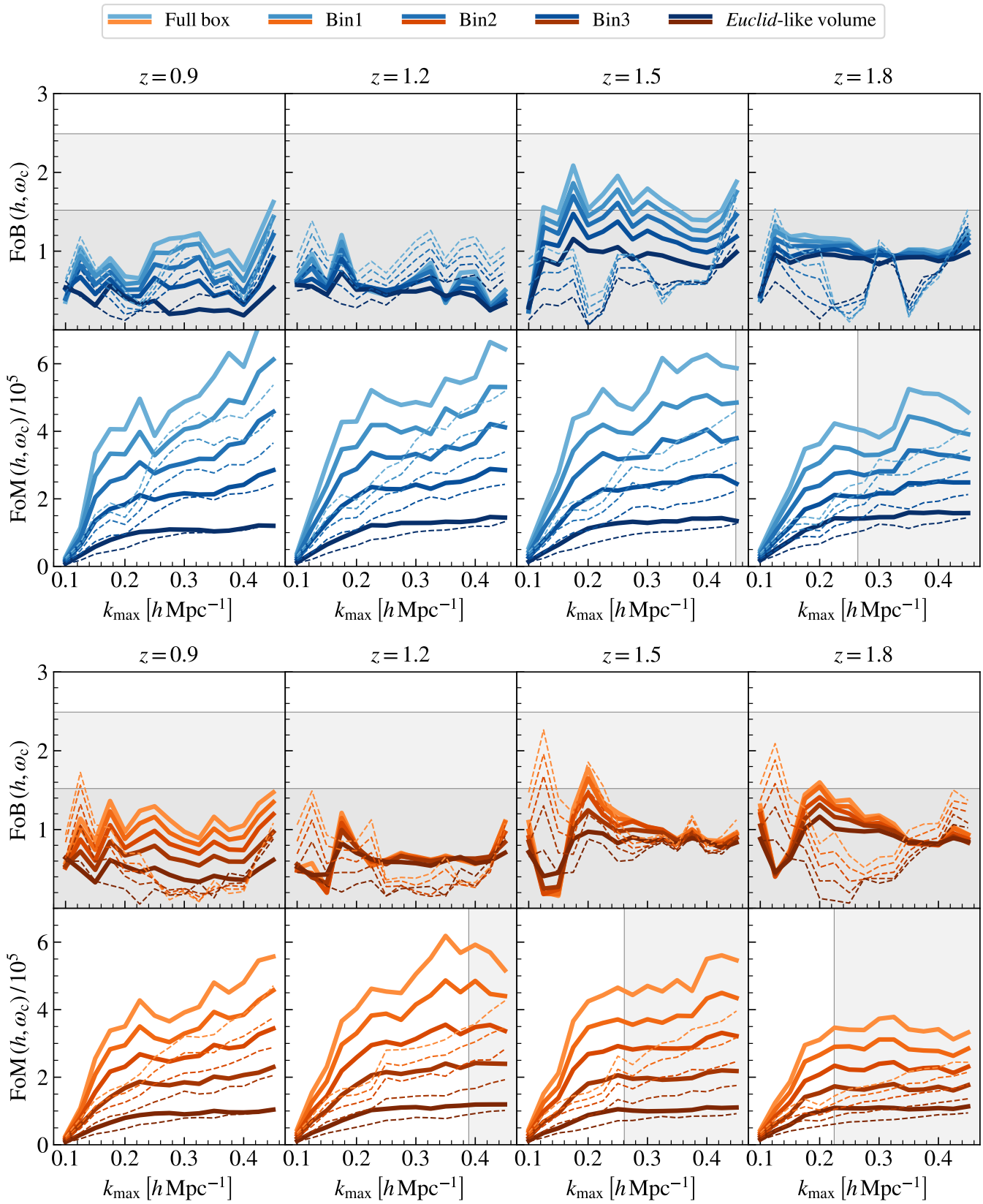


**Fig. 12.** Evolution of the FoB and FoM as a function of the different choice for the rescaling of the reference volume, as shown in the legend. Solid and dashed curves correspond to the case with $b_{\Gamma_3}$ fixed to the coevolution relation and to the case with all the parameters free to vary, respectively. All configurations of the top two rows correspond to the fits of the Model 1 HOD samples, while the bottom two rows do the same for the Model 3 HOD samples. Grey bands in the FoM panels identify the scale at which the Poisson shot-noise contribution assumes the same height of the underlying clustering signal, marking the transition to the shot-noise-dominated scales.

model features the linear bias $b_1$, the quadratic bias $b_2$, the tidal quadratic bias $b_{s^2}$, the Laplacian bias $b_{\nabla^2\delta}$, and a stochastic parameter representing a constant deviation from Poissonian shot noise, $\alpha_{P,1}$. In this work, we made use of the two implementations available in the public codes `COMET` (Eggemeier et al. 2022) for the EFT model – with accuracy tests carried out against external benchmarks, as shown in Appendix D – and `baccoemu` (Zennaro et al. 2023) for the hybrid model.

In the main section of this work, we present tests of the relative performance of these two galaxy bias models, while in the subsequent sections we show how we determined the range of validity of the EFT model and tested the impact of fixing one or more parameters of the Eulerian bias expansion to some physically motivated relations in a way that allows us to break strong parameter degeneracies and better constrain the cosmological parameters. In all cases, we determined the range of validity of a given bias relation and scale cut by means of three different performance metrics: the goodness of fit, the figure of bias, and the figure of merit. The latter two metrics are computed on the $\{h, \omega_c\}$ combination in order to quantify the accuracy and precision of the model in terms of these two parameters.

We compare the performance of the Eulerian and hybrid Lagrangian bias models using a rescaled covariance to match the size of *Euclid*-like redshift shells (assuming a full-sky area of $15\,000\,\mathrm{deg}^2$). Our results highlight how both models are capable of providing unbiased measurements of the cosmological parameters up to $k_{\rm max} = 0.45\,h\,\mathrm{Mpc}^{-1}$ for all four redshift snapshots, consistently with the $1\sigma$ confidence interval for the FoB distribution. In terms of FoM, `baccoemu` reaches the same amplitude of the maximally achievable FoM of the EFT case already at a lower $k_{\rm max}$ value, most likely due to the absence of the cubic tidal bias parameter that is instead free to vary in the EFT chains. As expected, when fixing this parameter, the EFT model performs similarly to `baccoemu`, and for most of the configurations it achieves a slightly larger FoM. When considering the covariance matrix corresponding to the full-box of the comoving snapshots (from three to seven times larger than the *Euclid*-like shells, depending on the considered redshift) we find that the EFT model manages to recover the cosmological parameters consistently with the $1\sigma$ confidence interval. However, there are some spurious cases at low values of $k_{\rm max}$ that are affected by projection effects, which can be alleviated by fixing the value of $b_{\Gamma_3}$. On the other hand, with this level of precision, we hit the intrinsic emulation error of `baccoemu`, which leads to bias in the inferred parameters when considering high $k_{\rm max}$ values at low redshift. Including an extra component to the covariance matrix – corresponding to 0.5% of the amplitude of the galaxy power spectrum, based on the combined systematic error from the emulation and from the measurements used to train the emulator – brings the FoB of `baccoemu` back within the $1\sigma$ confidence interval. Also in this case, on scales for which the intrinsic error of `baccoemu` is not the dominant contribution, that is, $k_{\rm max} \lesssim 0.2\,h\,\mathrm{Mpc}^{-1}$, its FoM is consistent with that of the EFT model. Not including the 0.5% extra error, we note that the FoM is consistent with – and in some cases even larger than – the corresponding EFT results, demonstrating the potential gain to be obtained by improving the accuracy of next-generation emulators of the full shape of the galaxy power spectrum.

We then focus exclusively on the EFT model: we first describe the fits we performed at fixed cosmology to check the self-consistency of the one-loop galaxy bias expansion in terms of the linear bias $b_1$ and the scale-independent shot-noise parameter $\alpha_{P,1}$. The fiducial values for these parameters were fitted from the measured galaxy-to-matter-power-spectrum ratio, assuming a leading-order recipe on scales of $k < 0.08\,h\,\mathrm{Mpc}^{-1}$. The result is that, when leaving all parameters free to vary, it is possible to recover – at better than $2\sigma$ – the value of both parameters for the majority of the samples – two HOD models times four different redshifts – and scale cuts up to $k_{\rm max} = 0.45\,h\,\mathrm{Mpc}^{-1}$. The only significant deviation takes place at the lowest redshift we consider, $z = 0.9$, for which we observe a departure from the fiducial values soon after $k_{\rm max} = 0.2\,h\,\mathrm{Mpc}^{-1}$. The latter is however consistent with sample variance expectations, as observed in a set of ten different noisy realisations of a synthetic theory data vector (see Appendix G). The systematic errors are partially alleviated when considering the position in parameter space corresponding to the maximum of the likelihood, showing how the deviations might be imputable to projection effects. Fixing one of the two tidal biases to either an excursion-set-derived relation or to the coevolution relation still results in constraints that are consistent with the $2\sigma$ confidence interval. In particular, the former is preferred in terms of constraints of the model parameters because of the simultaneous breaking of the strong degeneracies with both the quadratic bias $b_2$ and the cubic tidal bias $b_{\Gamma_3}$. We find that fixing both parameters at the same time still works extremely well (with a typical recovery of $b_1$ and $\alpha_{P,1}$ within the 68% confidence interval) even for the largest $k_{\rm max}$ values when considering the high-redshift snapshots. However, at low redshift, this choice can lead to deviations of more than $3\sigma$ for some of the configurations we test, especially when the maximum scale included in the fit is above a typical scale of $k_{\rm max} \sim 0.3\,h\,\mathrm{Mpc}^{-1}$.

We explicitly checked whether or not a next-to-leading-order correction to the shot-noise contribution ($\alpha_{P,2}$) can improve the model performance. In all cases considered, the marginalised posterior distribution for $\alpha_{P,2}$ is consistent with zero within $2\sigma$, suggesting this additional parameter is not needed, at least for the description of these galaxy samples. Additionally, the scales that can constrain this parameter soon become dominated by the underlying shot-noise correction, effectively breaking the perturbative description of the latter in a Taylor expansion. A more significant test should be carried out considering more realistic galaxy samples, in terms of galaxy number density, and also including RSDs and observational systematic effects.

When the parameter space is extended to also include the two cosmological parameters, we note a good recovery of the fiducial values across the whole range of separations we test. The FoB exhibits an increasing trend moving towards high redshifts and low values of $k_{\rm max}$, which is due to projection effects when marginalising over all the nuisance parameters. This trend can be partially corrected by fixing the cubic tidal bias $b_{\Gamma_3}$ to the coevolution relation, and indeed with this configuration it is possible to consistently recover unbiased – within the 68% confidence interval – constraints on the $(h, \omega_c)$ pair. Also in this case, fixing both tidal biases at the same time can lead to biased cosmological constraints, with the amplitude of the systematic errors increasing towards lower redshifts, which suggests a premature breaking of the tested relations. In terms of goodness of fit, we do not observe a significant change in the average $\chi^2$ when fixing some of the model parameters to the relations presented in Sect. 3.1.3. Finally, the FoM of the cosmological parameters clearly increases when reducing the degrees of freedom of the model. However, the configuration with both tidal biases fixed results in biased constraints for the low-redshift samples, while the case with only $b_{\Gamma_3}$ fixed can be employed down to smaller scales. This configuration displays a relative gain in FoM that ranges from 1.5 to 2 times, with respect to the case with all the parameters free. Relative gains in the FoM are more

concentrated at $k_{\rm max} \lesssim 0.3\,h\,{\rm Mpc}^{-1}$, where the model with fixed $b_{\Gamma_3}$ experiences a steep increase that is then followed by a more modest growth. The configuration with $b_{\mathcal{G}_2}$ fixed also displays a FoM larger than the case with all parameters free to vary, but with a steady slope that manages to catch up with the other configuration only for the largest $k_{\rm max}$ values that we tested. Overall, we find that the one-loop galaxy bias expansion is sufficiently accurate on the redshift range that we are exploring, $1 \lesssim z \lesssim 2$, even deep within the mildly non-linear regime, at $k_{\rm max} \sim 0.4\,h\,{\rm Mpc}^{-1}$, with a statistical significance on the cosmological parameters that can be enhanced by fixing some of the degrees of freedom of the model.

In order to understand the impact of a different statistical precision on the input data vectors, we rescaled the Gaussian covariance matrix used in the fitting procedure to match the volume of a *Euclid*-like spectroscopic bin, with three additional intermediate volume choices selected between the *Euclid*-like bin and the original volume of the comoving box. A smaller volume therefore corresponds to a reduced amplitude in the covariance matrix, resulting in a decrease in both FoB and FoM in a way that is proportional to the fraction of lost volume. We therefore confirm that these models of galaxy bias can be eventually used to analyse the real spectroscopic data collected by *Euclid*.

This paper stands as the first installment of a series of works meant to validate the theoretical framework that will be used to analyse the large-scale galaxy distribution as observed in the actual measurements of *Euclid*. Here we focus on the modelling of the real-space galaxy power spectrum of the spectroscopic sample, which stands as an important test for the complementary photometric analysis that is going to be carried out by *Euclid*, in the shape of the popular $3\times2$-point data combination. In parallel, in Euclid Collaboration: Eggemeier et al. (in prep.) we consider the joint analysis of the real-space galaxy power spectrum and bispectrum, exploring a consistent description of non-linear bias in both observables. Two additional papers in the series (Euclid Collaboration: Camacho et al., in prep., Euclid Collaboration: Pardede et al., in prep.) will extend the modelling tests to redshift space. In parallel, a different set of papers will be devoted to a similar analysis of configuration-space statistics (Euclid Collaboration: Guidi et al., in prep., Euclid Collaboration: Kärcher et al., in prep., Euclid Collaboration: Pugno et al., in prep.).

