Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **514,** 4363–4378 (2022)
Advance Access publication 2022 June 15
https://doi.org/10.1093/mnras/stac1544

Non-linear reconstruction of features in the primordial power spectrum from large-scale structure

Yuhao Li[1]★ Hong-Ming Zhu[2] and Baojiu Li[3]

[1]*Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Falmer, Brighton BN1 9RH, UK*
[2]*Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, Ontario M5S 3H8, Canada*
[3]*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*

Accepted 2022 June 1. Received 2022 June 1; in original form 2021 February 21

ABSTRACT

Potential features in the primordial power spectrum have been searched for in galaxy surveys in recent years since these features can assist in understanding the nature of inflation. The null detection to date suggests that any such features should be fairly weak, and next-generation galaxy surveys, with their unprecedented sizes and precisions, are in a position to place stronger constraints than before. However, even if such primordial features once existed in the early Universe, they would have been significantly damped in the non-linear regime at low redshift due to structure formation, which makes them difficult to be directly detected in real observations. A potential way to tackle this challenge for probing the features is to undo the cosmological evolution, i.e. using reconstruction to obtain an approximate linear density field. By employing a set of *N*-body simulations, here we show that a recently proposed non-linear reconstruction algorithm can effectively retrieve damped oscillatory features from halo catalogues and improve the accuracy of the measurement of feature parameters (assuming that such primordial features do exist). We do a Fisher analysis to forecast how non-linear reconstruction affects the constraining power, and find that it can lead to significantly more robust constraints on the feature amplitude for a DESI-like survey. Comparing non-linear reconstruction with other ways of improving constraints, such as increasing the survey volume and range of scales, this shows that it is possible to achieve what the latter do, but at a lower cost.

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

1 INTRODUCTION

Inflation, the most successful theory to solve the problems of the hot big bang model and to explain the seeding of the observed large-scale structures today, plays a crucial role in the development of modern cosmology. The single-field slow-roll inflation (Guth 1981; Albrecht & Steinhardt 1982; Linde 1982) predicts that primordial density fluctuations obey Gaussian statistics and the corresponding power spectrum follows a simple power law, which is favoured by the cosmic microwave background (CMB) data released by the WMAP (Peiris et al. 2003; Spergel et al. 2007; Komatsu et al. 2009; Hinshaw et al. 2013) and Planck (Ade et al. 2014a, 2016b; Akrami et al. 2020b) collaborations.

However, the physical origin of the inflation field, which is believed to have driven inflation, is not fully understood yet, and the fact that the very high energy scale in the early Universe makes it an ideal place to probe the imprints of the laws of fundamental physics offers the possibility that new physics can be revealed by cosmological observations of the large-scale structure (LSS). Certain sophisticated models of inflation and its alternatives developed over the last decades predict scale-dependent features in the power spectrum of primordial density fluctuations (see e.g. Bartolo et al. 2004; Chen 2010; Chluba, Hamann & Patil 2015; Slosar et al. 2019, for some reviews). Such 'feature models' can be mainly classified into three types with specific templates of oscillations added to the scale-invariant primordial power spectrum, each of which can be attributed to various mechanisms (see e.g. Chen 2010; Chluba et al. 2015; Slosar et al. 2019, for some reviews). 'Sharp-feature' models have sinusoidal wiggles in the power spectrum, $P(k)$, that oscillate linearly in wavenumber k at a fixed frequency, ω, which can be generated by a minimal local singularity such as a step in the inflationary potential that breaks the slow-roll condition (e.g. Starobinsky 1992; Adams, Cresswell & Easther 2001; Chen, Easther & Lim 2007; Hazra et al. 2010; Adshead et al. 2012; Hazra et al. 2014), or produced in particular cases of multifield models of inflation (e.g. Achucarro et al. 2011; Gao, Langlois & Mizuno 2012). Another type is the 'resonant-feature' model whose oscillatory features are in logarithmic k, which can be realised in, e.g. the axion monodromy inflation (Flauger et al. 2010; Flauger & Pajer 2011), or brane inflation (Bean et al. 2008), models. The last type is the so-called standard clock signal, which is a combination of the previous two feature models (e.g. Chen 2012; Chen & Namjoo 2014; Chen, Namjoo & Wang 2015).

These feature models have been continuously tested with the updated release of data from the Planck mission (Ade et al. 2014b, 2016a; Akrami et al. 2020a), but none of them has been found to be preferable to the scale-invariant power spectrum predicted by simple single-field slow-roll inflation models so far, which suggests

---

★ E-mail: yl700@sussex.ac.uk

Published by Oxford University Press on behalf of Royal Astronomical Society

that such features should be fairly weak if they do exist. Since the primordial features are not only imprinted in the CMB, but some of them can also leave a signature in the matter and galaxy distribution, future LSS surveys, such as *Euclid* (Racca et al. 2016), DESI (DESI Collaboration 2016), SPHEREx (Doré et al. 2014), and LSST (Ivezić et al. 2019), will provide the opportunity to search for, or tighten the constraints on, them, complementary to CMB data (e.g. Huang, Verde & Vernizzi 2012; Ballardini et al. 2016; Chen et al. 2016; Ballardini et al. 2018; L'Huillier et al. 2018; Palma, Sapone & Sypsas 2018; Zeng et al. 2019). More recently, this idea has been put into practice by making forecast (e.g. Beutler et al. 2019; Ballardini et al. 2020; Debono et al. 2020) or performing real LSS data analysis (Beutler et al. 2019).

However, any feature imprinted in the primordial density or curvature field by inflation is subject to the impact of cosmic evolution that last until today. In particular, even if such primordial features once existed in the very early Universe, they would have been modified in the late-time Universe due to non-linear structure formation. Meanwhile, the available information on large scales, where the evolution can be described by linear perturbation theory, is limited due to the cosmic variance, i.e. the poor statistics caused by the finite number of Fourier modes probed in that regime. This can affect the confidence level at which to measure or constrain these features. In order to maximally extract useful information from the observed galaxy distributions, several studies of the primordial features in the non-linear regime has been conducted. Vasudevan et al. (2019) and Beutler et al. (2019) analytically computed the damping effect by gravitational non-linearities, making a considerable contribution to the forecast of constraints on primordial feature from future galaxy surveys. Ballardini et al. (2020) employed *N*-body simulations to show a compatible non-linear damping effect with the analytical results above to leading order. Beutler et al. (2019) and Ballardini et al. (2020) made forecasts for future galaxy surveys by taking the damping effect into account. Besides, Beutler et al. (2019) performed the first LSS data analysis for the primordial features, which showed that LSS can surpass the CMB as a probe of such features. Furthermore, Vlah et al. (2016) and Chen, Vlah & White (2020) showed that different perturbation theories, including Lagrangian and Eulerian perturbation theories and the effective field theory, can model the non-linear evolution of primordial features well for $k \lesssim 0.25\ h\,\mathrm{Mpc}^{-1}$ at $z = 1$ and for $k \lesssim 0.2\ h\,\mathrm{Mpc}^{-1}$ at $z = 0$, but no oscillatory features survive past $k \approx 0.5\ h\,\mathrm{Mpc}^{-1}$. Thus, it would be beneficial to develop other approaches which can potentially allow us to exploit the LSS data in the range of scales, $0.2 \lesssim k(h\,\mathrm{Mpc}^{-1}) \lesssim 0.5$ even at low redshifts.

A potential method mentioned in Vasudevan et al. (2019), Ballardini et al. (2020) and implemented in Beutler et al. (2019) to address the issue of non-linear damping and further improve the constraints on primordial features is to undo the cosmological evolution in a process called reconstruction, which can partially retrieve the initial density field and therefore the information that existed there. A well-known example is the reconstruction of baryonic acoustic oscillation (BAO) features, which sharpens these features in the galaxy correlation function which provides a standard ruler for distance measurements (e.g. Eisenstein et al. 2007; Kazin et al. 2014; Schmittfull et al. 2015; Wang et al. 2017; Zhu et al. 2017; Shi, Cautun & Li 2018; Sarpa et al. 2019; Mao et al. 2021). While reconstructing the primordial power spectrum from observed galaxies has been shown to be beneficial for probing the primordial features from LSS data (Beutler et al. 2019), this study made use of one particular (the standard) reconstruction method, and it will be interesting to also assess how other reconstruction methods work in this regard.

In this work, as a first step towards assessing the potential benefit of non-linear reconstruction, we assume additional simple oscillatory features in the power-law primordial power spectrum. By utilizing a small set of *N*-body simulations, we study the performance of the non-linear reconstruction algorithm proposed recently by Shi et al. (2018) and Birkin et al. (2019) in retrieving the damped primordial features from the halo catalogues. In particular, by quantifying this damping caused by structure formation based on the functional form in Vasudevan et al. (2019) and Beutler et al. (2019), we will carry out parameter fittings to the damped and reconstructed wiggles, the comparison of which allows us to assess whether non-linear reconstruction can lead to more robust constraints on the feature parameters. To investigate the impact of non-linear reconstruction in real galaxy surveys, we also forecast the constraints on the feature parameters for a DESI-like survey using the Fisher matrix approach, and compare the cases with and without reconstruction.