*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 Academy of Finland, the Agenzia Spaziale Italiana, the Belgian Science Policy, the Canadian Euclid, Consortium, the French Centre National d'Etudes Spatiales, the Deutsches Zentrum für Luft- und Raumfahrt, the Danish Space Research Institute, the Fundação para a Ciência e a Tecnologia, the Ministerio de Ciencia e Innovación, the National Aeronautics and Space Administration, the National Astronomical Observatory of Japan, the Netherlandse Onderzoekschool Voor Astronomie, the Norwegian Space Agency, the Romanian Space Agency, the State Secretariat for Education, Research and Innovation (SERI) at the Swiss Space Office (SSO), and the United Kingdom Space Agency. A complete and detailed list is available on the *Euclid* web site (`http://www.euclid-ec.org`). The majority of the analysis carried out in this manuscript has been produced by a joint effort among several *Euclid* members and work centers. A special acknowledgement goes to the Max Planck Computing and Data Facility (MPCDF) in Garching (Germany), where most of the results presented in this article have been obtained. We acknowledge the hospitality of the Institute for Fundamental Physics of the Universe (IFPU) of Trieste for the group meeting held there in November 2022. C.M.'s research for this project was supported by a UK Research and Innovation Future Leaders Fellowship [grant MR/S016066/2]. C.M.'s work is supported by the Fondazione ICSC, Spoke 3 Astrophysics and Cosmos Observations, National Recovery and Resilience Plan (Piano Nazionale di Ripresa e Resilienza, PNRR) Project ID CN_00000013 "Italian Research Center on High-Performance Computing, Big Data and Quantum Computing" funded by MUR Missione 4 Componente 2 Investimento 1.4: Potenziamento strutture di ricerca e creazione di "campioni nazionali di R&S (M4C2-19 )" – Next Generation EU (NGEU). A.E. is supported at the Argelander Institut für Astronomie by an Argelander Fellowship. A.B. acknowledges support from the Excellence Cluster ORIGINS which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC-2094-390783311. R.E.A. acknowledges the support of the ERC-StG number 716151 (BACCO), and the project PID2021-128338NB-I00 from the Spanish Ministry of Science. M.B. is supported by the Programma Nazionale della Ricerca (PNR) grant J95F21002830001 with title "FAIR-by-design". M.K. is funded by the Excellence Initiative of Aix-Marseille University – A*MIDEX, a French "Investissements d'Avenir" programme (AMX-19-IET-008 – IPhU). This research made use of `matplotlib`, a Python library for publication quality graphics (Hunter 2007).

## References

Abidi, M. M., & Baldauf, T. 2018, JCAP, 07, 029
Alam, S., Ata, M., Bailey, S., et al. 2017, MNRAS, 470, 2617
Angulo, R. E., & White, S. D. M. 2010, MNRAS, 405, 143
Angulo, R., Fasiello, M., Senatore, L., & Vlah, Z. 2015, JCAP, 09, 029
Angulo, R. E., Zennaro, M., Contreras, S., et al. 2021, MNRAS, 507, 5869
Assassi, V., Baumann, D., Green, D., & Zaldarriaga, M. 2014, JCAP, 08, 056
Bagley, M. B., Scarlata, C., Mehta, V., et al. 2020, ApJ, 897, 98
Baldauf, T., Seljak, U., & c. v., Desjacques, V., & McDonald, P., 2012, Phys. Rev. D, 86, 083540
Baldauf, T., Seljak, U., Smith, R. E., Hamaus, N., & Desjacques, V. 2013, Phys. Rev. D, 88, 083507
Baldauf, T., Mercolli, L., & Zaldarriaga, M. 2015a, Phys. Rev. D, 92, 123007
Baldauf, T., Mirbabayi, M., Simonović, M., & Zaldarriaga, M. 2015b, Phys. Rev. D, 92, 043514
Baldauf, T., Codis, S., Desjacques, V., & Pichon, C. 2016, MNRAS, 456, 3985
Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
Barreira, A., Lazeyras, T., & Schmidt, F. 2021, JCAP, 08, 029
Baumann, D., Nicolis, A., Senatore, L., & Zaldarriaga, M. 2012, JCAP, 07, 051
Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379
Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013, ApJ, 762, 109
Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, Phys. Rep., 367, 1
Bertolini, D., Schutz, K., Solon, M. P., Walsh, J. R., & Zurek, K. M. 2016, Phys. Rev. D, 93, 123505
Beutler, F., Blake, C., Colless, M., et al. 2012, MNRAS, 423, 3430
Beutler, F., Seo, H.-J., Ross, A. J., et al. 2016, MNRAS, 464, 3409
Blake, C., Brough, S., Colless, M., et al. 2011, MNRAS, 415, 2876
Blas, D., Garny, M., Ivanov, M. M., & Sibiryakov, S. 2016, JCAP, 07, 028
Blot, L., Corasaniti, P. S., Alimi, J.-M., Reverdy, V., & Rasera, Y. 2015, MNRAS, 446, 1756
Blot, L., Corasaniti, P. S., Amendola, L., & Kitching, T. D. 2016, MNRAS, 458, 4462
Blot, L., Crocce, M., Sefusatti, E., et al. 2019, MNRAS, 485, 2806
Bonoli, S., & Pen, U. L. 2009, MNRAS, 396, 1610
Bose, B., Pourtsidou, A., Markovič, K., & Beutler, F. 2020, MNRAS, 493, 5301
Bouchet, F. R., Juszkiewicz, R., Colombi, S., & Pellat, R. 1992, ApJ, 394, L5
Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125
Carrasco, J. J. M., Hertzberg, M. P., & Senatore, L. 2012, J. High Energy Phys., 09, 82
Carrasco, J. J. M., Foreman, S., Green, D., & Senatore, L. 2014, JCAP, 07, 057
Carrilho, P., Moretti, C., & Pourtsidou, A. 2023, JCAP, 01, 028
Castorina, E., Carbone, C., Bel, J., Sefusatti, E., & Dolag, K. 2015, JCAP, 07, 043
Catelan, P., Lucchin, F., Matarrese, S., & Porciani, C. 1998, MNRAS, 297, 692
Catelan, P., Porciani, C., & Kamionkowski, M. 2000, MNRAS, 318, L39
Chan, K. C., Scoccimarro, R., & Sheth, R. K. 2012, Phys. Rev. D, 85, 083509
Chartier, N., Wandelt, B., Akrami, Y., & Villaescusa-Navarro, F. 2021, MNRAS, 503, 1897
Chen, S.-F., Vlah, Z., & White, M. 2022, JCAP, 02, 008
Chuang, C.-H., Yepes, G., Kitaura, F.-S., et al. 2019, MNRAS, 487, 48
Chudaykin, A., Ivanov, M. M., Philcox, O. H., & Simonović, M. 2020, Phys. Rev. D, 102, 063533
Cole, S., & Kaiser, N. 1989, MNRAS, 237, 1127
Coles, P. 1993, MNRAS, 262, 1065
Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039
Contreras, C., Blake, C., Poole, G. B., et al. 2013, MNRAS, 430, 924
Contreras, S., Angulo, R. E., Zennaro, M., Aricò, G., & Pellejero-Ibañez, M. 2020, MNRAS, 499, 4905
Cooray, A., & Sheth, R. 2002, Phys. Rep., 372, 1
Crocce, M., & Scoccimarro, R. 2006, Phys. Rev. D, 73, 063520
Crocce, M., & Scoccimarro, R. 2008, Phys. Rev. D, 77, 023533

Croton, D. J., Gao, L., & White, S. D. M. 2007, MNRAS, 374, 1303
d'Amico, G., Gleyzes, J., Kokron, N., et al. 2020, JCAP, 05, 005
Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2012, AJ, 145, 10
Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, AJ, 151, 44
Dekel, A., & Lahav, O. 1999, ApJ, 520, 24
de la Bella, L. F., Regan, D., Seery, D., & Parkinson, D. 2020, JCAP, 07, 011
DeRose, J., Wechsler, R. H., Tinker, J. L., et al. 2019, ApJ, 875, 69
DeRose, J., Kokron, N., Banerjee, A., et al. 2023, JCAP, 2023, 054
DESI Collaboration (Aghamousa, A., et al.) 2016, ArXiv e-prints [arXiv:1611.00036]
Desjacques, V. 2008, Phys. Rev. D, 78, 103503
Desjacques, V., Crocce, M., Scoccimarro, R., & Sheth, R. K. 2010, Phys. Rev. D, 82, 103529
Desjacques, V., Jeong, D., & Schmidt, F. 2018, Phys. Rep., 733, 1
Donath, Y., & Senatore, L. 2020, JCAP, 2020, 039
Drinkwater, M. J., Jurek, R. J., Blake, C., et al. 2010, MNRAS, 401, 1429
Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, MNRAS, 413, 971
Eggemeier, A., Scoccimarro, R., & Smith, R. E. 2019, Phys. Rev. D, 99, 123514
Eggemeier, A., Scoccimarro, R., Crocce, M., Pezzotta, A., & Sánchez, A. G. 2020, Phys. Rev. D, 102, 103530
Eggemeier, A., Scoccimarro, R., Smith, R. E., et al. 2021, Phys. Rev. D, 103, 123550
Eggemeier, A., Camacho-Quevedo, B., Pezzotta, A., et al. 2022, MNRAS, 519, 2962
Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605
Eisenstein, D. J., Seo, H.-J., & White, M. 2007, ApJ, 664, 660
Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, AJ, 142, 72
Euclid Collaboration (Knabenhans, M., et al.) 2019, MNRAS, 484, 5509
Euclid Collaboration (Blanchard, A., et al.) 2020, A&A, 642, A191
Euclid Collaboration (Knabenhans, M., et al.) 2021, MNRAS, 505, 2840
Feldman, H. A., Frieman, J. A., Fry, J. N., & Scoccimarro, R. 2001, Phys. Rev. Lett., 86, 1434
Feroz, F., Hobson, M. P., & Bridges, M. 2009a, MNRAS, 398, 1601
Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. 2019b, Open J. Astrophys., 2, 10
Fisher, K. B. 1995, ApJ, 448, 494
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Frusciante, N., & Sheth, R. K. 2012, JCAP, 11, 016
Fry, J. N. 1996, ApJ, 461, L65
Fry, J. N., & Gaztanaga, E. 1993, ApJ, 413, 447
Gelman, A., Carlin, J. B., Stern, H. S., et al. 2014, Bayesian Data Analysis (Chapman & Hall/CRC)
Giblin, B., Cataneo, M., Moews, B., & Heymans, C. 2019, MNRAS, 490, 4826
Gil-Marín, H., Noreña, J., Verde, L., et al. 2015, MNRAS, 451, 539
Gil-Marín, H., Guy, J., Zarrouk, P., et al. 2018, MNRAS, 477, 1604
Ginzburg, D., Desjacques, V., & Chan, K. C. 2017, Phys. Rev. D, 96, 083528
Goodman, J., & Weare, J. 2010, Commun. Appl. Math. Comput. Sci., 5, 65
Goroff, M. H., Grinstein, B., Rey, S. J., & Wise, M. B. 1986, ApJ, 311, 6
Grieb, J. N., Sánchez, A. G., Salazar-Albornoz, S., & Dalla Vecchia, C. 2016, MNRAS, 457, 1577
Grieb, J. N., Sánchez, A. G., Salazar-Albornoz, S., et al. 2017, MNRAS, 467, 2085
Gualdi, D., Gil-Marín, H., & Verde, L. 2021, JCAP, 07, 008
Guzzo, L., Pierleoni, M., Meneux, B., et al. 2008, Nature, 451, 541
Guzzo, L., Scodeggio, M., Garilli, B., et al. 2014, A&A, 566, A108
Hadzhiyska, B., García-García, C., Alonso, D., Nicola, A., & Slosar, A. 2021a, JCAP, 2021, 020
Hadzhiyska, B., Liu, S., Somerville, R. S., et al. 2021b, MNRAS, 508, 698
Hahn, T. 2005, Comput. Phys. Commun., 168, 78
Hamann, J., Hannestad, S., Lesgourgues, J., Rampf, C., & Wong, Y. Y. 2010, JCAP, 07, 022
Hamaus, N., Seljak, U., Desjacques, V., Smith, R. E., & Baldauf, T. 2010, Phys. Rev. D, 82, 043515
Hamilton, A. J. S. 1992, ApJ, 385, L5
Hamilton, A. J. S. 2000, MNRAS, 312, 257
Hastings, W. K. 1970, Biometrika, 57, 97
Heitmann, K., Lawrence, E., Kwan, J., Habib, S., & Higdon, D. 2013, ApJ, 780, 111
Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
Hou, J., Sánchez, A. G., Scoccimarro, R., et al. 2018, MNRAS, 480, 2521
Howlett, C., Ross, A. J., Samushia, L., Percival, W. J., & Manera, M. 2015, MNRAS, 449, 848
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Ishiyama, T., Prada, F., Klypin, A. A., et al. 2021, MNRAS, 506, 4210
Ivanov, M. M., Simonović, M., & Zaldarriaga, M. 2020, JCAP, 05, 042
Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111
Izard, A., Crocce, M., & Fosalba, P. 2016, MNRAS, 459, 2327
Jain, B., & Bertschinger, E. 1994, ApJ, 431, 495
Jones, D. H., Read, M. A., Saunders, W., et al. 2009, MNRAS, 399, 683
Kaiser, N. 1984, ApJ, 284, L9
Kaiser, N. 1987, MNRAS, 227, 1
Kokron, N., DeRose, J., Chen, S.-F., White, M., & Wechsler, R. H. 2021, MNRAS, 505, 1422
Kokron, N., Chen, S.-F., White, M., DeRose, J., & Maus, M. 2022, JCAP, 09, 059
Kuhlen, M., Vogelsberger, M., & Angulo, R. 2012, Phys. Dark Univ., 1, 50
Lazeyras, T., & Schmidt, F. 2018, JCAP, 09, 008
Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv e-prints [arXiv:1110.3193]
Lazeyras, T., Barreira, A., & Schmidt, F. 2021, JCAP, 10, 063
Lazeyras, T., Barreira, A., Schmidt, F., & Desjacques, V. 2023, JCAP, 2023, 023
Le Fèvre, O., Cassata, P., Cucciati, O., et al. 2013, A&A, 559, A14
Liu, J., Bird, S., Zorrilla Matilla, J. M., et al. 2018, JCAP, 03, 049
Maksimova, N. A., Garrison, L. H., Eisenstein, D. J., et al. 2021, MNRAS, 508, 4017
Manera, M., & Gaztañaga, E. 2011, MNRAS, 415, 383
Manera, M., Sheth, R. K., & Scoccimarro, R. 2010, MNRAS, 402, 589
Markovič, K., Bose, B., & Pourtsidou, A. 2019, Open J. Astrophys., 2, 13
Matsubara, T. 1999, ApJ, 525, 543
Matsubara, T. 2008, Phys. Rev. D, 77, 063530
McDonald, P. 2006, Phys. Rev. D, 74, 103512
McDonald, P., & Roy, A. 2009, JCAP, 08, 020
McEwen, J. E., Fang, X., Hirata, C. M., & Blazek, J. A. 2016, JCAP, 09, 015
Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, J. Chem. Phys., 21, 1087
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347
Modi, C., Chen, S.-F., & White, M. 2020, MNRAS, 492, 5754
Monaco, P., Theuns, T., & Taffoni, G. 2002, MNRAS, 331, 587
Moradinezhad Dizgah, A., Biagetti, M., Sefusatti, E., Desjacques, V., & Noreña, J. 2021, JCAP, 05, 015
Moretti, C., Tsedrik, M., Carrilho, P., & Pourtsidou, A. 2023, JCAP, 2023, 025
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
Nicola, A., Hadzhiyska, B., Findlay, N., et al. 2024, JCAP, 2024, 015
Nishimichi, T., Takada, M., Takahashi, R., et al. 2019, ApJ, 884, 29
Nusser, A., & Davis, M. 1994, ApJ, 421, L1
Oddo, A., Sefusatti, E., Porciani, C., Monaco, P., & Sánchez, A. G. 2020, JCAP, 03, 056
Oddo, A., Rizzo, F., Sefusatti, E., Porciani, C., & Monaco, P. 2021, JCAP, 11, 038
Okumura, T., Hikage, C., Totani, T., et al. 2016, PASJ, 68, 38
Osato, K., Nishimichi, T., Bernardeau, F., & Taruya, A. 2019, Phys. Rev. D, 99, 063530
Pandey, S., Krause, E., DeRose, J., et al. 2022, Phys. Rev. D, 106, 043520
Peacock, J. A., Cole, S., Norberg, P., et al. 2001, Nature, 410, 169
Pellejero-Ibañez, M., Angulo, R. E., Zennaro, M., et al. 2023, MNRAS, 520, 3725
Perko, A., Senatore, L., Jennings, E., & Wechsler, R. H. 2016, ArXiv e-ptints [arXiv:1610.09321]
Pezzotta, A., de la Torre, S., Bel, J., et al. 2017, A&A, 604, A33
Pezzotta, A., Crocce, M., Eggemeier, A., Sánchez, A. G., & Scoccimarro, R. 2021, Phys. Rev. D, 104, 043531
Philcox, O. H., & Ivanov, M. M. 2022, Phys. Rev. D, 105, 043517
Planck Collaboration VI. 2020, A&A, 641, A6
Pollack, J. E., Smith, R. E., & Porciani, C. 2012, MNRAS, 420, 3469
Pollack, J. E., Smith, R. E., & Porciani, C. 2014, MNRAS, 440, 555
Porredon, A., Crocce, M., Elvin-Poole, J., et al. 2022, Phys. Rev. D, 106, 103530
Potter, D., Stadel, J., & Teyssier, R. 2017, CompAC, 4, 2
Pozzetti, L., Hirata, C. M., Geach, J. E., et al. 2016, A&A, 590, A3
Pueblas, S., & Scoccimarro, R. 2009, Phys. Rev. D, 80, 043504
Reid, B. A., Samushia, L., White, M., et al. 2012, MNRAS, 426, 2719
Reid, B., Ho, S., Padmanabhan, N., et al. 2016, MNRAS, 455, 1553
Rizzo, F., Moretti, C., Pardede, K., et al. 2023, JCAP, 01, 031
Roth, N., & Porciani, C. 2011, MNRAS, 415, 829
Saito, S., Baldauf, T., Vlah, Z., et al. 2014, Phys. Rev. D, 90, 123522
Sánchez, A. G., Montesano, F., Kazin, E. A., et al. 2014, MNRAS, 440, 2692
Sánchez, A. G., Scoccimarro, R., Crocce, M., et al. 2016, MNRAS, 464, 1640
Sánchez, A. G., Ruiz, A. N., Jara, J. G., & Padilla, N. D. 2022, MNRAS, 514, 5673
Scaramella, R., Amiaux, J., Mellier, Y., et al. 2022, A&A, 662, A112
Schechter, P. 1976, ApJ, 203, 297
Scherrer, R. J., & Weinberg, D. H. 1998, ApJ, 504, 607
Schmidt, F. 2016, Phys. Rev. D, 93, 063512
Schmidt, F., Jeong, D., & Desjacques, V. 2013, Phys. Rev. D, 88, 023515
Schmittfull, M., Simonović, M., Assassi, V., & Zaldarriaga, M. 2019, Phys. Rev. D, 100, 043514

Schneider, A., Teyssier, R., Potter, D., et al. 2016, JCAP, 04, 047
Scoccimarro, R. 2004, Phys. Rev. D, 70, 083007
Scoccimarro, R., Zaldarriaga, M., & Hui, L. 1999a, ApJ, 527, 1
Scoccimarro, R., Couchman, H. M. P., & Frieman, J. A. 1999b, ApJ, 517, 531
Scoccimarro, R., Sheth, R. K., Hui, L., & Jain, B. 2001, ApJ, 546, 20
Sefusatti, E., Crocce, M., Pueblas, S., & Scoccimarro, R. 2006, Phys. Rev. D, 74, 023522
Sefusatti, E., Crocce, M., Scoccimarro, R., & Couchman, H. M. P. 2016, MNRAS, 460, 3624
Semenaite, A., Sánchez, A. G., Pezzotta, A., et al. 2022, MNRAS, 512, 5657
Semenaite, A., Sánchez, A. G., Pezzotta, A., et al. 2023, MNRAS, 521, 5013
Senatore, L. 2015, JCAP, 11, 007
Senatore, L., & Zaldarriaga, M. 2014, ArXiv e-prints [arXiv:1409.1225]
Seo, H.-J., Siegel, E. R., Eisenstein, D. J., & White, M. 2008, ApJ, 686, 13
Sheth, R. K., & Lemson, G. 1999, MNRAS, 304, 767
Sheth, R. K., & Tormen, G. 1999, MNRAS, 308, 119
Sheth, R. K., Chan, K. C., & Scoccimarro, R. 2013, Phys. Rev. D, 87, 083002
Simonović, M., Baldauf, T., Zaldarriaga, M., Carrasco, J. J., & Kollmeier, J. A. 2018, JCAP, 04, 030
Skilling, J. 2006, Bayesian Anal., 1, 833
Smith, R. E., Scoccimarro, R., & Sheth, R. K. 2007, Phys. Rev. D, 75, 063512
Springel, V., Pakmor, R., Zier, O., & Reinecke, M. 2021, MNRAS, 506, 2871
Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, AJ, 124, 1810
Szalay, A. S. 1988, ApJ, 333, 21
Taruya, A., & Soda, J. 1999, ApJ, 522, 46
Taruya, A., Nishimichi, T., & Saito, S. 2010, Phys. Rev. D, 82, 063522
Tassev, S., Zaldarriaga, M., & Eisenstein, D. J. 2013, JCAP, 06, 036
Tegmark, M., Eisenstein, D. J., Strauss, M. A., et al. 2006, Phys. Rev. D, 74, 123507
Tröster, T., Sánchez, A. G., Asgari, M., et al. 2020, A&A, 633, L10
Villaescusa-Navarro, F., Hahn, C., Massara, E., et al. 2020, ApJS, 250, 2
Vlah, Z., Seljak, U., Yat Chu, M., & Feng, Y. 2016, JCAP, 03, 057
Wadekar, D., & Scoccimarro, R. 2020, Phys. Rev. D, 102, 123517
Wadekar, D., Ivanov, M. M., & Scoccimarro, R. 2020, Phys. Rev. D, 102, 123521
Wang, Y. 2008, Phys. Rev. D, 77, 123525
Wang, Y., Zhao, G.-B., Chuang, C.-H., et al. 2018, MNRAS, 481, 3160
Wibking, B. D., Salcedo, A. N., Weinberg, D. H., et al. 2018, MNRAS, 484, 989
Winther, H. A., Casas, S., Baldi, M., et al. 2019, Phys. Rev. D, 100, 123540
York, D. G., Adelman, J., Anderson, John E., J., , et al. 2000, AJ, 120, 1579
Yuan, S., Garrison, L. H., Hadzhiyska, B., Bose, S., & Eisenstein, D. J. 2022, MNRAS, 510, 3301
Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, ApJ, 736, 59
Zel'dovich, Y. B. 1970, A&A, 5, 84
Zennaro, M., Angulo, R. E., Aricò, G., Contreras, S., & Pellejero-Ibáñez, M. 2019, MNRAS, 489, 5938
Zennaro, M., Angulo, R. E., Pellejero-Ibáñez, M., et al. 2023, MNRAS, 524, 2407
Zhao, G.-B., Wang, Y., Saito, S., et al. 2018, MNRAS, 482, 3497

[1] Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 1, 85748 Garching, Germany
e-mail: pezzotta@mpe.mpg.de
[2] SISSA, International School for Advanced Studies, Via Bonomea 265, 34136 Trieste TS, Italy
[3] ICSC – Centro Nazionale di Ricerca in High Performance Computing, Big Data e Quantum Computing, Via Magnanelli 2, Bologna, Italy
[4] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
[5] INAF-Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, 34143 Trieste, Italy
[6] IFPU, Institute for Fundamental Physics of the Universe, via Beirut 2, 34151 Trieste, Italy
[7] Department of Physics, Oxford University, Keble Road, Oxford OX1 3RH, UK
[8] Université de Genève, Département de Physique Théorique and Centre for Astroparticle Physics, 24 quai Ernest-Ansermet, 1211 Genève 4, Switzerland
[9] Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain
[10] Institut de Ciencies de l'Espai (IEEC-CSIC), Campus UAB, Carrer de Can Magrans, s/n Cerdanyola del Vallés, 08193 Barcelona, Spain
[11] INFN, Sezione di Trieste, Via Valerio 2, 34127 Trieste TS, Italy
[12] Institute of Theoretical Astrophysics, University of Oslo, PO Box 1029 Blindern, 0315 Oslo, Norway
[13] International Centre for Theoretical Physics (ICTP), Strada Costiera 11, 34151 Trieste, Italy
[14] Universität Bonn, Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany
[15] Excellence Cluster ORIGINS, Boltzmannstrasse e 2, 85748 Garching, Germany
[16] Ludwig-Maximilians-University, Schellingstrasse 4, 80799 Munich, Germany
[17] Donostia International Physics Center (DIPC), Paseo Manuel de Lardizabal, 4, 20018 Donostia-San Sebastián, Guipuzkoa, Spain
[18] IKERBASQUE, Basque Foundation for Science, 48013 Bilbao, Spain
[19] Dipartimento di Scienze Matematiche, Fisiche e Informatiche, Università di Parma, Viale delle Scienze 7/A, 43124 Parma, Italy
[20] Technion Israel Institute of Technology, Israel
[21] Institut d'Estudis Espacials de Catalunya (IEEC), Carrer Gran Capitá 2-4, 08034 Barcelona, Spain
[22] Aix-Marseille Université, CNRS, CNES, LAM, Marseille, France
[23] Area Science Park, Località Padriciano 99, 34149 Trieste, Italy
[24] Laboratoire Univers et Théorie, Observatoire de Paris, Université PSL, Université Paris Cité, CNRS, 92190 Meudon, France
[25] Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano, Via Celoria 16, 20133 Milano, Italy
[26] INFN-Sezione di Milano, Via Celoria 16, 20133 Milano, Italy
[27] INFN Gruppo Collegato di Parma, Viale delle Scienze 7/A, 43124 Parma, Italy
[28] Department of Mathematics and Physics, Roma Tre University, Via della Vasca Navale 84, 00146 Rome, Italy
[29] INFN-Sezione di Roma Tre, Via della Vasca Navale 84, 00146 Roma, Italy
[30] Aix-Marseille Université, Université de Toulon, CNRS, CPT, Marseille, France
[31] Dipartimento di Fisica – Sezione di Astronomia, Università di Trieste, Via Tiepolo 11, 34131 Trieste, Italy
[32] INAF-Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano, Italy
[33] INFN-Sezione di Genova, Via Dodecaneso 33, 16146 Genova, Italy
[34] Division of Theoretical Physics, Ruđer Bošković Institute, Zagreb 10000, Croatia
[35] Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[36] DAMTP, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK
[37] School of Mathematics and Physics, University of Surrey, Guildford, Surrey GU2 7XH, UK
[38] INAF-Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, 40129 Bologna, Italy
[39] Dipartimento di Fisica e Astronomia, Università di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy
[40] INFN-Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
[41] Universitäts-Sternwarte München, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstrasse 1, 81679 München, Germany
[42] INAF-Osservatorio Astrofisico di Torino, Via Osservatorio 20, 10025 Pino Torinese (TO), Italy
[43] Dipartimento di Fisica, Università di Genova, Via Dodecaneso 33, 16146 Genova, Italy
[44] Department of Physics "E. Pancini", University Federico II, Via Cinthia 6, 80126 Napoli, Italy
[45] INAF-Osservatorio Astronomico di Capodimonte, Via Moiariello 16, 80131 Napoli, Italy
[46] INFN section of Naples, Via Cinthia 6, 80126 Napoli, Italy
[47] Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, 4150-762 Porto, Portugal
[48] Dipartimento di Fisica, Università degli Studi di Torino, Via P. Giuria 1, 10125 Torino, Italy