This paper is organized as follows: in Section 2, we describe the model of primordial features, the simulations used in this work, and the methodology of assessing the performance of the non-linear reconstruction method to retrieve the damped primordial features due to structure formation. In Section 3, we give more details on the approach used to forecast the constraints on the feature parameters for the DESI-like survey. In Section 4, we show the results of non-linear reconstruction and forecast and discuss the implications of them. Finally, in Section 5 we conclude our findings and discuss potential future improvements.

2 METHODOLOGY

We start with presenting the primordial power spectrum models with oscillatory features that we adopt in this paper for illustration purpose. We then describe the simulation runs for these models. It is followed by a brief review of the non-linear reconstruction method which will be used to recover the small-scale oscillation features from evolved dark matter and halo fields. Finally, we describe the analytical model to quantify the features measured in the power spectrum before giving the details of the Fisher matrix forecast in the next section.

2.1 Models of featured primordial power spectrum

We take a power law-type primordial power spectrum to be our fiducial no-wiggle model (note that the BAO wiggles are still included), given by

$$P_{\mathrm{nw}}^{\mathrm{ini}}(k) = A_s \left(\frac{k}{k_*}\right)^{n_s - 1}, \tag{1}$$

where k is the comoving wavenumber, A_s and n_s are, respectively, the scalar amplitude and spectral index with the pivot scale given by $k_* = 0.05\ \mathrm{Mpc}^{-1}$. To explore whether the non-linear reconstruction algorithm employed in this paper can lead to improvements compared with the unreconstructed cases in Ballardini et al. (2020), we consider four wiggled models that are based on the template of the sharp feature model (Slosar et al. 2019), i.e. oscillations in linear k, given by

$$P_{\mathrm{w}}^{\mathrm{ini}}(k) = P_{\mathrm{nw}}^{\mathrm{ini}}(k)\left[1 + A\cos(\omega k^m + \phi)\right], \tag{2}$$

where A, ω, and ϕ are, respectively, the amplitude, frequency, and phase of the oscillation. We extend the sharp feature model by introducing m for a particular purpose explained later; when $m = 1$, equation (2) is related to the equation (2.1) in Ballardini et al. (2020).

Table 1. The oscillation parameters used for the no-wiggle model and four wiggled models. Columns respectively denote (1) the power of the comoving wavenumber; (2) the amplitude, (3) frequency, and (4) phase of the oscillation.

	m	A	ω (Mpc^{m})	ϕ/π
Fiducial		0		
Model 1	1	0.05	40	0
Model 2	1	0.05	70	0
Model 3	1	0.05	150	0
Model 4	0.631	0.05	28.9	0

Note that even if the primordial features exist, they could be more complicated than any phenomenological models that we are currently using. For now, we cannot determine the precise form of the features, thus we aim at something narrow, which is assuming that we know the functional form and verifying if non-linear reconstruction can improve the accuracy of measuring the feature parameters.

The oscillation parameters of the five models are listed in Table 1. Note that the frequencies of the wiggled models here are in units of Mpc^{m} due to m introduced above. The initial oscillations of the four wiggled models are shown in the red-dashed lines in the right-hand panel of Fig. 1, where we have presented the difference between $P_{\mathrm{w}}^{\mathrm{ini}}$ and $P_{\mathrm{nw}}^{\mathrm{ini}}$. Within our interested range of scales, $k = (0.05\text{–}0.5)\, h\,\mathrm{Mpc}^{-1}$, Model 1 has the first peak at the smallest scale, followed by Models 2 and 3, the frequency used in Model 3 is the same as BAO frequency. Model 4 is particularly adopted to have the first two peaks at the same positions of the first and third peaks of Model 2. The reason why this special model is designed will be explained in Section 4.2. By comparing the reconstructed wiggles of the four wiggled models later, we would be able to comprehend the effect of the non-linear reconstruction method on different scales.

2.2 *N*-body simulations

In the regime of linear perturbations, the primordial wiggles preserve their shapes and amplitude $P_{\mathrm{w}}^{\mathrm{ini}}/P_{\mathrm{nw}}^{\mathrm{ini}}$. However, non-linear large-scale structure evolution will change this behaviour, leading to damping of $P_{\mathrm{w}}^{\mathrm{ini}}/P_{\mathrm{nw}}^{\mathrm{ini}}$ at late times. This makes it harder to measure the properties of these primordial oscillations directly from an evolved density field, even more so for a late-time tracer (e.g. galaxy or halo) field. In order to quantify such effects, *N*-body cosmological simulations can prove to be a useful tool.

We have run five simulation runs including the no-wiggle model and four wiggled models. First, we assume a flat universe and adopt Planck 2018 cosmology, with $h = 0.674$, $\Omega_{\mathrm{m}} = 0.3135$, $\Omega_c h^2 = 0.120$, $\Omega_b h^2 = 0.0224$, $\Omega_\Lambda = 0.6865$, $n_s = 0.965$, and $A_s = 2 \times 10^{-9}$ (Aghanim et al. 2020). The value of σ_8 is approximately 0.79 though it varies a little bit across different models. We then customize the function of the primordial power spectrum in the Einstein–Boltzmann solver code CAMB (Lewis & Challinor 2011) to be equation (1) for the no-wiggle model and equation (2) for the wiggled models. We calculate the linear theory matter power spectrum at $z = 49$ using this version of the CAMB code, which is used as the input matter power spectrum for the publicly available code 2LPTic (Crocce, Pueblas & Scoccimarro 2006) to generate the initial conditions used for the *N*-body simulations. In the left-hand panel of Fig. 1, we compare the initial matter power spectrum given by CAMB and the matter power spectrum measured from the initial conditions generated using 2LPTic; it can be seen that they are in good agreement for all models within the range of scales of our interest (the blowing up at small scales is due to the finite particle resolution).

To more conveniently describe the oscillatory features for the wiggled models, as mentioned above, we define the relative wiggle pattern as

$$O_{\mathrm{rw}}^{\mathrm{ini}}(k) = \frac{P_{\mathrm{w}}^{\mathrm{ini}}(k)}{P_{\mathrm{nw}}^{\mathrm{ini}}(k)} - 1, \qquad (3)$$

which are shown in the right-hand panel of Fig. 1. This indicates that the oscillatory features have been reliably created in the initial conditions of the simulations within our interested range of scales, e.g. $k \lesssim 0.5\, h\,\mathrm{Mpc}^{-1}$.

Next, we run the simulations using the parallel *N*-body code RAMSES (Teyssier 2002) which is based on the adaptive mesh refinement (AMR) technique. Each simulation is performed with $N = 1024^3$ dark matter particles in a box of size $1024\, h^{-1}\,\mathrm{Mpc}$, and we output four snapshots at different redshifts, respectively, as $z = 0$, 0.5, 1, and 1.5. For each snapshot, we use the halo finder ROCKSTAR (Behroozi, Wechsler & Wu 2013) to identify the haloes with the definition of the halo mass M_{200c}, where M_{200c} is the mass within a sphere whose average density is 200 times the critical density. Since the low-mass haloes are unable to be fully probed due to the limited simulation resolution, we measure the cumulative halo mass functions (cHMFs) from the main haloes with more than 100 particles to check the validity of the simulation, which show very good agreement with the analytical formulae in Tinker et al. (2008). For each snapshot we establish one dark matter particle catalogue (hereafter DM) and two halo catalogues, respectively, with the number density of $1 \times 10^{-3} (h^{-1}\,\mathrm{Mpc})^{-3}$ (hereafter H1) and $5 \times 10^{-4} (h^{-1}\,\mathrm{Mpc})^{-3}$ (hereafter H2). Both host haloes and subhaloes are included in the halo catalogues. The number density of $5 \times 10^{-4} (h^{-1}\,\mathrm{Mpc})^{-3}$ is chosen to be an approximate value according to the current observations such as CMASS or LOWZ despite not being exactly the same, and $1 \times 10^{-3} (h^{-1}\,\mathrm{Mpc})^{-3}$ is a representative value of emission line galaxies (ELGs) in DESI survey; these choices are also somehow limited by the resolution of our simulations, though the use of the dark matter density field serves as a catalogue that has a much larger number density. Many realistic mock galaxy catalogues would give something between H1 and DM.

We achieve the number density by applying a mass cutoff, i.e. neglecting the haloes with smaller masses than the cutoff. By using the power spectrum estimator tool POWMES (Colombi & Novikov 2011), we measure the non-linear matter power spectrum from DM and non-linear halo power spectrum separately from H1 and H2. Finally, we take the ratio of the power spectrum of the wiggled models to the corresponding power spectrum of the no-wiggle model to obtain the quantity O_{rw} for all cases.