[49] INFN-Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy
[50] INAF-IASF Milano, Via Alfonso Corti 12, 20133 Milano, Italy
[51] INAF-Osservatorio Astronomico di Roma, Via Frascati 33, 00078 Monteporzio Catone, Italy
[52] INFN-Sezione di Roma, Piazzale Aldo Moro, 2 – c/o Dipartimento di Fisica, Edificio G. Marconi, 00185 Roma, Italy
[53] Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra, (Barcelona), Spain
[54] Port d'Informació Científica, Campus UAB, C. Albareda s/n, 08193 Bellaterra (Barcelona), Spain
[55] Institute for Theoretical Particle Physics and Cosmology (TTK), RWTH Aachen University, 52056 Aachen, Germany
[56] Dipartimento di Fisica e Astronomia "Augusto Righi" – Alma Mater Studiorum Università di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
[57] Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK
[58] European Space Agency/ESRIN, Largo Galileo Galilei 1, 00044 Frascati, Roma, Italy
[59] ESAC/ESA, Camino Bajo del Castillo, s/n., Urb. Villafranca del Castillo, 28692 Villanueva de la Cañada, Madrid, Spain
[60] Université Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, UMR 5822, Villeurbanne 69100, France
[61] Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
[62] UCB Lyon 1, CNRS/IN2P3, IUF, IP2I Lyon, 4 rue Enrico Fermi, 69622 Villeurbanne, France
[63] Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Edifício C8, Campo Grande, 1749-016 Lisboa, Portugal
[64] Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal
[65] Department of Astronomy, University of Geneva, ch. d'Ecogia 16, 1290 Versoix, Switzerland
[66] INAF-Istituto di Astrofisica e Planetologia Spaziali, via del Fosso del Cavaliere, 100, 00100 Roma, Italy
[67] INFN-Padova, Via Marzolo 8, 35131 Padova, Italy
[68] Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191 Gif-sur-Yvette, France
[69] Istituto Nazionale di Fisica Nucleare, Sezione di Bologna, Via Irnerio 46, 40126 Bologna, Italy
[70] INAF-Osservatorio Astronomico di Padova, Via dell'Osservatorio 5, 35122 Padova, Italy
[71] University Observatory, Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany
[72] von Hoerner & Sulger GmbH, SchloßPlatz 8, 68723 Schwetzingen, Germany
[73] Technical University of Denmark, Elektrovej 327, 2800 Kgs. Lyngby, Denmark
[74] Cosmic Dawn Center (DAWN), Copenhagen, Denmark
[75] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
[76] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
[77] Department of Physics and Helsinki Institute of Physics, Gustaf Hällströmin katu 2, 00014 University of Helsinki, Finland
[78] Aix-Marseille Université, CNRS/IN2P3, CPPM, Marseille, France
[79] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[80] AIM, CEA, CNRS, Université Paris-Saclay, Université de Paris, 91191 Gif-sur-Yvette, France
[81] Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK
[82] Department of Physics, PO Box 64, 00014 University of Helsinki, Finland
[83] Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, Helsinki, Finland
[84] NOVA optical infrared instrumentation group at ASTRON, Oude Hoogeveensedijk 4, 7991PD Dwingeloo, The Netherlands
[85] Dipartimento di Fisica e Astronomia "Augusto Righi" – Alma Mater Studiorum Università di Bologna, via Piero Gobetti 93/2, 40129 Bologna, Italy
[86] Department of Physics, Institute for Computational Cosmology, Durham University, South Road, DH1 3LE, UK
[87] Institut d'Astrophysique de Paris, 98bis Boulevard Arago, 75014 Paris, France
[88] Institut d'Astrophysique de Paris, UMR 7095, CNRS, and Sorbonne Université, 98 bis boulevard Arago, 75014 Paris, France
[89] European Space Agency/ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
[90] Department of Physics and Astronomy, University of Aarhus, Ny Munkegade 120, 8000 Aarhus C, Denmark
[91] Centre for Astrophysics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
[92] Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
[93] Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada
[94] Université Paris-Saclay, Université Paris Cité, CEA, CNRS, Astrophysique, Instrumentation et Modélisation Paris-Saclay, 91191 Gif-sur-Yvette, France
[95] Space Science Data Center, Italian Space Agency, via del Politecnico snc, 00133 Roma, Italy
[96] CEA Saclay, DFR/IRFU, Service d'Astrophysique, Bat. 709, 91191 Gif-sur-Yvette, France
[97] Université Paris Cité, CNRS, Astroparticule et Cosmologie, 75013 Paris, France
[98] Centre National d'Etudes Spatiales – Centre spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France
[99] Institute of Space Science, Str. Atomistilor, nr. 409 Măgurele, Ilfov 077125, Romania
[100] Dipartimento di Fisica e Astronomia "G. Galilei", Università di Padova, Via Marzolo 8, 35131 Padova, Italy
[101] Departamento de Física, FCFM, Universidad de Chile, Blanco Encalada 2008, Santiago, Chile
[102] Universität Innsbruck, Institut für Astro- und Teilchenphysik, Technikerstr. 25/8, 6020 Innsbruck, Austria
[103] Satlantis, University Science Park, Sede Bld 48940, Leioa-Bilbao, Spain
[104] Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Avenida Complutense 40, 28040 Madrid, Spain
[105] Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Tapada da Ajuda, 1349-018 Lisboa, Portugal
[106] Universidad Politécnica de Cartagena, Departamento de Electrónica y Tecnología de Computadoras, Plaza del Hospital 1, 30202 Cartagena, Spain
[107] Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, CNRS, UPS, CNES, 14 Av. Edouard Belin, 31400 Toulouse, France
[108] Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands
[109] INFN-Bologna, Via Irnerio 46, 40126 Bologna, Italy
[110] Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA
[111] INAF, Istituto di Radioastronomia, Via Piero Gobetti 101, 40129 Bologna, Italy
[112] Instituto de Astrofísica de Canarias, Calle Vía Láctea s/n, 38204 San Cristóbal de La Laguna, Tenerife, Spain
[113] Centre de Calcul de l'IN2P3/CNRS, 21 avenue Pierre de Coubertin, 69627 Villeurbanne Cedex, France
[114] Institut für Theoretische Physik, University of Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany
[115] Université St Joseph; Faculty of Sciences, Beirut, Lebanon
[116] Junia, EPA department, 41 Bd Vauban, 59800 Lille, France
[117] Instituto de Física Teórica UAM-CSIC, Campus de Cantoblanco, 28049 Madrid, Spain

[118] CERCA/ISO, Department of Physics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA
[119] Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, Via Giuseppe Saragat 1, 44122 Ferrara, Italy
[120] Istituto Nazionale di Fisica Nucleare, Sezione di Ferrara, Via Giuseppe Saragat 1, 44122 Ferrara, Italy
[121] Institut de Physique Théorique, CEA, CNRS, Université Paris-Saclay, 91191 Gif-sur-Yvette Cedex, France
[122] Minnesota Institute for Astrophysics, University of Minnesota, 116 Church St SE, Minneapolis, MN 55455, USA
[123] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice cedex 4, France
[124] Institute Lorentz, Leiden University, PO Box 9506, Leiden 2300 RA, The Netherlands
[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] Université Paris-Saclay, CNRS, Institut d'astrophysique spatiale, 91405 Orsay, France
[129] Departamento Física Aplicada, Universidad Politécnica de Cartagena, Campus Muralla del Mar, 30202 Cartagena, Murcia, Spain
[130] Department of Computer Science, Aalto University, PO Box 15400, Espoo 00 076, Finland
[131] Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France
[132] Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK
[133] Department of Physics and Astronomy, Vesilinnantie 5, 20014 University of Turku, Finland
[134] Serco for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada 28692, Madrid, Spain
[135] ARC Centre of Excellence for Dark Matter Particle Physics, Melbourne, Australia
[136] Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Victoria 3122, Australia
[137] W.M. Keck Observatory, 65-1120 Mamalahoa Hwy, Kamuela, HI, USA
[138] Oskar Klein Centre for Cosmoparticle Physics, Department of Physics, Stockholm University, Stockholm 106 91, Sweden
[139] Astrophysics Group, Blackett Laboratory, Imperial College London, London SW7 2AZ, UK
[140] Univ. Grenoble Alpes, CNRS, Grenoble INP, LPSC-IN2P3, 53, Avenue des Martyrs, 38000 Grenoble, France
[141] INAF-Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125 Firenze, Italy
[142] Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 2, 00185 Roma, Italy
[143] Centro de Astrofísica da Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal
[144] Dipartimento di Fisica, Università di Roma Tor Vergata, Via della Ricerca Scientifica 1, Roma, Italy
[145] INFN, Sezione di Roma 2, Via della Ricerca Scientifica 1, Roma, Italy
[146] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[147] Department of Astrophysics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
[148] Higgs Centre for Theoretical Physics, School of Physics and Astronomy, The University of Edinburgh, Edinburgh EH9 3FD, UK
[149] Dipartimento di Fisica, Università degli studi di Genova, and INFN-Sezione di Genova, via Dodecaneso 33, 16146 Genova, Italy
[150] Theoretical astrophysics, Department of Physics and Astronomy, Uppsala University, Box 515, 751 20 Uppsala, Sweden
[151] Department of Astrophysical Sciences, Peyton Hall, Princeton University, Princeton, NJ 08544, USA
[152] Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 Copenhagen, Denmark
[153] Center for Cosmology and Particle Physics, Department of Physics, New York University, New York, NY 10003, USA
[154] Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, 10010 New York, NY, USA

## Appendix A: Matter power spectrum fits and the fiducial cosmology

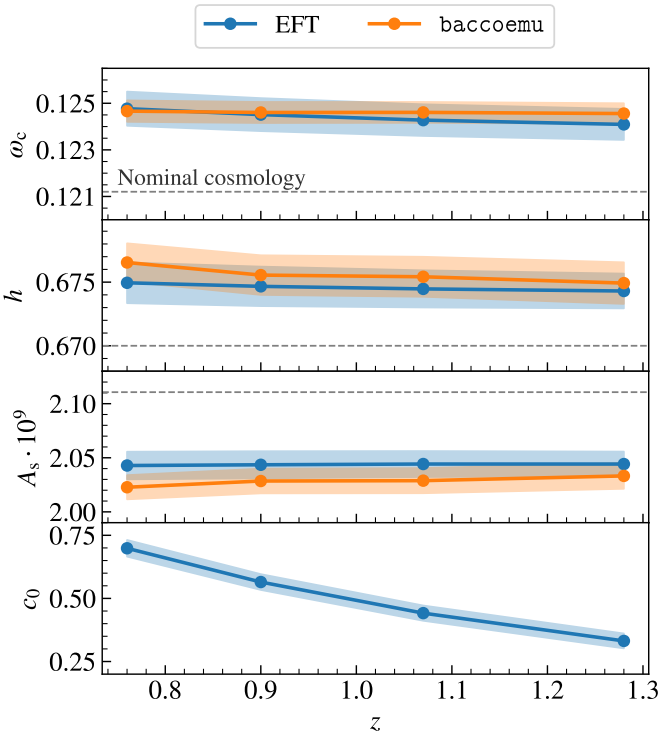


**Fig. A.1.** Marginalised one-dimensional constraints as a function of redshift obtained by fitting the measured matter power spectrum with the EFT model (blue) and the `baccoemu` emulator (orange). Solid lines and shaded bands mark the mean and the standard deviation of the posterior distribution, respectively. For both models, the fit is carried out up to maximum wave mode $k_{\rm max} = 0.25\,h\,{\rm Mpc}^{-1}$, using a Gaussian covariance matrix corresponding to the full box volume of the Flagship I simulation. Dashed lines denote the nominal fiducial values of the parameters $\{h, \omega_c, A_s\}$. In the bottom panel we show the marginalised constraints on the $c_0$ EFT counterterm parameter.

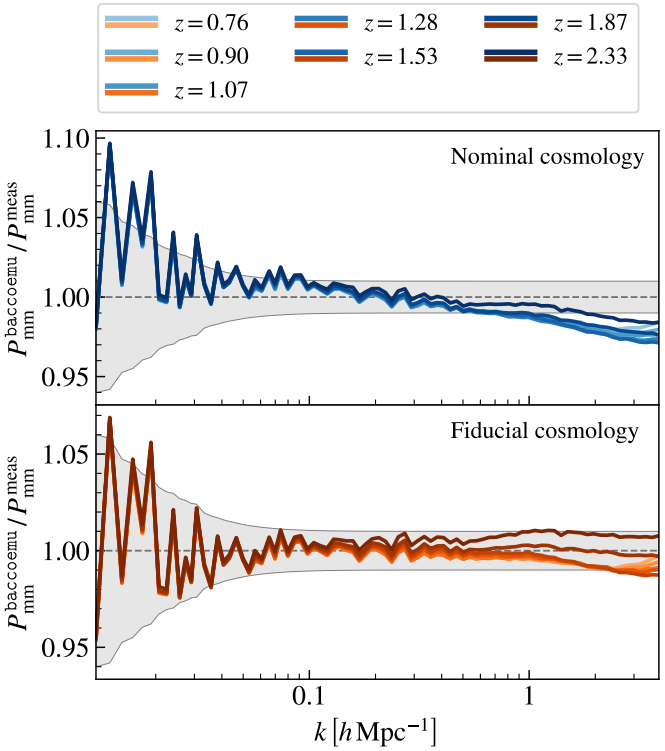


**Fig. A.2.** Ratio between the Flagship matter power spectrum measurements and predictions obtained using `baccoemu`, considering both the nominal cosmology (upper panel) and the fiducial one (bottom panel). In both cases, different colours correspond to a different redshift, as shown in the legend. The grey shaded band mark the intrinsic error of the power spectrum measurements to which an additional 1% contribution has been added to include the error contribution of the emulator, as explained in Sect. 5.1.

In the main body of this article, we carried out tests meant to assess the level of accuracy of different models for the one-loop galaxy power spectrum, using measurements coming from the Flagship I simulation. In addition to this analysis, we also tested our model for the matter power spectrum on a set of measurements obtained on the fly while running the `PKDGRAV3` code in the redshift range $[0.7, 2.4]$. Each measurement consists of the matter power spectrum, measured in the range of wave modes defined by the interval $[0.01, 4]\ h\,{\rm Mpc}^{-1}$, using 18 linearly spaced bins up to $k \sim 0.03\,h\,{\rm Mpc}^{-1}$ and other 84 logarithmically spaced bins after.