2.3 Reconstruction

In order to partially retrieve the primordial features damped during structure formation, we perform reconstruction of the initial density field from the late-time density field using the non-linear reconstruction algorithm described in Shi et al. (2018). This reconstruction method is based on mass conservation. Without assuming a cosmological model or having free parameters except the size of the mesh used to calculate the density field, it employs multigrid Gauss–Seidel relaxation to solve the non-linear partial differential equation which governs the mapping between the initial Lagrangian and final Eulerian coordinates of particles in evolved density fields. Previous tests show that the reconstructed density field is over ∼ 80 per cent correlated with the initial density field for $k \lesssim 0.6\, h\,\mathrm{Mpc}^{-1}$, if

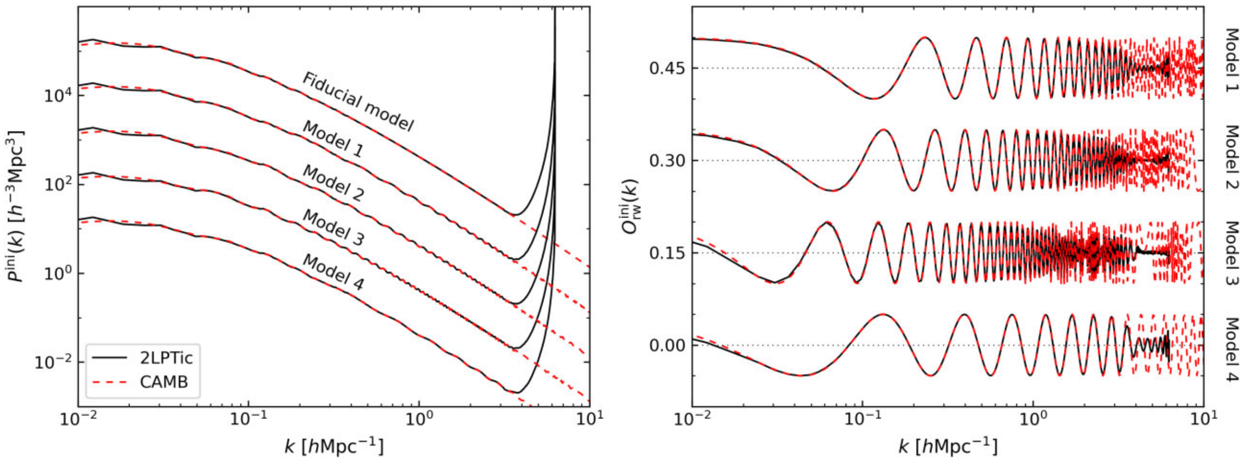

Figure 1. The left-hand panel shows the comparison between the initial matter power spectra given by CAMB (red dashed lines) and the matter power spectra measured from the initial conditions of the simulations generated using 2LPTic (black lines), from the top down they are, respectively, the fiducial model, Models 1, 2, 3, and 4, each model is shifted upwards by a factor of 10 successively to avoid the clutter of all curves. The right-hand panel shows the $O^{\rm ini}_{\rm rw}$ results, cf. equation (3), obtained from the left-hand panel for the four wiggled models, for instance, the top curve shows the ratio of Model 1 to the fiducial model, followed by the ones for Models 2, 3, and 4 downwards; each model is shifted upwards by a constant of 0.15 successively for the same reason as above.

reconstruction is performed on the dark matter density field, which cover the scales of our interest, but the performance becomes poorer when the method is instead applied on the density fields calculated from sparse tracers (Birkin et al. 2019; Liu, Yu & Li 2020; Wang, Li & Cautun 2020). This method is implemented in a modified version of the ECOSMOG code (Li et al. 2012, 2013), which itself is based on RAMSES.

We reconstruct the initial density field separately from the catalogues DM, H1 and H2 for each snapshot. The halo catalogues, which contain both main and subhaloes, are assumed to be the same as mock galaxy catalogues hereafter unless otherwise stated.[1] The procedure for the reconstruction from the halo catalogue is principally similar to that from the dark matter particle catalogue, apart from two things at the beginning. One is that we prepare the GADGET-format particle data for the ECOSMOG code in two ways. The halo catalogue is directly written into Gadget-format tracer particles due to its small number density. However, the very large number of the simulation particles, along with their strongly non-uniform spatial distribution, in the dark matter particle catalogues, leads to the requirement of large memory footprint when processing the data. To avoid this problem, we use the publicly available DTFE code (Cautun & van de Weygaert 2011), based on Delaunay tessellation, to calculate the density field on a regular mesh with 512^3 cells employing the triangular shaped cloud (TSC) mass assignment scheme; then the mesh cells are regarded as uniformly distributed fake particles with different masses, which are transformed to GADGET format that can be directly read by ECOSMOG.

The other particular thing is that we calculate the linear halo bias used for the reconstruction from the halo catalogue. The estimate of the halo bias is based on the relation

$$b_1(r) = \frac{\xi_{\rm hh}(r)}{\xi_{\rm hm}(r)}, \tag{4}$$

where $\xi_{\rm hh}(r)$ is the autocorrelation function of haloes and $\xi_{\rm hm}(r)$ is the cross-correlation function between the haloes and the dark matter particles. We use the publicly available CUTE code (Alonso 2012) to measure $\xi_{\rm hh}(r)$ and $\xi_{\rm hm}(r)$ from a given simulation snapshot, and take the ratio between them to obtain the value of linear halo bias as a function of the distance r. Since the linear halo bias is theoretically a constant on large scales, we apply the method of least squares to the values on scales $r \gtrsim 10\ h^{-1}$ Mpc to obtain an estimate of it. Note that when dealing with observational data we do not necessarily have such an accurate measurement of the linear halo or galaxy bias; however, Birkin et al. (2019) find that the exact value of linear bias is not very important for this reconstruction method to recover the phases of the initial density field.

The following steps of reconstruction are then the same for both dark matter particle catalogue and halo catalogues. First, ECOSMOG calculates the density field in the Eulerian coordinates using the TSC mass assignment scheme, and solves the mapping between the Eulerian and Lagrangian coordinates, to get the displacement potential as well as the displacement field on a regular mesh with 512^3 cells. We then use a PYTHON code to transfer the output fields from the Eulerian coordinates to the Lagrangian coordinates. After that, because the Lagrangian coordinates are not uniform, we feed the DTFE code with the Lagrangian coordinates and displacement field of the mesh cells to calculate the reconstructed density field as the divergence of the displacement field w.r.t. the Lagrangian coordinates. Finally, we measure the reconstructed power spectrum from the reconstructed density field using a post-processing code.

2.4 Parameter fitting to the damped wiggles

As we discussed above, cosmic structure formation leads to damping of the primordial wiggles. Reconstruction is expected to revert some of this damping, but cannot completely undo it. So we need a model for the wiggles of the reconstructed matter or halo power spectrum. Ideally this should be an analytical model since it can be more easily used in the Fisher analysis later. In this section, we describe how this is achieved by using a fitting function.

A functional form of the feature damping is analytically computed in Vasudevan et al. (2019) and Beutler et al. (2019) to be a Gaussian. We combine it with the oscillatory feature model described above, in order to directly fit the wiggle pattern $O_{\rm rw}$. The fitting function that

---

[1] As a result, we will use 'haloes' and 'galaxies' interchangeable throughout the rest of this paper: 'galaxies' will be used where we refer to observational quantities, while 'haloes' will be used for simulated quantities.

is used to described the damped wiggles is given by

$$O^{\rm fit}_{\rm rw}(k, z) = A\cos(\omega k^m + \phi)\exp\left[-\frac{k^2\zeta(z)^2}{2}\right], \tag{5}$$

where $\zeta(z)$ is the damping parameter that depends on the redshift z. For the fitting of each measured $O_{\rm rw}(k) = P_{\rm w}(k)/P_{\rm nw}(k) - 1$, we let ω, ϕ, and ζ be the free parameters because ω and ϕ play an essential role in determining the position of the peaks, and ζ quantifies the extent of the damping effect. The parameters A and m are taken to be their theoretical values in Table 1. In principle, A is also a free parameter here and should be allowed to vary in our parameter fitting. We have explicitly checked this 4-parameter fitting and found that, compared with the 3-parameter fitting, in the vast majority of cases of Table 2, the best-fitting values of ω and ϕ are not more accurate, which is as expected. There is a degeneracy between the amplitude A and the damping scale ζ, with the fitted values of the latter having larger uncertainties in the case of the 4-parameter fitting. Since for our forecast work the value of ζ is more important, we stick with the results obtained from the 3-parameter fitting.

We apply the least-squares estimator to obtain the best-fitting parameters by minimizing

$$\chi^2 = \sum_{i=1}^{N}\left[O_{\rm rw,i}(z) - O^{\rm fit}_{\rm rw}(k_{\rm i}, z; \omega, \phi, \zeta)\right]^2, \tag{6}$$

where $O_{\rm rw,\,i}(z)$ are the data points of wiggle spectrum in the ith k bin at reshift z. Since there is only one realization of simulation for each model, we assume that the uncertainties of all data points $O_{\rm rw,\,i}(z)$ are the same and follow the same Gaussian distribution. Note that, as the quantity we fit is $O_{\rm rw} = P_{\rm w}/P_{\rm nw} - 1$, this is equivalent to doing the fitting of $P_{\rm w}$ with $\sqrt{P_{\rm nw}}$ as uncertainty (e.g. Feldman, Kaiser & Peacock 1994).

We calculate the uncertainties of the best-fitting parameters based on 95 per cent confidence interval, as a rough estimate of the size of the errors. To minimize the influence of the cosmic variance on very large scales, we fit the data within the interval of $k = (0.04-0.6)\ h\,{\rm Mpc}^{-1}$, which covers our intended range of scales.

3 FORECAST FOR THE DESI-LIKE SURVEY

In order to investigate the impact of reconstruction, we will forecast the constraints on the feature parameters for the DESI-like survey using the Fisher information matrix, and compare with the case of doing no reconstruction. For this purpose, we first model the observed broad-band galaxy power spectrum. Then we describe how to calculate the Fisher information matrix, followed by its analytical marginalization. Finally, we give the specifications of the DESI-like survey.

3.1 Modelling the observed galaxy power spectrum

Combining the equations (3) and (5), the featured non-linear matter power spectrum in real space can be modelled as

$$P_{\rm mod}(k, z) = P_{\rm nl}(k, z)\left[1 + A\cos(\omega k^m + \phi)\exp\left(-\frac{k^2\zeta(z)^2}{2}\right)\right], \tag{7}$$

where $P_{\rm nl}(k, z)$ is the non-linear matter power spectrum without the primordial oscillatory features at z, which includes the BAO wiggles and is equivalent to the non-linear matter power spectrum of the no-wiggle model. However, since there is only one simulation realization for a single no-wiggle model, which cannot provide a smooth non-linear matter power spectrum, and since a fast method to get $P_{\rm mod}$ is more convenient in the Fisher analysis, we use the HALOFIT model in the CAMB code to calculate $P_{\rm nl}(k, z)$ instead later in this work. We have checked that the fractional difference between the simulated no-wiggle power spectrum and the one computed by HALOFIT is below 10 per cent within the entire fitting range.

The broad-band galaxy power spectrum in real space is not a direct observable due to the measurement in the angular and redshift coordinates instead of the 3D comoving coordinates. In order to relate the observed galaxy power spectrum $P_{\rm obs}(\boldsymbol{k}, z)$ to the modelled matter power spectrum $P_{\rm mod}(k, z)$, the standard practice is to project the galaxies to their comoving positions assuming some reference cosmology via the coordinate transformation based on the relations

$$k^{\rm ref}_{\perp} = \frac{D_{\rm A}(z)}{D^{\rm ref}_{\rm A}(z)}k_{\perp}, \quad k^{\rm ref}_{\parallel} = \frac{H^{\rm ref}(z)}{H(z)}k_{\parallel}, \tag{8}$$

where $k_{\parallel}$ and $k_{\perp}$ are, respectively, the line of sight and transverse components of the wavevector $\boldsymbol{k}$, i.e. $k^2 = |\boldsymbol{k}|^2 = k^2_{\perp} + k^2_{\parallel}$, the superscript $^{\rm ref}$ denotes the reference cosmology, note that the reference cosmology hereafter is the same one used in the simulations unless otherwise stated; $D_{\rm A}(z) = r(z)/(1+z)$ is the angular diameter distance at z with the comoving distance $r(z)$: under the assumption of flat universe it is given by

$$r(z) = \frac{c}{H_0}\int_0^z {\rm d}z'\left[\Omega_{\rm m}(1+z)^3 + \Omega_{\Lambda}\right]^{-\frac{1}{2}}, \tag{9}$$

where $\Omega_{\Lambda} = 1 - \Omega_{\rm m}$ is the current density parameter of the cosmological constant, and the Hubble parameter $H(z)$ is given by

$$H(z) = H_0\left[\Omega_{\rm m}(1+z)^3 + \Omega_{\Lambda}\right]^{\frac{1}{2}}. \tag{10}$$

Along with several main factors being considered, i.e. the redshift-space distortions (RSD) and shot noise, one can model the observed galaxy power spectrum as

$$P_{\rm obs}(k, \mu, z) = \left[\frac{D^{\rm ref}_{\rm A}(z)}{D_{\rm A}(z)}\right]^2 \frac{H(z)}{H^{\rm ref}(z)}\frac{F_{\rm FoG}(k, \mu, z)}{\sigma^2_8(z)}P_{\rm mod}(k, z) + N_{\rm gal}(z), \tag{11}$$

where $\sigma_8(z)$ is the R.M.S. linear density fluctuations on the scale of $8\,h^{-1}$Mpc, $N_{\rm gal}(z) = 1/\overline{n}_{\rm g}(z)$ is the shot noise with $\overline{n}_{\rm g}(z)$ being the galaxy number density, and the Finger-of-God factor $F_{\rm FoG}(k, \mu, z)$ describing the effect of RSD is modelled as Ballardini et al. (2020)

$$F_{\rm FoG}(k, \mu, z) = \frac{\left[b(z)\sigma_8(z) + f(z)\sigma_8(z)\mu^2\right]^2}{1 + k^2\mu^2\sigma^2_{r,p}/2}\exp\left(-k^2\mu^2\sigma^2_{r,z}\right), \tag{12}$$

where we have included the linear halo bias at z, $b(z)$, to make $P_{\rm obs}$ the 'galaxy' [remember that in our simulations we treat (sub)haloes as mock galaxies] power spectrum, and

$$f(z) = \frac{{\rm d}\ln D(a)}{{\rm d}\ln a}, \tag{13}$$

is the linear growth rate at z with $D(a)$ and a respectively being the linear growth factor and the scale factor (note that we normalize $D(a)$ so that $D(a=1) = 1$ in this work), $\mu = \cos\theta$ with θ being the angle between the wavevector $\boldsymbol{k}$ and the line of sight, i.e. $\mu = k_{\parallel}/k$, $\sigma_{r,\,p} = \sigma_p/[H(z)a]$ is the distance dispersion corresponding to the physical velocity dispersion σ_p whose fiducial value is taken to be 290 km s^{-1}. The last exponential factor represents an additional damping to account for the observational redshift error $\sigma(z)$ with $\sigma_{r,\,z} = c\sigma(z)/H(z)$ specific to a given survey, which is very close to 1 for our intended range of scales given that the DESI survey assumes

Table 2. The best-fitting parameters of ω, ϕ, and ζ and their 95% uncertainties for the four wiggled models studied in this work. The values of ω and ϕ are, respectively, in the units of Mpc^m and π, and their theoretical values are shown below the title of each model on the top of the table. DM denotes the dark matter particle catalogue, H1 the halo catalogue with $n_{\rm halo} = 1 \times 10^{-3} (h^{-1}\,{\rm Mpc})^{-3}$ and H2 the halo catalogue with $n_{\rm halo} = 5 \times 10^{-4} (h^{-1}\,{\rm Mpc})^{-3}$. Each group of six rows includes the unreconstructed and reconstructed cases for the same redshift.