We run two independent analyses, the first using the next-to-leading order matter power spectrum obtained in the EFTofLSS framework (Eq. 19) and the second using the non-linear matter power spectrum from the `baccoemu` emulator. We use an analytical Gaussian covariance matrix to describe the error on the matter power spectrum measurements, assuming the full volume of the simulation box, that is, $(3780\,h^{-1}\,{\rm Mpc})^3$. For both models, we limit the maximum mode of the fit to $k_{\rm max} = 0.25\,h\,{\rm Mpc}^{-1}$, within the expected PT range of validity for the relevant redshifts.

In Fig. A.1 we show the marginalised 1d posterior distributions for $\{h, \omega_c, A_s\}$ as a function of redshift. For this test we limit the redshift range to the first four snapshots ($z < 1.3$), since the results are sufficient to draw conclusions on the agreement between data vector and theory models. Also shown, with dashed lines, are the nominal values of the parameters provided in Potter et al. (2017). We find that all cosmological parameters are obtained with a bias of 3–4$\sigma$ from their fiducial values, in a way that is consistent across redshifts, as highlighted by the almost constant trends in each of the panels. In addition, the EFT model and `baccoemu` are consistent at better than $1\sigma$ with each other, pointing to a systematic effect that cannot be attributed to the particular model used to describe the data vectors.

In the upper panel of Fig. A.2 we show the ratio of the non-linear matter power spectrum from the `baccoemu` emulator assuming the nominal cosmology to the measured one for the different redshifts. We note that the discrepancy among the two sets of curves is apparently due to a different tilt in the full-shape of the matter power spectrum, here corresponding to the nominal value of $n_s = 0.96$. Indeed we note that the disagreement between nominal and recovered cosmology can be alleviated for all parameters by assuming a different value of $n_s$, while keeping all the remaining parameters fixed to their nominal value (with the exception of the scalar amplitude $A_s$, since the latter has to be modified to recover the nominal value of $\sigma_8$). This is also supported by the matter transfer function files that were generated to set-up the initial conditions for the Flagship I simulation, which confirmed how all the parameters affecting the shape of the transfer function $\{\omega_c, \omega_b, M_\nu\}$ are consistent with the nominal values listed in Table 1.

We therefore performed new fits to the Flagship I matter power spectra keeping all parameters fixed (including the Hubble parameter $h$) to the nominal values, and only sampling those controlling the primordial matter power spectrum, that is, the scalar amplitude $A_s$ and index $n_s$. The best-fit value that we found for the spectral index is close to the value of $n_s = 0.97$. This leads to consistent results across the redshift range we are considering, as well as for a broad range of scales, up to $k \sim 4\,h\,\mathrm{Mpc}^{-1}$, as shown in the bottom panel of Fig. A.2.

The new values of the spectral index and scalar amplitude, along with nominal values for the other parameters constitute what we refer to as the fiducial cosmology. This is the one given in Table 1 and adopted for all predictions throughout this paper.

## Appendix B: Standard perturbation theory

In this section we report the full expressions for several quantities that are defined in Sect. 3 and used throughout the rest of the paper. For a more detailed and exhaustive description of the framework on which cosmological PT is based, we refer the reader to the comprehensive review by Bernardeau et al. (2002).

The main idea behind cosmological PT is that the generic solution to the growth of non-linear density and velocity fluctuations – $\delta(\boldsymbol{k})$ and $\theta(\boldsymbol{k})$ – in an expanding universe can be expressed in terms of linear theory solutions, $\delta_{\rm L}(\boldsymbol{k})$. Assuming an EdS universe, it is possible to perfectly separate the time- and space-dependence of $\delta(\boldsymbol{k})$ and $\theta(\boldsymbol{k})$ (Goroff et al. 1986; Jain & Bertschinger 1994). In an arbitrary ΛCDM cosmology, it can be shown that approximate solutions can be found with the same separation (Donath & Senatore 2020), especially in a range of redshift for which the EdS cosmology is still a valid approximation. In this case we can write the density and velocity divergence fields using an expansion of the form

$$\delta\left(\boldsymbol{k},\tau\right) = \sum_{n=1}^{\infty} D_n(\tau)\,\delta^{(n)}(\boldsymbol{k})\,, \tag{B.1}$$

$$\theta\left(\boldsymbol{k},\tau\right) = -\mathcal{H}(\tau) f(\tau) \sum_{n=1}^{\infty} D_n(\tau)\,\theta^{(n)}(\boldsymbol{k})\,, \tag{B.2}$$

where $\tau$ is the conformal time defined via $\mathrm{d}t = a(\tau)\,\mathrm{d}\tau$, $a(\tau)$ is the cosmic scale factor, $\mathcal{H}(\tau) \equiv \mathrm{d}\ln a(\tau)/\mathrm{d}\tau$ is the conformal Hubble expansion factor, and $f(\tau) \equiv \mathrm{d}\ln D_1(\tau)/\mathrm{d}\ln a(\tau)$ is the growth rate. The $n$-th order growth factor $D_n$ characterises the time-dependence of the density and velocity field, and reduces to $D_n = a^n$ in the EdS limit. Assuming the conservation of mass, momentum, and the Poisson equation, the individual $n$-th order corrections to the density and velocity fields can be written as

$$\begin{aligned}\delta^{(n)}(\boldsymbol{k}) = &\int_{\boldsymbol{q}_1} \dots \int_{\boldsymbol{q}_n} \delta_{\rm D}^{(3)}(\boldsymbol{k} - \boldsymbol{q}_{1\dots n})\, F_n(\boldsymbol{q}_1,\dots,\boldsymbol{q}_n) \\ &\times \delta_{\rm L}(\boldsymbol{q}_1)\dots\delta_{\rm L}(\boldsymbol{q}_n)\,,\end{aligned} \tag{B.3}$$

$$\begin{aligned}\theta^{(n)}(\boldsymbol{k}) = &\int_{\boldsymbol{q}_1} \dots \int_{\boldsymbol{q}_n} \delta_{\rm D}^{(3)}(\boldsymbol{k} - \boldsymbol{q}_{1\dots n})\, G_n(\boldsymbol{q}_1,\dots,\boldsymbol{q}_n) \\ &\times \delta_{\rm L}(\boldsymbol{q}_1)\dots\delta_{\rm L}(\boldsymbol{q}_n)\,,\end{aligned} \tag{B.4}$$

where the $n$-th order PT kernels $F_n$ and $G_n$ are homogeneous functions of the wave vectors $(\boldsymbol{q}_1,\dots,\boldsymbol{q}_n)$, and are built starting from the fundamental mode-coupling functions,

$$\alpha\left(\boldsymbol{k}_1,\boldsymbol{k}_2\right) \equiv \frac{\boldsymbol{k}_{12}\cdot\boldsymbol{k}_1}{k_1^2}\,, \tag{B.5}$$

$$\beta\left(\boldsymbol{k}_1,\boldsymbol{k}_2\right) \equiv \frac{k_{12}^2(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)}{2\,k_1^2 k_2^2}\,, \tag{B.6}$$

with $\boldsymbol{k}_{12} = \boldsymbol{k}_1 + \boldsymbol{k}_2$. At linear order these quantities clearly become unity, $F_1 = G_1 = 1$, so to recover linear theory predictions, that is, $\delta^{(1)}(\boldsymbol{k}) = \theta^{(1)}(\boldsymbol{k}) = \delta_{\rm L}(\boldsymbol{k})$. At higher order, these kernels can be derived using recursive relations, which read

$$\begin{aligned}F_n\left(\boldsymbol{q}_1,\dots,\boldsymbol{q}_n\right) = &\sum_{m=1}^{n-1} \frac{G_m(\boldsymbol{q}_1,\dots,\boldsymbol{q}_m)}{(2n+3)\,(n-1)} \\ &\times\Big[(2n+1)\,\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2)\,F_{n-m}\left(\boldsymbol{q}_{m+1},\dots,\boldsymbol{q}_n\right) \\ &+ 2\beta(\boldsymbol{k}_1,\boldsymbol{k}_2)\,G_{n-m}(\boldsymbol{q}_{m+1},\dots,\boldsymbol{q}_n)\Big]\,,\end{aligned} \tag{B.7}$$

$$\begin{aligned}G_n\left(\boldsymbol{q}_1,\dots,\boldsymbol{q}_n\right) = &\sum_{m=1}^{n-1} \frac{G_m(\boldsymbol{q}_1,\dots,\boldsymbol{q}_m)}{(2n+3)\,(n-1)} \\ &\times\Big[3\,\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2)\,F_{n-m}(\boldsymbol{q}_{m+1},\dots,\boldsymbol{q}_n) \\ &+ 2n\beta(\boldsymbol{k}_1,\boldsymbol{k}_2)\,G_{n-m}(\boldsymbol{q}_{m+1},\dots,\boldsymbol{q}_n)\Big]\,,\end{aligned} \tag{B.8}$$

where $\boldsymbol{k}_1 = \boldsymbol{q}_1 + \dots + \boldsymbol{q}_m$ and $\boldsymbol{k}_2 = \boldsymbol{q}_{m+1} + \dots + \boldsymbol{q}_n$. As a classical example, needed for the calculation of the one-loop galaxy power spectrum, the second-order PT kernels for the non-linear evolution of the matter density and velocity fields are defined as

$$F_2(\boldsymbol{k}_1,\boldsymbol{k}_2) = \frac{5}{7} + \frac{1}{2}\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_2}{k_1 k_2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right) + \frac{2}{7}\frac{(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)^2}{k_1^2\,k_2^2}\,, \tag{B.9}$$

$$G_2(\boldsymbol{k}_1,\boldsymbol{k}_2) = \frac{3}{7} + \frac{1}{2}\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_2}{k_1 k_2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right) + \frac{4}{7}\frac{(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)^2}{k_1^2\,k_2^2}\,, \tag{B.10}$$

while the explicit expression for the third-order kernel of the matter density field, $F_3(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)$, can be found in Goroff et al. (1986), and it is not reported here for practical purposes.

When considering the galaxy power spectrum $P_{\rm gg}$, we need to evaluate the usual two-point statistics defined by the ensemble average of the galaxy density field times itself,

$$\left\langle \delta_{\rm g}(\boldsymbol{k})\,\delta_{\rm g}(\boldsymbol{k}')\right\rangle = (2\pi)^3\,\delta_{\rm D}^{(3)}(\boldsymbol{k}+\boldsymbol{k}')\,P_{\rm gg}(k)\,. \tag{B.11}$$

At third order in the perturbations of $\delta$, we obtain the one-loop expression for the power spectrum presented in Eq. (27), where all the next-to-leading order corrections are grouped into mode-coupling and propagator-like contributions. In the former, the loop integrand is proportional to $P_{\rm L}(|\boldsymbol{k}-\boldsymbol{q}|)\,P_{\rm L}(q)$, reflecting the mixing of modes due to non-linear evolution, while in the latter the integrals are carried out on the factor $P_{\rm L}(k)\,P_{\rm L}(q)$, corresponding to a time propagation of the initial density field. Separating these two groups into individual corrections, each one multiplied by a given combination of bias parameters, we end up with the following scheme,

$$\begin{aligned}P_{\rm gg}^{\text{1-loop}}(k) = &\,b_1^2\,P^{\text{1-loop}}(k) \\ &+ b_1 b_2\,P_{b_1 b_2}(k) + b_1 b_{\mathcal{G}_2}\,P_{b_1 b_{\mathcal{G}_2}}(k) \\ &+ b_1 b_{\Gamma_3}\,P_{b_1 b_{\Gamma_3}}(k) + b_2^2\,P_{b_2 b_2}(k) \\ &+ b_2 b_{\mathcal{G}_2}\,P_{b_2 b_{\mathcal{G}_2}}(k) + b_{\mathcal{G}_2}^2\,P_{b_{\mathcal{G}_2} b_{\mathcal{G}_2}}(k)\,,\end{aligned} \tag{B.12}$$

where all the previous terms can be represented as loop integrals,

$$\begin{aligned}P^{\text{1-loop}}(k) &= P^{\text{1-loop, MC}}(k) + P^{\text{1-loop, Prop}}(k) \\ &= 2\int_{\boldsymbol{q}} F_2^2(\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q})\,P_{\rm L}(|\boldsymbol{k}-\boldsymbol{q}|)\,P_{\rm L}(q) \\ &\quad + 6\,P_{\rm L}(k)\int_{\boldsymbol{q}} F_3(\boldsymbol{q},-\boldsymbol{q},\boldsymbol{k})\,P_{\rm L}(q)\,,\end{aligned} \tag{B.13}$$

$$P_{b_1 b_2}(k) = 2 \int_{\boldsymbol{q}} F_2(\boldsymbol{k}-\boldsymbol{q}, \boldsymbol{q})\, P_{\rm L}(|\boldsymbol{k}-\boldsymbol{q}|)\, P_{\rm L}(q)\,, \tag{B.14}$$

$$\begin{aligned} P_{b_1 b_{\mathcal{G}_2}}(k) &= P^{\rm MC}_{b_1 b_{\mathcal{G}_2}}(k) + P^{\rm Prop}_{b_1 b_{\mathcal{G}_2}}(k) \\ &= 4 \int_{\boldsymbol{q}} F_2(\boldsymbol{k}-\boldsymbol{q}, \boldsymbol{q})\, S(\boldsymbol{k}-\boldsymbol{q}, \boldsymbol{q})\, P_{\rm L}(|\boldsymbol{k}-\boldsymbol{q}|)\, P_{\rm L}(q) \\ &\quad + 8\, P_{\rm L}(k) \int_{\boldsymbol{q}} F_2(\boldsymbol{k}, -\boldsymbol{q})\, S(\boldsymbol{k}-\boldsymbol{q}, \boldsymbol{q})\, P_{\rm L}(q)\,, \end{aligned} \tag{B.15}$$

$$P_{b_1 b_{\Gamma_3}}(k) = -\frac{16}{7} P_{\rm L}(k) \int_{\boldsymbol{q}} S(\boldsymbol{k}-\boldsymbol{q}, \boldsymbol{q})\, S(\boldsymbol{k}, \boldsymbol{q})\, P_{\rm L}(q)\,, \tag{B.16}$$

$$P_{b_2 b_2}(k) = \frac{1}{2} \int_{\boldsymbol{q}} P_{\rm L}(|\boldsymbol{k}-\boldsymbol{q}|)\, P_{\rm L}(q)\,, \tag{B.17}$$

$$P_{b_2 b_{\mathcal{G}_2}}(k) = 2 \int_{\boldsymbol{q}} S(\boldsymbol{k}-\boldsymbol{q}, \boldsymbol{q})\, P_{\rm L}(\boldsymbol{k}-\boldsymbol{q}, \boldsymbol{q})\, P_{\rm L}(q)\,, \tag{B.18}$$

$$P_{b_{\mathcal{G}_2} b_{\mathcal{G}_2}}(k) = 2 \int_{\boldsymbol{q}} S^2(\boldsymbol{k}-\boldsymbol{q}, \boldsymbol{q})\, P_{\rm L}(\boldsymbol{k}-\boldsymbol{q}, \boldsymbol{q})\, P_{\rm L}(q)\,. \tag{B.19}$$

In the previous set of equations, the $P_{b_1 b_{\Gamma_3}}$ contribution is characterised by a single propagator-like term, which is perfectly degenerate with the second addend contributing to $P_{b_1 b_{\mathcal{G}_2}}$. Following the expansion that has been adopted in this paper, the degeneracy results from the equality

$$P_{b_1 b_{\Gamma_3}}(k) = \frac{2}{5} P^{\rm MC}_{b_1 b_{\mathcal{G}_2}}(k)\,. \tag{B.20}$$

As shown in the main body of this article, breaking this degeneracy with the information coming from the galaxy power spectrum alone is not possible, and for this reason we often keep one of the two tidal bias parameters fixed to some physically motivated relation, such as the excursion-set-based relation (Eq. 37) for $b_{\mathcal{G}_2}$, or the coevolution relations for both $b_{\mathcal{G}_2}$ and $b_{\Gamma_3}$ (Eqs. 35 – 36).