				Model 1			Model 2	
				$\omega = 40, \phi = 0$			$\omega = 70, \phi = 0$	
z		para	DM	H1	H2	DM	H1	H2
0.0	unrec	ω	33.0 ± 3.4	36.5 ± 1.6	36.2 ± 1.9	68.5 ± 1.1	69.0 ± 1.7	68.4 ± 1.7
		ϕ	0.09 ± 0.08	0.02 ± 0.04	0.04 ± 0.05	0.00 ± 0.03	0.01 ± 0.05	0.00 ± 0.04
		ζ	7.23 ± 0.95	6.21 ± 0.44	6.90 ± 0.53	7.43 ± 0.26	6.95 ± 0.40	7.27 ± 0.41
	rec	ω	40.0 ± 0.1	40.3 ± 0.4	40.8 ± 1.1	70.1 ± 0.1	70.1 ± 1.0	69.9 ± 0.9
		ϕ	0.00 ± 0.01	− 0.02 ± 0.02	− 0.04 ± 0.04	− 0.01 ± 0.01	0.00 ± 0.05	0.01 ± 0.04
		ζ	2.05 ± 0.05	3.41 ± 0.13	4.07 ± 0.32	2.06 ± 0.03	3.58 ± 0.28	4.00 ± 0.25
0.5	unrec	ω	37.6 ± 1.4	37.3 ± 1.1	37.7 ± 1.0	69.1 ± 0.4	68.0 ± 0.9	68.2 ± 1.2
		ϕ	0.02 ± 0.04	0.03 ± 0.03	0.02 ± 0.03	0.00 ± 0.01	0.05 ± 0.03	0.02 ± 0.04
		ζ	5.53 ± 0.40	5.27 ± 0.32	5.71 ± 0.29	5.88 ± 0.10	5.72 ± 0.23	6.18 ± 0.30
	rec	ω	40.0 ± 0.1	40.1 ± 0.3	40.3 ± 0.7	70.1 ± 0.1	70.1 ± 0.4	69.7 ± 0.6
		ϕ	0.00 ± 0.01	− 0.01 ± 0.02	− 0.01 ± 0.03	− 0.01 ± 0.01	0.00 ± 0.02	0.01 ± 0.03
		ζ	1.52 ± 0.04	3.17 ± 0.10	3.75 ± 0.20	1.51 ± 0.04	3.12 ± 0.10	3.83 ± 0.17
1.0	unrec	ω	38.6 ± 0.6	38.0 ± 0.7	37.5 ± 0.9	69.5 ± 0.2	68.7 ± 0.8	69.0 ± 0.8
		ϕ	0.01 ± 0.02	0.01 ± 0.03	0.01 ± 0.03	0.00 ± 0.01	0.02 ± 0.03	0.00 ± 0.03
		ζ	4.39 ± 0.18	4.38 ± 0.22	5.08 ± 0.26	4.73 ± 0.05	4.90 ± 0.21	5.48 ± 0.20
	rec	ω	40.0 ± 0.1	40.1 ± 0.3	39.7 ± 0.7	70.0 ± 0.1	70.0 ± 0.5	70.0 ± 0.7
		ϕ	0.00 ± 0.01	− 0.01 ± 0.02	0.01 ± 0.03	0.00 ± 0.01	0.00 ± 0.02	− 0.01 ± 0.03
		ζ	1.09 ± 0.04	3.09 ± 0.11	3.56 ± 0.20	1.11 ± 0.03	3.23 ± 0.14	3.81 ± 0.20
1.5	unrec	ω	39.3 ± 0.3	38.1 ± 0.7	36.5 ± 0.8	69.6 ± 0.1	69.1 ± 0.8	68.2 ± 1.0
		ϕ	0.00 ± 0.01	0.02 ± 0.03	0.07 ± 0.03	0.00 ± 0.01	0.01 ± 0.03	0.03 ± 0.03
		ζ	3.60 ± 0.10	4.07 ± 0.20	4.92 ± 0.25	3.90 ± 0.02	4.43 ± 0.21	5.17 ± 0.26
	rec	ω	40.0 ± 0.1	39.6 ± 0.5	40.1 ± 1.1	70.0 ± 0.1	70.2 ± 0.4	69.6 ± 1.1
		ϕ	0.00 ± 0.01	0.02 ± 0.02	− 0.01 ± 0.05	0.00 ± 0.01	− 0.01 ± 0.02	0.01 ± 0.05
		ζ	0.84 ± 0.03	3.08 ± 0.15	3.66 ± 0.32	0.88 ± 0.03	3.24 ± 0.12	3.90 ± 0.29
				Model 3			Model 4	
				$\omega = 150, \phi = 0$			$\omega = 28.9, \phi = 0$	
z		para	DM	H1	H2	DM	H1	H2
0.0	unrec	ω	149.3 ± 0.4	151.2 ± 2.0	151.6 ± 2.1	26.1 ± 0.6	27.2 ± 1.0	25.9 ± 1.0
		ϕ	0.01 ± 0.01	− 0.04 ± 0.05	− 0.05 ± 0.04	0.11 ± 0.04	0.06 ± 0.06	0.14 ± 0.06
		ζ	7.95 ± 0.09	7.44 ± 0.51	8.51 ± 0.50	6.96 ± 0.23	6.62 ± 0.35	6.68 ± 0.36
	rec	ω	150.1 ± 0.1	150.7 ± 0.6	150.7 ± 1.0	29.0 ± 0.1	29.5 ± 0.4	28.7 ± 0.6
		ϕ	− 0.01 ± 0.01	− 0.04 ± 0.03	− 0.03 ± 0.04	− 0.01 ± 0.01	− 0.05 ± 0.03	0.03 ± 0.05
		ζ	2.13 ± 0.04	3.70 ± 0.19	4.04 ± 0.28	2.11 ± 0.04	3.41 ± 0.14	3.75 ± 0.22
0.5	unrec	ω	149.6 ± 0.2	150.4 ± 0.9	150.4 ± 1.5	27.2 ± 0.2	27.2 ± 0.6	26.2 ± 0.7
		ϕ	0.01 ± 0.01	− 0.01 ± 0.02	− 0.02 ± 0.03	0.07 ± 0.02	0.06 ± 0.05	0.12 ± 0.05
		ζ	6.30 ± 0.05	6.02 ± 0.23	6.85 ± 0.39	5.33 ± 0.08	4.96 ± 0.23	5.58 ± 0.26
	rec	ω	150.1 ± 0.1	150.0 ± 0.5	150.1 ± 0.8	29.0 ± 0.1	29.1 ± 0.3	29.3 ± 0.4
		ϕ	− 0.01 ± 0.01	− 0.01 ± 0.02	− 0.02 ± 0.03	− 0.01 ± 0.01	− 0.03 ± 0.03	− 0.03 ± 0.04
		ζ	1.63 ± 0.03	3.37 ± 0.15	3.96 ± 0.23	1.53 ± 0.04	3.04 ± 0.11	3.46 ± 0.16
1.0	unrec	ω	149.7 ± 0.1	149.9 ± 0.9	149.3 ± 0.9	27.7 ± 0.1	27.1 ± 0.5	26.7 ± 0.5
		ϕ	0.01 ± 0.01	0.00 ± 0.03	0.02 ± 0.03	0.05 ± 0.01	0.08 ± 0.04	0.10 ± 0.03
		ζ	5.08 ± 0.03	5.29 ± 0.25	5.95 ± 0.24	4.19 ± 0.03	4.30 ± 0.20	4.99 ± 0.17
	rec	ω	150.1 ± 0.1	150.6 ± 1.7	149.9 ± 0.6	29.0 ± 0.1	28.7 ± 0.4	28.3 ± 0.4
		ϕ	− 0.01 ± 0.01	− 0.04 ± 0.08	0.00 ± 0.03	− 0.01 ± 0.01	0.02 ± 0.03	0.04 ± 0.04
		ζ	1.27 ± 0.03	3.31 ± 0.52	3.85 ± 0.17	1.12 ± 0.03	3.04 ± 0.13	3.55 ± 0.15
1.5	unrec	ω	149.7 ± 0.1	149.6 ± 0.8	149.6 ± 0.9	28.0 ± 0.1	27.4 ± 0.6	27.0 ± 0.6
		ϕ	0.01 ± 0.01	0.01 ± 0.03	0.01 ± 0.03	0.04 ± 0.01	0.05 ± 0.05	0.08 ± 0.04
		ζ	4.21 ± 0.02	4.88 ± 0.23	5.71 ± 0.24	3.41 ± 0.02	3.92 ± 0.22	4.73 ± 0.21
	rec	ω	150.1 ± 0.1	150.0 ± 0.7	150.0 ± 1.2	28.9 ± 0.1	29.0 ± 0.4	28.9 ± 0.5
		ϕ	− 0.01 ± 0.01	− 0.01 ± 0.03	0.00 ± 0.05	0.00 ± 0.01	− 0.02 ± 0.04	− 0.01 ± 0.04
		ζ	1.07 ± 0.03	3.38 ± 0.20	4.07 ± 0.34	0.83 ± 0.02	3.19 ± 0.16	3.65 ± 0.17

$\sigma(z) = 0.0005/(1 + z)$ (DESI Collaboration 2016), so we neglect it in the calculation.

Additional effects involved in real observational constraints, such as the survey window function and finite bandwidths, which would influence the forecasted constraining power to some extent (see e.g. Beutler et al. 2019, for a more detailed discussion), should be taken into account when dealing with real surveys in future works, but these are not included in the forecast here. The present work is therefore a simplified proof-of-concept study which is likely to lead to optimistic forecasts.

3.2 Fisher information matrix

The Fisher matrix approach provides a method to propagate the uncertainties of the observable to the constraints on the cosmological parameters. Our calculation of the Fisher matrix is based on Tegmark (1997) and Seo & Eisenstein (2003), assuming that the power spectrum of a given k mode satisfies a Gaussian distribution which has a variance equal to the power spectrum itself, and that different bins of k are independent of each other for large surveys, the Fisher matrix for each redshift bin, with bin centre at $z = z_{\rm c}$, can be approximated as

$$F_{ij}(z_{\rm c}) = \frac{V_{\rm eff}(z_{\rm c})}{4\pi^2}\int_0^1 {\rm d}\mu \times \int_{k_{\rm min}}^{k_{\rm max}} {\rm d}k k^2 \frac{\partial \ln P_{\rm obs}(k,\mu,z_{\rm c})}{\partial\theta_i}\frac{\partial \ln P_{\rm obs}(k,\mu,z_{\rm c})}{\partial\theta_j}, \qquad (14)$$

where $k_{\rm min}$, $k_{\rm max}$ are, respectively, the minimum and maximum values of k used for the forecast. We set $k_{\rm min} = 0.05\,h{\rm Mpc}^{-1}$ and adopt two values of $k_{\rm max}$, respectively, 0.25 and 0.5 $h{\rm Mpc}^{-1}$, to compare the constraints for different ranges of scales. The effective volume of the redshift bin $V_{\rm eff}(z_{\rm c})$ is expressed as

$$V_{\rm eff}(z_{\rm c}) = \left[1 + \frac{1}{\overline{n}_{\rm g}(z)P_{\rm obs}(k,\mu,z)}\right]^{-2} V_{\rm surv}(z_{\rm c}), \qquad (15)$$

where $\overline{n}_{\rm g}(z)P_{\rm obs}(k,\mu,z)$ is the signal to noise, the comoving survey volume $V_{\rm surv}(z_{\rm c})$ with the redshift bin width Δz is given by

$$V_{\rm surv}(z_{\rm c}) = \frac{4\pi}{3}\left[r\left(z_{\rm c} + \frac{\Delta z}{2}\right)^3 - r\left(z_{\rm c} - \frac{\Delta z}{2}\right)^3\right]\frac{\Omega_{\rm surv}}{\Omega_{\rm sky}}, \qquad (16)$$

where $\Omega_{\rm surv}$ and $\Omega_{\rm sky}$ are, respectively, the survey area and the area of the full sky. Additionally, θ is the 8D parameter vector which consists of five cosmological parameters and three oscillation parameters

$$\omega_{\rm c} = \Omega_{\rm c}h^2, \omega_{\rm b} = \Omega_{\rm b}h^2, h, n_{\rm s}, A_{\rm s}, A, \omega, \phi. \qquad (17)$$

The partial derivatives of $P_{\rm obs}(k,\ \mu,\ z_{\rm c})$ w.r.t. the cosmological parameters are calculated numerically using the finite difference

$$\frac{\partial P_{\rm obs}(k,\mu,z_{\rm c})}{\partial\theta_i} = \frac{P_{\rm obs}(\theta_i^{\rm fid} + \Delta\theta_i) - P_{\rm obs}(\theta_i^{\rm fid} - \Delta\theta_i)}{2\Delta\theta_i}, \qquad (18)$$

where $\Delta\theta_i$ is taken to be 10 per cent of the fiducial value of $\theta_i^{\rm fid}$, though we have explicitly checked that the partial derivative is insensitive to the size of $\Delta\theta_i$. By contrast, the partial derivatives w.r.t. the oscillation parameters can be calculated analytically due to the analytical form of the oscillations.

The Fisher matrices of the different redshift bins are summed up to get a 8 × 8 matrix, and then we can calculate the covariance matrix by taking the inverse of this Fisher matrix and the uncertainties of the parameters are given by the square roots of its diagonal elements. Since we are mainly interested in the constraints on the oscillation parameters, we marginalize the cosmological parameters using the analytical marginalization method given by Taylor & Kitching (2010), which marginalizes the nuisance parameters and preserves the information about the target parameters. The marginalized Fisher matrix is given by

$$F^{\rm M}_{\alpha\beta} = F_{\alpha\beta} - F_{\alpha m}F^{-1}_{mn}F_{n\beta}, \qquad (19)$$

where the subscripts α and β denote the target parameters, while the subscripts m and n denote the nuisance parameters. Finally, we get the uncertainties of the oscillation parameters from the marginalized Fisher matrix.

3.3 Parameters used in the Fisher analysis

The parameters used in the Fisher analysis, including those associated with the specifications of the DESI-like survey (DESI Collaboration 2016) are discussed here.

We start with the most crucial parameter, the damping parameter ζ displayed in Table 2, which depends not only on the redshifts but also on the halo number densities and – more importantly – whether the reconstruction is applied. We only have values of ζ for four redshifts, i.e. $z = 0$, 0.5, 1, 1.5, and two different halo number densities, i.e. $n_{\rm halo} = 1 \times 10^{-3}(h^{-1}\,{\rm Mpc})^{-3}$ and $5 \times 10^{-4}(h^{-1}\,{\rm Mpc})^{-3}$, but the forecasted number density achievable in the DESI-like survey varies over the redshift range, so the values of ζ may not apply to the entire redshift range. As a result, we cut off some high-redshift bins which have the number density much smaller than $5 \times 10^{-4}(h^{-1}\,{\rm Mpc})^{-3}$. We use a bilinear interpolation between the redshift and the number density to estimate an appropriate value of ζ for a given combination of the redshift and number density. For those the number density is larger than $1 \times 10^{-3}(h^{-1}\,{\rm Mpc})^{-3}$ or smaller than $5 \times 10^{-4}(h^{-1}\,{\rm Mpc})^{-3}$, we simply adopt the values of ζ for $n_{\rm halo} = 1 \times 10^{-3}(h^{-1}\,{\rm Mpc})^{-3}$ or $n_{\rm halo} = 5 \times 10^{-4}(h^{-1}\,{\rm Mpc})^{-3}$ instead. In this work, we use different values of ζ for the different models as obtained using the fitting method described in Section 2.4, and we will comment on this point again later.

As we consider both emission line galaxies (ELGs) and luminous red galaxies (LRGs) in the DESI-like survey, which have different number densities and redshift distributions, different range of redshift bins is chosen for ELGs and LRGs in the Fisher analysis. After throwing away the redshift bins with very small number densities, we take the range of $z = (0.6\text{–}1.3)$ for ELGs and $z = (0.6\text{–}0.9)$ for LRGs, and the redshift bin width is by default $\Delta z = 0.1$. In addition to the calculation of effective survey volume, by following the DESI-like survey, the fixed values of $\overline{n}_{\rm g}(z)P_{\rm obs}(0.14, 0.6, z)$ are used for the signal to noise, two survey areas are considered including the expected survey area of 14 000 and 9000 deg^2 as the pessimistic case (DESI Collaboration 2016). As for the Finger-of-God factor, the linear halo bias for ELGs and LRGs is simply defined in terms of the growth factor via (DESI Collaboration 2016)

$$b_{\rm ELG}(z)D(z) = 0.84 \quad \text{and} \quad b_{\rm LRG}(z)D(z) = 1.70. \qquad (20)$$

4 RESULTS AND DISCUSSION

In this section, we will first compare the linear, non-linear and reconstructed $O_{\rm rw}$ measured for all models and redshifts. Then we present the results of the analytical fit to more quantitatively demonstrate the improvement by the reconstruction. Finally, we show the results of the constraints on the oscillation parameters and give forecast for the DESI-like survey.

4.1 Comparisons among wiggle spectra

In Fig. 2, we compare the results of the linear, non-linear and reconstructed $O_{\rm rw}(k)$ obtained from DM, H1 and H2 at the four redshifts for the four wiggled models. The black solid lines represent the linear $O_{\rm rw}(k)$ obtained from the initial conditions of the simulations, which are equivalent to the primordial oscillatory features. The blue dashed lines represent the non-linear $O_{\rm rw}(k)$ obtained from the output snapshots of the simulations, which are also referred to