The behaviour of all the loop integrals listed above is to consistently converge to zero at infrared modes, since the non-linear kernel $F_n(\boldsymbol{k}_1, \ldots, \boldsymbol{k}_n)$ scales as $k^2$ when $\boldsymbol{k} \equiv \boldsymbol{k}_1 + \ldots + \boldsymbol{k}_n$ goes to zero, reflecting the range of validity of linear theory predictions. The only exception is represented by the $P_{b_2 b_2}$ term, which features a non-zero asymptote for $k \to 0$. This limit can be manually set to zero via a redefinition of the loop integral, such that

$$P_{b_2 b_2}(k) = \frac{1}{2} \int_{\boldsymbol{q}} P_{\rm L}(q)\, \left[P_{\rm L}(|\boldsymbol{k}-\boldsymbol{q}|) - P_{\rm L}(q)\right]\,. \tag{B.21}$$

In turn, the extra contribution

$$P^{\rm noise}_{b_2 b_2}(k) = \int_{\boldsymbol{q}} P^2_{\rm L}(q) \tag{B.22}$$

can be absorbed by the constant shot-noise parameter, $\alpha_{P,1}$, which we have defined in Sect. 3.1.2, as they both correspond to constant shift in the amplitude of the galaxy power spectrum.

## Appendix C: Implementation of the wiggle vs. no-wiggle split

In this section, we investigate the prescriptions used to obtain a smooth template $P_{\rm nw}$ starting from the linear matter power spectrum. This is an important aspect of the theoretical recipe that we adopt, as the wiggle-no wiggle split is essential for the correct implementation of IR-resummation, as shown in Sect. 3.1. While several different algorithms can be found in the literature, here we test three different methods.

The first one is based on a one-dimensional Gaussian smoothing (GS1D), and consists of a rescaling of the original formula for the featureless matter power spectrum $P_{\rm EH}$, originally presented in Eisenstein & Hu (1998), to match the broadband amplitude of the linear matter power spectrum. In practice, we follow the approach of Vlah et al. (2016), who defines the smooth component of the linear matter power spectrum as

$$P_{\rm nw}(k) = P_{\rm EH}(k)\, \mathcal{F}\left[\frac{P_{\rm L}(k)}{P_{\rm EH}(k)}\right]\,, \tag{C.1}$$

where $\mathcal{F}$ is meant to filter out the broadband difference between $P_{\rm L}$ and $P_{\rm EH}$. We choose a functional form for $\mathcal{F}$ corresponding to a Gaussian filter, that is,

$$\mathcal{F}\left[f(k)\right] = \frac{\log_{10}({\rm e})}{\sqrt{2\pi}\lambda} \int {\rm d}q\, \frac{f(q)}{q} \exp\left[-\frac{1}{2\lambda^2} \log^2_{10}\left(\frac{k}{q}\right)\right]\,, \tag{C.2}$$

where $\lambda$ determines the variance of the Gaussian filter used to rescale the ratio of the linear to the featureless power spectrum. In this analysis, we fix its value to $\lambda = 0.25$.

The result of this approach is presented in the top panel of Fig. C.1, where we show a comparison between the shape of the original Eisenstein & Hu (1998) smooth function and the one presented in Eq. (C.1) using the fiducial cosmology from Table 1 at $z = 0$. From the comparison, it is clear that the broadband of $P_{\rm EH}$ features a non-negligible tilt with respect to the one of the linear matter power spectrum, reaching deviations of up to 2% across the whole BAO wave mode interval. This difference can be corrected using the filtering strategy, whose output is completely consistent with the broadband shape of $P_{\rm L}$, with the only exception of a small residual of $\sim 0.5\%$ peaking at $0.5\,h\,{\rm Mpc}^{-1} \lesssim k \lesssim 1\,h\,{\rm Mpc}^{-1}$.

The second approach makes use of a discrete sine transform (DST), and has been originally proposed in Hamann et al. (2010) (see Chudaykin et al. 2020 and Ivanov et al. 2020 for an application to real data). This is based on a fast Fourier transformation of the input matter power spectrum, and in the removal of the bump corresponding to the BAO peak. This step is carried out based on the value of the second derivative of the sine transform, and the generated gap is subsequently filled using a cubic-spline interpolation. The new function is finally transformed back into Fourier space to deliver a power spectrum shape deprived from BAO oscillations.

The third and final approach is based on the approximation of the BAO wiggles with a basis spline (B-spline) curve, starting from a set of knots $\{k_i, P_i\}$, and subsequently finding the spline coefficients that maximise the likelihood with the original power spectrum (see e.g. Vlah et al. 2016, for a similar implementation).

The bottom panel of Fig. C.1 shows a comparison between the three methods summarised above. We note a non-negligible difference, with discrepancies on the BAO scales that can reach a fraction of percent, depending on the considered approach (see Moradinezhad Dizgah et al. 2021, for similar conclusions).

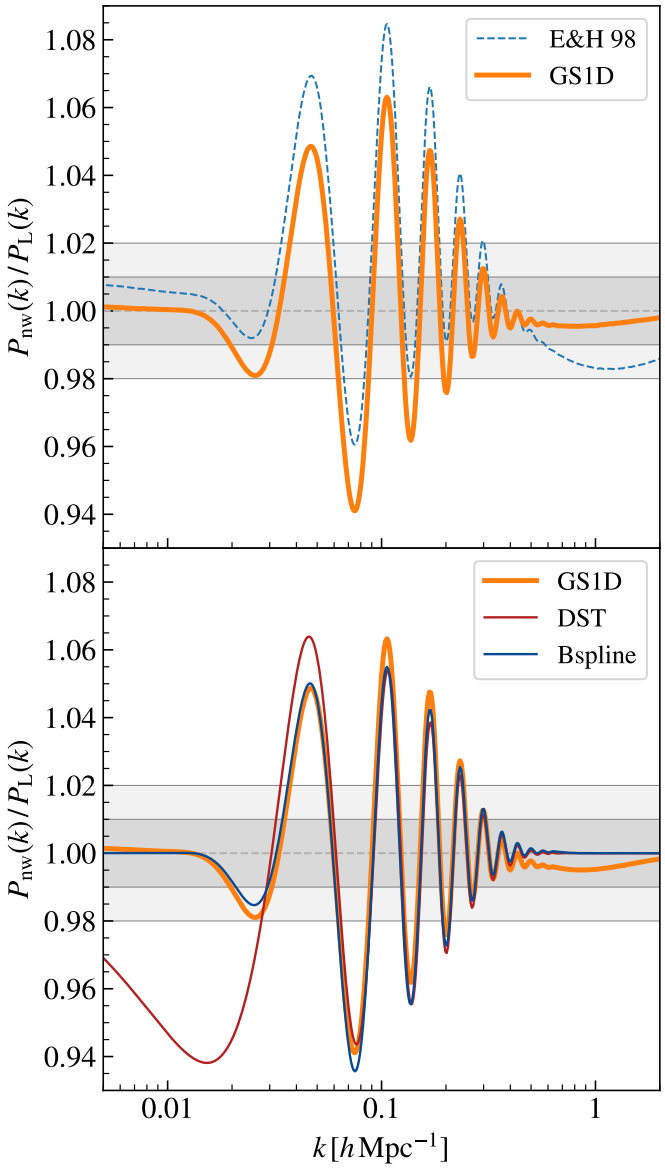


**Fig. C.1.** Ratio between the no-wiggle and the linear matter power spectrum. *Top*: Comparison between the raw featureless power spectrum $P_{\rm EH}$ and the no-wiggle power spectrum $P_{\rm nw}$ computed using Eq. (C.1). In both cases, the power spectra are computed at $z = 0$ using the fiducial cosmology from Table 1. *Bottom*: Comparison between the three different methods we tested to obtain the no-wiggle power spectrum $P_{\rm nw}$. The thick orange line corresponds to the method we selected, i.e. the convolution with a Gaussian smoothing function, while the other two lines represent a Discrete Sine Transform (red) and a basis spline (blue). In both panels, the grey shaded bands represent the 1% (dark grey) and 2% (light grey) thresholds.

Nevertheless, at linear order these differences are diluted by the recombination of the smooth and wiggling component at a later stage, so that the net result on the IR-resummed non-linear galaxy power spectrum is going to be much smaller than the values exhibited in this plot. We highlight a major discrepancy between the DST method and the other two that reaches its maximum at a scale $k \sim 0.01\,h\,{\rm Mpc}^{-1}$, which roughly coincides with the position of the turnaround in the matter power spectrum. Because of this major discrepancy, and since all three methods behave in a slightly different way on the BAO scales, we decided to adopt the GS1D method throughout the analysis presented in this paper. We observe a marginal residuals of this method on mildly non-linear scales, at $0.5\,h\,{\rm Mpc}^{-1} \lesssim k \lesssim 1\,h\,{\rm Mpc}^{-1}$, which, once again, is not likely to significantly bias our results

At the same time we highlight how this approach is most likely going to perform worse when considering cosmologies beyond the vanilla ΛCDM model, such as those including massive neutrinos. In these cases, depending on the magnitude of the deviation from ΛCDM, the broadband shape of $P_{\rm EH}$ can deviate from that of $P_{\rm L}$ by up to 10% (as quoted in Eisenstein & Hu 1998). We therefore adopt one of the other two methods for future analyses, such as for the upcoming analysis of the Flagship II simulation.

## Appendix D: Comparison between model implementation

In order to estimate the systematic error budget due to the implementation of the algorithm for the one-loop model presented in Sect. 3.1 and Appendix B, we make use of four individual implementations, based on four independent codes provided by several group members. Each code features a different way to compute the loop corrections presented in Sect. B.

These include a two-dimensional integration implemented within the `Cuba` library[30] (Hahn 2005) used to generate the training set of the `COMET` emulator (Eggemeier et al. 2022), and implemented in the `CosmoSIS-gClust`[31] code (courtesy of A. Moradinezhad Dizgah). Alternative methods take advantage of a Fast Fourier Transform approach, such as `FastPT`[32] (McEwen et al. 2016) – implemented in the `PBJ` code (Oddo et al. 2020, 2021; Rizzo et al. 2023; Moretti et al. 2023) – and `FFTLog`[33] (Hamilton 2000; Simonović et al. 2018) – used in the `TNSToolBox`[34] code (courtesy of S. de la Torre) and again in `CosmoSIS-gClust`.

In Fig. D.1 we show the systematic deviation between different computations of the same terms, including the leading- and next-to-leading order correction to the IR-resummed matter power spectrum, and the six one-loop bias corrections defined in Eqs. (B.14 – B.19). For this exercise, we compute a common data vector using the fiducial Flagship cosmology from Table 1 at $z = 0$. We assume as a reference measurement the one performed using a direct integration of the wave number $q$ within the range of scales $[0.00001, 100]\ h\,{\rm Mpc}^{-1}$, which proved to be a sensible choice to achieve the convergence of the different loop integrals. Finally, we compare the different terms evaluated using the four codes described in the previous paragraph. We find an overall optimal consistency among the different theory implementations, with all terms being in agreement at better than 0.01% on the overall range of scales shown in Fig. D.1. We observe a slightly worse concordance between different computations of the propagator-like terms, that is, $P_{\rm mm}^{\rm IR\text{-}(LO+NLO)}$, which contains the contribution $P_{13}$, and more importantly $P_{b_1 b_{\Gamma_3}}$, for which the discrepancy can become as large as ∼ 0.05% on scales of $k \sim 1\,h\,{\rm Mpc}^{-1}$. However, we consider this difference completely negligible, since this value is much smaller than the statistical precision of the data vectors, and since this scale is completely outside of the range of scales considered in this work.

[30] http://www.feynarts.de/cuba/
[31] To be made publicly available soon.
[32] https://github.com/JoeMcEwen/FAST-PT/
[33] https://jila.colorado.edu/~ajsh/FFTLog/
[34] https://github.com/sdlt/TNS_ToolBox/

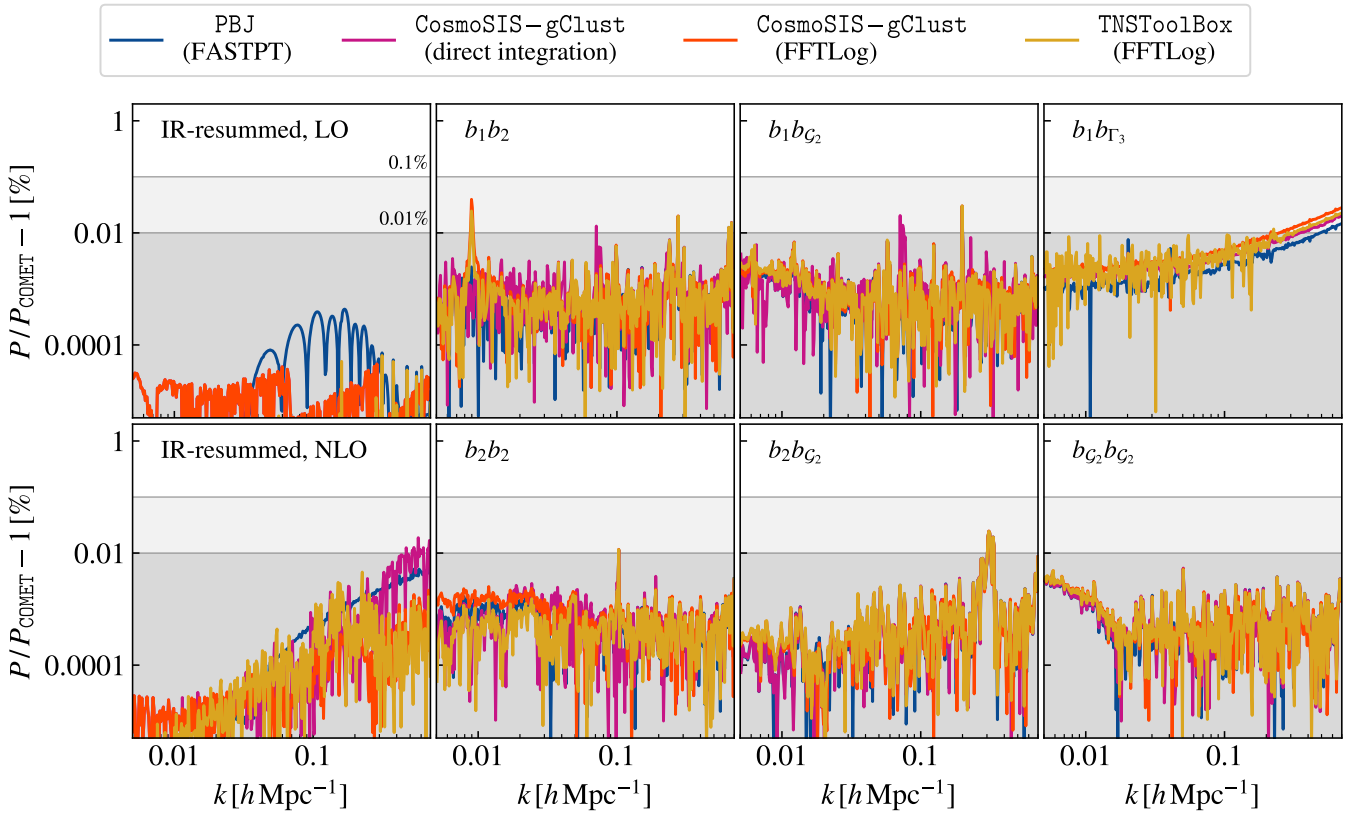


**Fig. D.1.** Comparison among the different methods to obtain the one-loop bias integrals, colour-coded according to the individual code/method, as listed in the legend. The reference code for these plots is the COMET emulator. Different panels corresponds to different diagrams: the first column shows the leading order IR-resummed matter power spectrum – for which no actual integration is required – and the next-to-leading order correction. The remaining columns show the bias diagrams from Eqs. (B.14 – B.19).

## Appendix E: Comparison between posterior averaged and maximum likelihood position

In the main body of this article, we determine the goodness of fit of a given model configuration in terms of the posterior-averaged $\chi^2$ statistics. For this purpose, we simply iterate through the sampled positions of the posterior distribution and define the weighted mean as

$$\left\langle \chi^2 \right\rangle = \sum_i w_i \chi^2_i \,, \tag{E.1}$$

where $\chi^2_i$ and $w_i$ correspond to the $\chi^2$ and corresponding weight for the $i$-th sampled position, respectively. As already mentioned in Sec. 4.1.2, this quantity can be estimated with less uncertainty from a sampled posterior distribution with respect to the maximum-likelihood position. However, the conclusions drawn in Sec. 5 are independent from which definition is used to quantify the goodness-of-fit statistics.

In Fig. E.1 we show the goodness-of-fit performance metric for the case of the Model 3 sample at $z = 1.2$, which is one of the few cases with a posterior-averaged $\chi^2$ getting above the $2\sigma$ in Fig. 9 and 10. Differently from those two plots, in this case we compare the trend when using the posterior-averaged $\chi^2$ (solid lines) and maximum-likelihood value (dash-dotted lines). As expected, the latter is constantly outperforming the former, with a relative improvement that becomes larger at low $k_{\rm max}$ values. Moving to larger $k_{\rm max}$, we note that for this specific case the maximum-likelihood position makes all the points consistent within $2\sigma$ of the considered $\chi^2$ distribution. However, the marginal gains are not significantly changing the interpretations that can be made when using the posterior-averaged values.

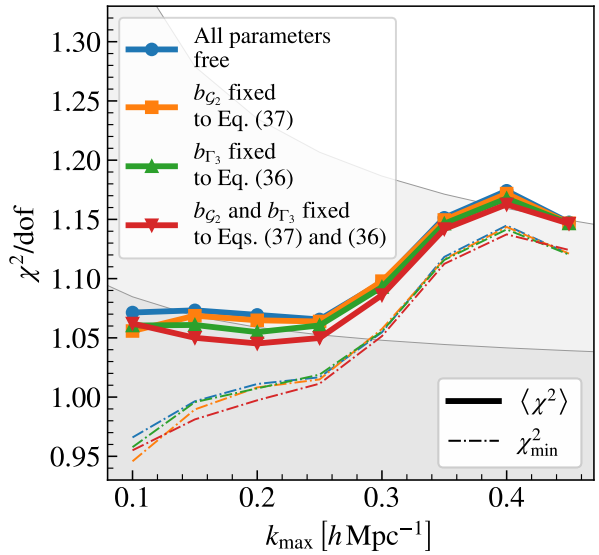


**Fig. E.1.** Normalised $\chi^2$ values for the Model 3 sample at $z = 1.2$. Solid and dot-dashed lines corresponds to the posterior-averaged and minimum $\chi^2$, respectively. Different colours represent different bias modelling assumtpions, as listed in the legend. The two grey shaded areas represent the 68% and 95% of the $\chi^2$ distribution with the same number of degrees of freedom of the considered model.

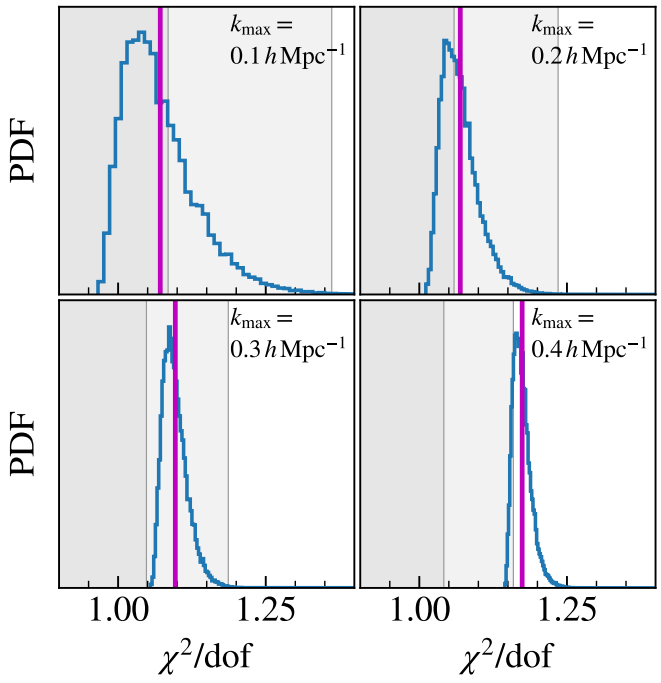


**Fig. E.2.** Normalised $\chi^2$ distribution for the Model 3 sample at $z = 1.2$ and four different value of $k_{max}$, as shown in the upper right corner of the individual panels. The shaded grey areas mark the same confidence intervals as in Fig. E.1. The magenta vertical line marks the posterior-averaged value.

As a complement, in Fig. E.2 we show the complete distribution of $\chi^2$ values, and we compare it with the 68th and 95th percentile of the reference $\chi^2$ distribution with the same numbers of degrees of freedom. Different panels correspond to different $k_{max}$ values, while the solid magenta line marks the position of the posterior-averaged value. The only configurations for which there is a partially significant gain in picking the maximum-likelihood value is clearly the one at $k_{max} = 0.4\,h\,Mpc^{-1}$. Otherwise, our conclusions are not affected by the choice of which $\chi^2$ value is used to estimate the goodness-of-fit performance metric. For this reason, in Sec. 5 we show results employing the posterior-averaged $\chi^2$ statistics.

## Appendix F: Dependency on the selected sampler

As anticipated in Sect. 4, in order to obtain a posterior distribution for each of the model configurations that we test, we need to select a robust algorithm to sample the large multi-dimensional parameter space of the models described in Sect. 3. Indeed, the choice of the sampler is critical in presence of multi-modal distributions. As already thoroughly discussed in the main body of this paper, this situation is prevalent when we focus solely on the full shape of the galaxy power spectrum, using only two-point statistics, due to the strong degeneracies between model parameters. In these situations traditional algorithms, which are meant to explore smaller and more well-behaved – Gaussian-like – random variables, may be under-performing, and therefore affecting the efficiency of the sampling.

In this analysis we compare three different samplers, testing them against a reference subset of the galaxy power spectrum data vectors already used to validate the theoretical model for one-loop galaxy bias. The properties of these sampling algorithms are listed hereafter.

### *Metropolis–Hastings sampler*

This approach (Metropolis et al. 1953; Hastings 1970) is a Markov-Chain Monte Carlo (MCMC) method based on the construction of a random walk inside the parameter space. Subsequent points in the Markov chain are determined based on a proposal function that has to be specified as a free parameter of the model. For each new point, the likelihood of the model determines whether the candidate state is accepted or discarded, that is, whether to move to the candidate state or stay in the current state. For this standard algorithm we make use of a non-public code,[35] which features an implementation of a Metropolis-Hastings algorithm coupled with the likelihood for the galaxy power spectrum described in Sect. 3.

### *Affine-invariant sampler*

This MCMC approach (Goodman & Weare 2010) is still based on a random walk across the parameter space, but with an improved ensemble sampler, in which a large number of walkers interact with each other in a way that reduces the dependence of the sampling on the aspect ratio of the particular posterior distribution under consideration. We use the affine-invariant sampler implemented in the public Python package `emcee` (Foreman-Mackey et al. 2013).[36]

### *Nested sampling*

This approach (Skilling 2006) relies on iteratively refining a set of live points, initially drawn from a prior distribution, to explore the parameter space. At each iteration, the point with the lowest likelihood is selected from the set and replaced with a new point. The latter is drawn from a constrained region of the prior distribution, where the likelihood must be higher than that of the replaced point. The selected point is then assigned a weight, based on its likelihood and the proportion of the prior volume it represents: several iterations will allow us to thus construct the posterior distribution. These algorithms are particularly suited for multi-modal distributions, given they are not subject to getting stuck in local minima of the loglikelihood, as it commonly happens with standard algorithms based on Markov chains. We use the public package `Multinest` (Feroz et al. 2009a, 2019b), which can be interfaced with Python using a dedicated wrapper module.[37]

### *Comparison*

The comparison between the marginalised two-dimensional distributions obtained with the different samplers is shown in Fig. F.1, using as data the galaxy power spectrum of the Model 3 HOD sample at $z = 0.9$ with $k_{max} = 0.2\,h\,Mpc^{-1}$. We test two different cases, one where we vary the nuisance parameters $(b_1, b_2, b_{\Gamma_3}, c_0, \alpha_{P,1})$ while keeping the quadratic tidal bias $b_{\mathcal{G}_2}$ fixed to its local Lagrangian relation (35), and one where the latter is also allowed to freely vary with the rest of the parameters.[38]

In the first case, we note an almost perfect match among the three different set of contours, with a statistically negligi-

[35] `COMPASS`, courtesy of Ariel G. Sánchez, Martín Crocce, and Román Scoccimarro.
[36] https://emcee.readthedocs.io/en/stable/
[37] https://johannesbuchner.github.io/PyMultiNest/
[38] As shown in the main body of this article, fixing $b_{\mathcal{G}_2}$ is the only case for which it is possible to completely break the strong degeneracies of the parameter space.