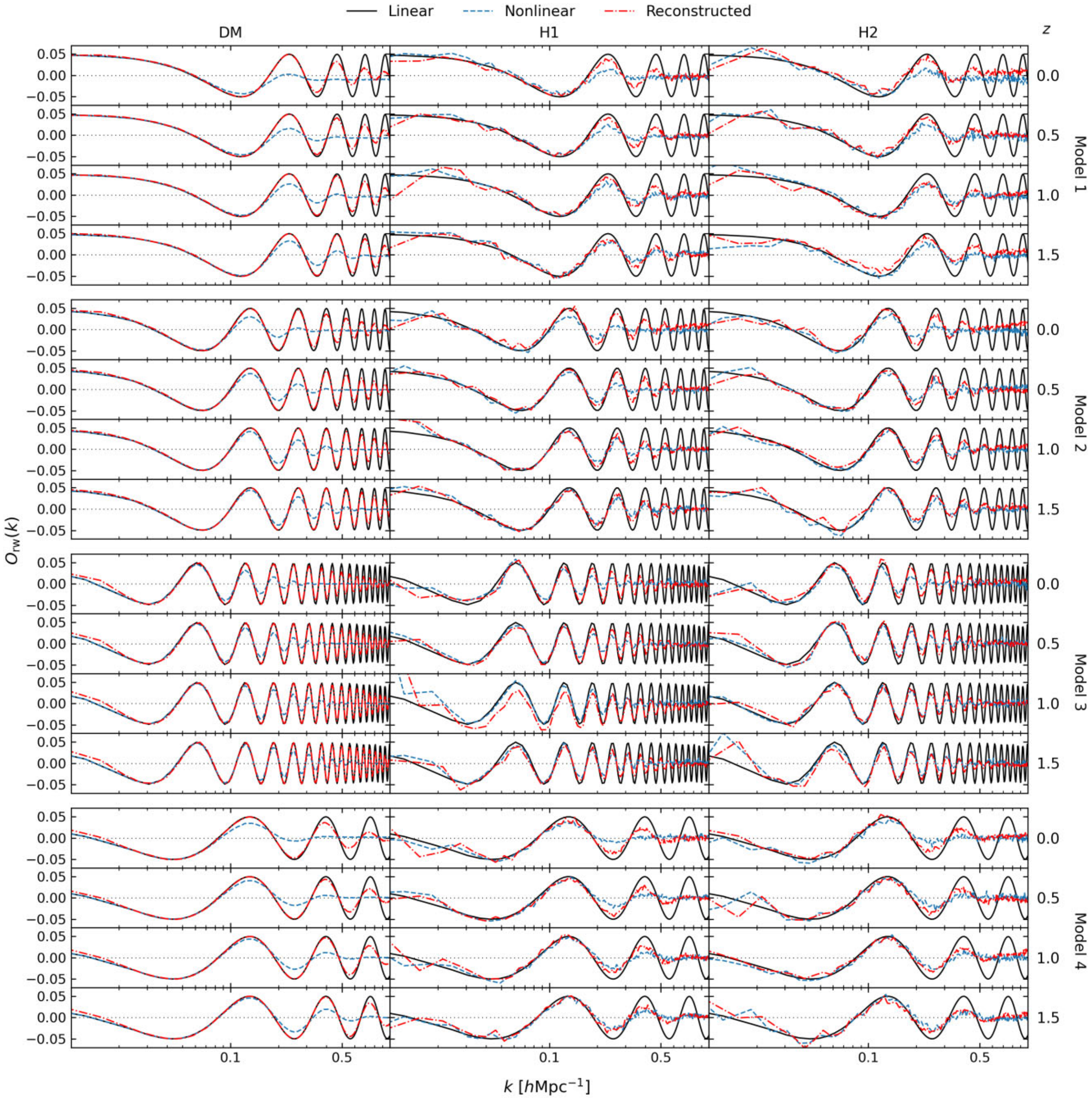

Figure 2. Comparisons among the linear (black solid line), non-linear (blue dash–dotted line) and reconstructed (red dashed line) $O_{\rm rw}$. The linear $O_{\rm rw}$ is measured from the initial conditions generated using 2LPTic, the non-linear $O_{\rm rw}$ is measured from the output snapshots of the simulations, and the reconstructed $O_{\rm rw}$ is obtained from the reconstructed density field. Each row represents one redshift z which is shown on the right side. The three columns denote, respectively, the results from the dark matter particle catalogue DM and the halo catalogues H1 and H2. Every four rows from the top down, respectively, belong to Models 1, 2, 3, and 4.

as the unreconstructed $O_{\rm rw}(k)$ for convenience. It can be seen that the wiggles on small scales are gradually damped as the redshift decreases. The red dash–dotted lines represent the reconstructed $O_{\rm rw}(k)$ obtained from the reconstructed density field, which helps to partially retrieve the damped wiggles.

The $O_{\rm rw}(k)$ results shown in the first column are obtained from DM, which exhibit some common characteristics for all three wiggled models. By comparing the unreconstructed results with the linear-theory predictions, it can be seen that the scale at which the wiggles start to be weakened becomes larger as time progresses. Furthermore, the wiggles on scales $k \gtrsim 0.3\,h\,{\rm Mpc}^{-1}$ are strongly damped at $z = 0$, and so the recovery of the wiggles on scales $0.3 \lesssim k \lesssim 0.5\,h\,{\rm Mpc}^{-1}$ would be an important objective of reconstruction. By comparing the reconstructed $O_{\rm rw}$ with the linear-theory prediction, we can see that, while the reconstructed power spectrum is not exactly the same as the linear spectrum, the reconstruction method to a certain extent helps retrieve the initial oscillations on our interested scales, $0.05 \lesssim k \lesssim 0.5\,h\,{\rm Mpc}^{-1}$. This agrees with the findings in Shi et al. (2018), which studied the performance of the same reconstruction method in dark matter reconstruction.

The success of the reconstruction from the dark matter particles is largely thanks to their high number density, which allows the late-time non-linear density field to be accurately produced: in this sense, reconstruction from DM can be considered as an idealised case or an upper limit, which will be difficult to achieve in real observations. For a rough comparison, we have shown, in the middle and right columns of Fig. 2, the $O_{\rm rw}$ results obtained from the two halo catalogues, H1 and H2, which have number densities similar to typical real galaxy catalogues. These results are less impressive than those for the dark matter particles because of the much smaller halo number densities. Also due to the small halo number densities, these results are noisier, which in theory can be made smoother by having more realizations of simulations, or equivalently a larger volume.

By comparing the results of H1 and H2 for the same model, we find that there is no significant difference in the unreconstructed $O_{\rm rw}(k)$ at the same redshift, because the number densities of these two halo catalogues only differ by a factor of 2. In most cases the reconstructed $O_{\rm rw}$ results of H1 seem slightly better compared to those of H2, as a result of the slightly larger halo number density in H1, though the difference is again insignificant visually. We shall revisit this point when discussing the analytical fit in the next subsection. Comparing the results with and without reconstruction, it is clear that the former does lead to less damped and sharper oscillation features, confirming that reconstruction can indeed help to partially retrieve the damped wiggles. This recovery seems more substantial at lower redshifts than at higher redshifts, since at higher redshifts there is less damping in the unreconstructed power spectra to start with. At lower redshifts, on the other hand, reconstruction can even recover some of the wiggles at $k \sim 0.5\,h\,{\rm Mpc}^{-1}$, where the wiggles are strongly damped in the unreconstructed case. We expect that this will help to improve the accuracy of the measurements of wiggle parameters, especially in models with few wiggles at $k \lesssim 0.3\,h\,{\rm Mpc}^{-1}$ – we will discuss this in the parameter fittings next.[2]

Finally, we notice that in rare cases, for example H1 at $z = 1$ and H2 at $z = 1.5$ for Model 3, the reconstructed $O_{\rm rw}$ seems to be poorer than the unreconstructed one. The exact cause of this is not clear, but we note that for these two cases the unreconstructed $O_{\rm rw}$ happens to be very noisy and deviate strongly from their theoretical values at large scales (a similar 'correlation' can be observed in certain other panels across Fig. 2, though to a lesser extent). It is possible that the halo power spectra in these cases have inaccurate amplitudes of the oscillations on large scales, which affect the reconstruction results. Given that in both H1 and H2 this only affects a particular snapshot and not all snapshots, we suspect that it is related to the only one realization per model we have used. Further investigation of this issue will be left for future works with more simulation realizations.

4.2 Wiggle parameter fitting

The corresponding best-fitting parameters of ω, ϕ, and $\zeta(z)$, as well as their uncertainties, are given in Table 2, which assist the understanding from a quantitative perspective. The relevant figures showing the analytical fit to the data can be found in the Appendix. As mentioned before, we will mainly focus on the results of H1 and H2, and so the results of DM would be taken as a reference and not be discussed in detail. The three parameters are mainly determined by the remaining peaks in the wiggles. We shall first discuss the results of the damping parameter, followed by the oscillation parameters, and then combine them to clarify the improvement given by reconstruction.

The damping parameter ζ effectively describes the extent of the damping effects caused by the gravitational non-linearities[3] and characterizes the suppression of the primordial oscillations. It is zero in the linear regime, such as at the initial redshift $z = 49$, and gradually increases at lower redshifts as the structures become progressively more non-linear and consequently more information of the wiggles in the primordial power spectrum gets damped. Thus reconstruction has the aim to reduce ζ and retrieve the primordial oscillations. Table 2 shows that the reconstructed values of ζ are evidently smaller than the unreconstructed values in all cases. Apart from a few high-redshift ($z > 1$) cases, the uncertainties of most cases are also reduced after reconstruction, which confirms that the reconstruction successfully retrieves the damped wiggles to an appreciable extent. Specifically, by comparing the cases among different models but the same catalogues and redshifts, the corresponding values after reconstruction seem to be nearly independent of the model, which implies that the improvement on the recovery of the wiggles does not strongly depend on the shape of the primordial oscillations.[4]

For a closer inspection, we show the ratios of unreconstructed to reconstructed ζ in Table 3, $\zeta_{\rm unrec}/\zeta_{\rm rec}$, which can be considered as an indicator of the reconstruction efficiency. We do this for all the cases (models, tracer types, and redshifts) listed in Table 2. The reconstruction efficiency of halo catalogues H1 and H2 increases with decreasing redshift, which shows that reconstruction is more beneficial for lower redshifts ($z < 1$). This is to be expected, given that the halo density field is more non-linear at low z and so the unreconstructed ζ is significantly larger than at high z; on the other hand, the reconstructed ζ depends more mildly on z, so that the ratio $\zeta_{\rm unrec}/\zeta_{\rm rec}$ increases with decreasing z. Also, among the low-redshift ($z < 1$) cases, the larger number density of H1 leads to higher efficiency when compared with H2 at the same redshift. For the DM case, the trend is reversed, with the ratio between unreconstructed and reconstructed ζ values increasing with redshift. Here, the behaviour is quite different from the halo cases, with the reconstructed ζ decreasing much faster with increasing redshift z. We have checked (though not shown here) that the values of $\zeta_{\rm unrec}/\zeta_{\rm rec}$ for the primordial features studied here are broadly consistent with the reconstruction efficiency defined in the same way applied to the reconstruction of BAO wiggles in Birkin et al. (2019), which uses the same reconstruction method and similar tracer number density.