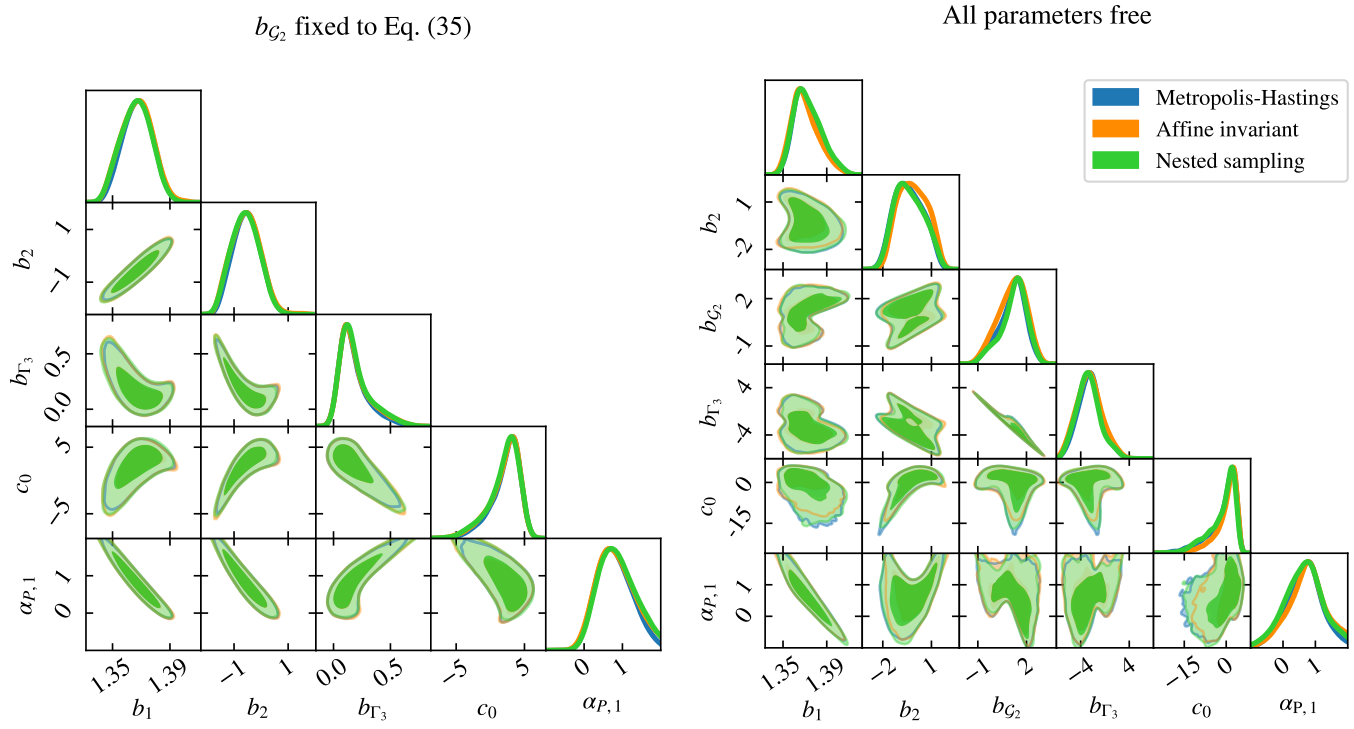


**Fig. F.1.** Comparison between runs at fixed cosmology obtained using different samplers, carried out using a reference galaxy power spectrum (Model 3 HOD sample at $z = 0.9$). The left panel shows the posterior distribution of the nuisance parameters with $b_{\mathcal{G}_2}$ fixed to the local Lagrangian relation (Eq. 35), while in the right panel all the parameters are left free to vary. Different colours correspond to different samplers, as listed in the legend.

ble difference only appreciable at the tails of the posterior distributions, which can be anyhow partially explained by the intrinsic variance of the individual realizations of the posterior distributions. The second case shows a slightly larger discrepancy, mostly due to the presence of a multi-modal profile driven by the strong degeneracy between $b_{\mathcal{G}_2}$ and $b_{\Gamma_3}$. In this case, the peak of the posterior distribution is less consistent when using different samplers, and we find that the size of the discrepancy is a very strong function of $k_{\rm max}$. In details, the main differences arise because of the non-homogeneous sampling of different peaks, for which traditional Markov chain algorithms may get stuck in a particular minimum of the distribution, therefore affecting the overall convergence of the chain. Given the purpose that nested sampling algorithms were initially developed for, we decided to employ the latter as our baseline sampler, and used it to sample the parameter space for all the results presented in this work.

## Appendix G: Sample variance effects

Even though the volume of the Flagship I simulation is quite large, most of our results are, to some extent, affected by the sample variance of the single $N$-body realization available.

In this section we quantify this effect reproducing some of the parameters fits starting from noiseless, synthetic data vectors obtained as the real-space galaxy power spectrum predictions from the `COMET` emulator. These are generated adopting the reference cosmological parameters from Table 1 at the four different redshifts of the comoving snapshots, $z = \{0.9, 1.2, 1.5, 1.8\}$. For each of the four samples, the nuisance parameters are derived as follows: $b_1$ and $\alpha_{P,1}$ are fixed to the best-fit value listed in Table 4 for the Model 1 HOD samples, $b_2$ is computed using a $b_2(b_1)$ relation derived from the corresponding HOD model, $b_{\mathcal{G}_2}$ and $b_{\Gamma_3}$ are obtained from the excursion-set relation (37) and the coevolution relation (36), respectively, $c_0$ is set to unity, and $\alpha_{P,2}$ is set to zero. The covariance matrix for each sample is computed assuming only the Gaussian component from the full Flagship I box, and therefore corresponds to a volume of about $58\,h^{-3}\,\mathrm{Gpc}^3$.

Figure G.1 shows the marginalised one-dimensional constraints of the $b_1$ and $\alpha_{P,1}$, similarly to Fig. 5, for fits at fixed cosmology of the synthetic data vectors. Even though the data vectors are smooth – no noise component has been added to any of the $P_{\rm gg}(k_i)$ bins – and also perfectly consistent with the theory model used to fit them, we observe some discrepancy of the constraints with the fiducial values, both with all parameters free as for the case with $b_{\Gamma_3}$ fixed to the coevolution relation. In both cases, the statistical constraints are less tight than the ones obtained with other bias relations, with a partial break of the degeneracies between parameters at higher redshift then the coevolution relation is applied. Such deviations are induced by projection – or prior volume – effects in the marginalised posteriors, due to the strong degeneracies between parameters. In particular, these affect $b_2$, $b_{\mathcal{G}_2}$ and $b_{\Gamma_3}$, and cannot be broken completely even when reducing the dimensionality of the parameter space by fixing $b_{\Gamma_3}$, as the $b_2$-$b_{\mathcal{G}_2}$ degeneracy is still present. Consistently among the different redshifts, we note that there is a trend for $b_1$ and $\alpha_{P,1}$ to be under- and overestimated, respectively, similarly to what we observe for the real Flagship I data vectors.

On the contrary, the case where $b_{\mathcal{G}_2}$ is fixed as a function of $b_1$ is systematically better in terms of amplitude of the error bars and accuracy in the recovery of the parameters, because the posterior distribution of this case is closer to a Gaussian distribution. This is due to the simultaneous breaking of both the $b_2$-$b_{\mathcal{G}_2}$ and $b_{\mathcal{G}_2}$-$b_{\Gamma_3}$ degeneracies, for which a clear example can be found in the left plot of Fig. F.1.

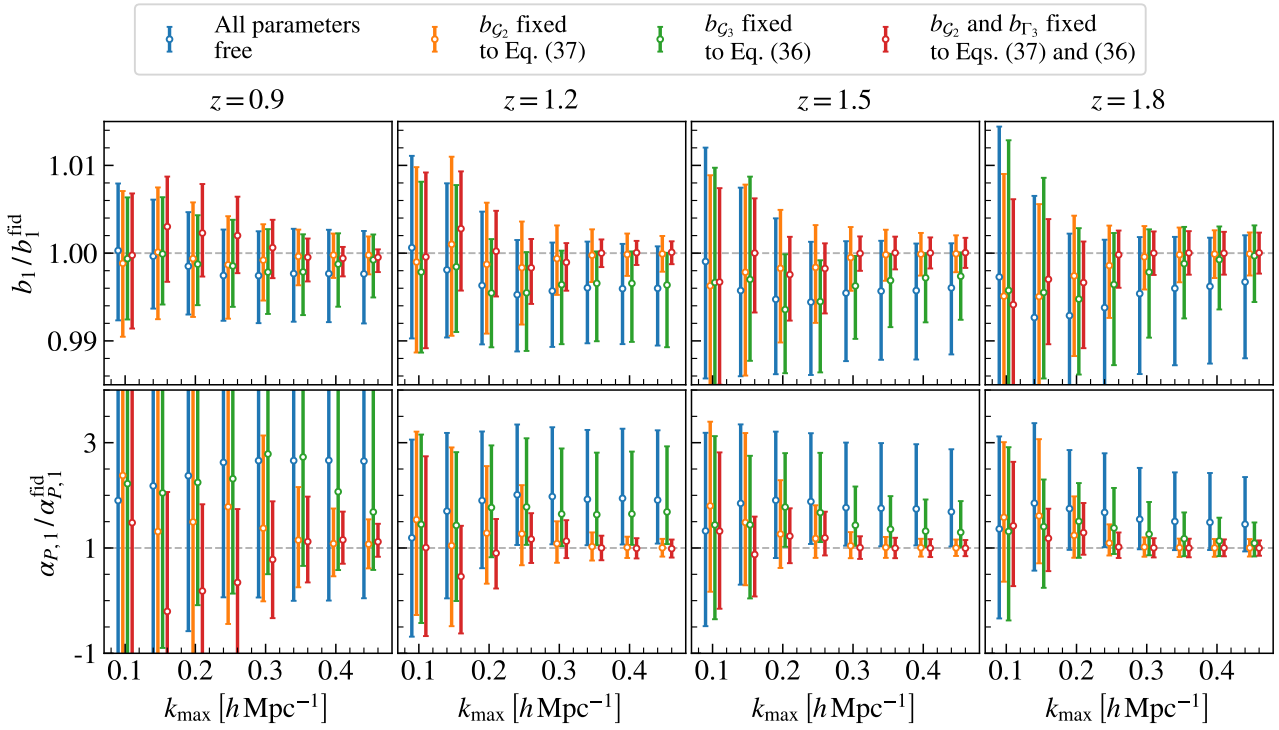


**Fig. G.1.** Marginalised 1d constraints on the linear bias $b_1$ and the shot-noise parameter $\alpha_{P,1}$ obtained from a set of synthetic theory vectors created using the same recipe from Eq. (34), for the four different redshifts already explored with the Flagship data vectors. Different colours correspond to different assumptions on the total number of degrees of freedom of the model, as listed in the legend.

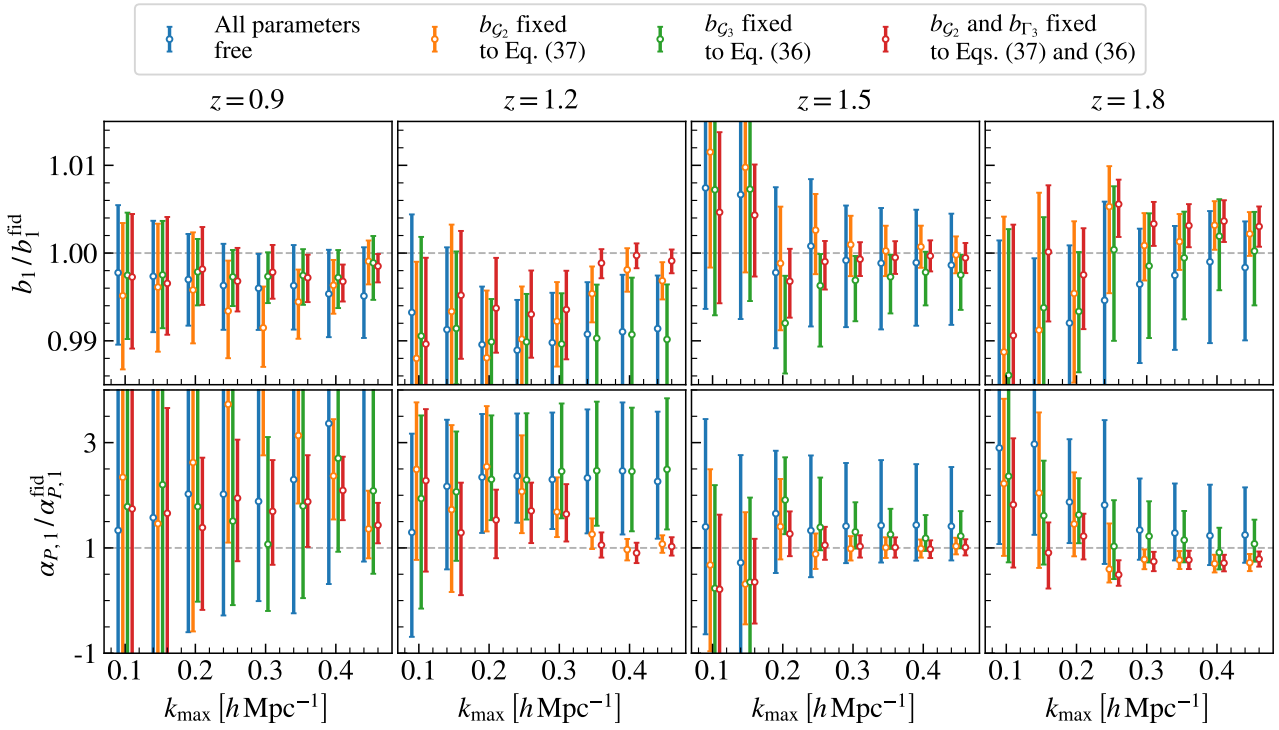


**Fig. G.2.** Same as in Fig. G.1, but using theory vectors displaced by Gaussian noise realisation from a box of size $3780\,h^{-1}$ Mpc.

As a follow-up test, we add to the synthetic data vectors a Gaussian noise consistent with the covariance assumed for the fits to the simulation. Figure G.2 shows the new marginalised constraints on $b_1$ and $\alpha_{P,1}$. In this case it is possible to observe a much larger discrepancy from the fiducial values, present also for the configurations in which $b_{\mathcal{G}_2}$ is fixed as a function of $b_1$. We remark how the case at $z = 1.2$ features a large fluctuation at intermediate scales, $k_{\rm max} \sim 0.2\,h\,{\rm Mpc}^{-1}$, which is symptomatic of the particular realization of the Gaussian noise. We repeated this exercise, on the sample at $z = 1.2$, with ten

different noise realizations, and in all cases we observe a slightly different trend as a function of $k_{\rm max}$ for the marginalised constraints of the configurations we are considering, that is, with all the parameters free to vary, or with either $b_{\mathcal{G}_2}$ or $b_{\Gamma_3}$ fixed in terms of $b_1$. For some of the configurations we observe trends as the one seen in the $z = 0.9$ Flagship I data vectors, for which we find a sharp running in the $(b_1, \alpha_{P,1})$ plane at $k_{\rm max} \gtrsim 0.2\,h\,{\rm Mpc}^{-1}$. We therefore conclude that a more realistic analysis, also including observational effects and a proper survey window function should be conducted using a wider set of simulations, in order to reduce the overall impact of cosmic variance.