Next, let us consider whether the 'sharpened' wiggles after reconstruction can lead to more accurate measurements of the oscillation parameters ω and ϕ. Regarding the oscillation frequency ω, the reconstructed values of ω are much closer to the theoretical values than the unreconstructed values in all cases, which is especially evident at low redshifts. Except for a few high-redshift cases, the improvement on the uncertainties after reconstruction is evident in most cases as well. The unreconstructed ω values of Models 2 and

---

[2] This is actually one of the motivations for our specific parameter choices in the feature models of equation (2), because we are particularly interested in cases where there are not many wiggles at $k \lesssim 0.3h\,{\rm Mpc}^{-1}$ to maximally show the power of reconstruction.

[3] Redistribution of matter due to baryonic processes, such as stellar and black hole feedback, could also lead to damping effects to the power spectrum, but that is less relevant for the range of scales we are interested in (some of the recent galaxy formation simulations, e.g. Schaye et al. 2015; Springel et al. 2018, predict that this affects the matter power spectrum at $k \gtrsim 1h\,{\rm Mpc}^{-1}$).

[4] This makes sense given that the amplitude of the primordial oscillations is relatively small in this work, so that the effects of the wiggles can be considered as small perturbations to the primordial and subsequently the evolved non-linear density field. Reconstruction, along with the reduction of ζ from the unreconstructed to the reconstructed cases that it leads to, is sensitive to the overall distribution of matter.

Table 3. The ratios of unreconstructed to reconstructed ζ, $\zeta_{\rm unrec}/\zeta_{\rm rec}$, obtained from Table 2, which can be used to describe the reconstruction efficiency, in all cases considered in Table 2.

	Model 1			Model 2			Model 3			Model 4		
z	DM	H1	H2	DM	H1	H2	DM	H1	H2	DM	H1	H2
0.0	3.52	1.82	1.69	3.61	1.94	1.82	3.73	2.01	2.11	3.30	1.94	1.78
0.5	3.64	1.66	1.52	3.89	1.83	1.61	3.87	1.79	1.73	3.48	1.63	1.61
1.0	4.03	1.42	1.43	4.26	1.52	1.44	4.00	1.60	1.54	3.74	1.41	1.41
1.5	4.29	1.32	1.34	4.43	1.37	1.33	3.93	1.44	1.40	4.11	1.23	1.30

3 appear to be closer to their theoretical values than in Models 1 and 4, which is probably because the former two models have more oscillation periods within the fitting range of scales than the latter two (see the right-hand panel of Fig. 1, or the blue lines in Fig. 2). After reconstruction, however, there is less clear difference among the four models, either in how close the reconstructed ω is to the theoretical value or in their uncertainties. Likewise, the difference between the best-fitting reconstructed ω values in H1 and H2 is rather mild, although the uncertainties are generally smaller for the former catalogue. Overall, the results indicate that reconstruction does indeed lead to a stronger improvement of the measurement of ω in Models 1 and 4, which have fewer visible peaks at $k \lesssim 0.5h\,{\rm Mpc}^{-1}$.

The situation is quite different in the case of the oscillation phase ϕ. The unreconstructed values of ϕ in Models 1, 2, and 3 are determined very well in most cases, so the reconstructed values only show a little improvement on the unreconstructed ϕ even for low-redshift cases. However, for Model 4 the unreconstructed values largely deviate from the theoretical value in all cases, and the unreconstructed values of H2 deviate even further than those of H1 at the same redshift. Although we cannot exclude the possibility that this discrepancy is an effect caused by the particular simulation, since we have only one realization for each model, we doubt this would be the cause, because the same random phases have been used to generate the ICs for all simulations. Instead, we suspect that this is more likely to be caused by the fact that $m \neq 1$ in Model 4, which means that the oscillation pattern is more complicated and thus leads to a less accurate fitting of ϕ. Regardless, based on the table, it seems that the reconstruction once again enables more accurate measurement of ϕ, especially for H2 at low redshift.

When considering the results of all three parameters, it seems that the reconstruction is most useful at low redshifts, $z < 1$, and Models 1 and 4 benefit more from it than Models 2 and 3 do. Although the peaks of Models 2 and 3 are better preserved after the cosmic evolution so that their reconstructed results are better than those of the other two models, the improvement is relatively limited, suggesting that the improvement depends not only on how clear-cut the reconstructed wiggles are, but also on how poorly the primordial wiggles are preserved before reconstruction. Overall, reconstruction seems more useful where the primordial wiggles are more damped.[5] As we mentioned before, the wiggles on scales $k \gtrsim 0.3\,h\,{\rm Mpc}^{-1}$ are strongly damped at $z = 0$; Models 2 and 3 have exactly the first several original peaks outside this range of scales, so these peaks are effectively preserved at low redshift. By contrast, we designed Model 4 so that it has one original peak at the same position of the first peak of Model 2 which is effectively preserved, and its second peak is at the same position of the third peak of Model 2, which is strongly damped. Therefore, the primordial wiggles of Model 4 are preserved less well than those of Model 2, and this Model benefits more from the reconstruction. Similarly, Model 1 has two original peaks in the range $k \lesssim 0.5\,h\,{\rm Mpc}^{-1}$: the first is at a smaller scale compared with the first peak of the other models and thus is not preserved as well as the first peak of the other models due to the stronger damping effect, while the second peak is completely damped. Therefore Models 1 and 4 both benefit from the reconstruction substantially more than Models 2 and 3.

Additionally, the values of ω used in Models 1, 2, and 3 imply that the reconstruction method is not only effective at low frequency, such as ~ 40 Mpc, but also working well at relatively higher frequency, such as ~ 150 Mpc.

4.3 Constraints on oscillation parameters for DESI-like survey

Since the four wiggled models have similar results of the constraints on the oscillation parameters, we shall take Model 1 as an example to illustrate and discuss how the reconstruction potentially improves the constraints in a real galaxy survey. Additionally, we also forecast how much the uncertainties of the feature amplitude can be reduced after reconstruction for the wiggled models.

Fig. 3 shows the forecasted constraints on the oscillation parameters for a DESI-like survey with a survey area of 14 000 deg^2, based on the primordial oscillations of Model 1. The marginalized posterior distribution of each parameter shown in the upper panels indicates that, without reconstruction, the case of $k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$ (the red lines) give better constraints than the case with $k_{\rm max} = 0.25\,h\,{\rm Mpc}^{-1}$ (grey), because in the former case more k modes are included in the Fisher matrix and increase the accuracy of the constraints. Additionally, by comparing the cases with the same $k_{\rm max}$ (red versus blue, or grey versus green lines), we find that reconstruction leads to stronger constraints on the parameters, especially with $k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$. This is because the oscillation wiggles on scales $k \gtrsim 0.25\,h\,{\rm Mpc}^{-1}$ are heavily damped at low redshift without any reconstruction, while the reconstructed wiggles at $k = (0.25 - 0.5)\,h\,{\rm Mpc}^{-1}$ significantly contribute to the constraints. By contrast, since the peaks on scales $k \lesssim 0.25\,h\,{\rm Mpc}^{-1}$ are preserved reasonably well, the reconstruction for $k_{\rm max} = 0.25\,h\,{\rm Mpc}^{-1}$ does not lead to as much benefit as in the case of $k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$. Furthermore, stronger constraints are shown for ELGs (right-hand panels) compared with LRGs (left-hand panels), because the former has more available redshift bins and larger number density for the same redshift bins.

In particular, every two out of three parameters show degeneracies in the confidence contours when $k_{\rm max} = 0.25\,h\,{\rm Mpc}^{-1}$, though these degeneracies are broken and replaced with stronger constraints when $k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$ in the A–ω and A–ϕ contours due to more k modes included. By contrast, the ω–ϕ contours keep the degeneracy which is a consequence caused by the oscillation model itself and by the fact that here we are trying to constrain both oscillatory frequency and phase over a limited range of k.

---

[5] This statement, of course, is based on the limited range of models we have studied here.

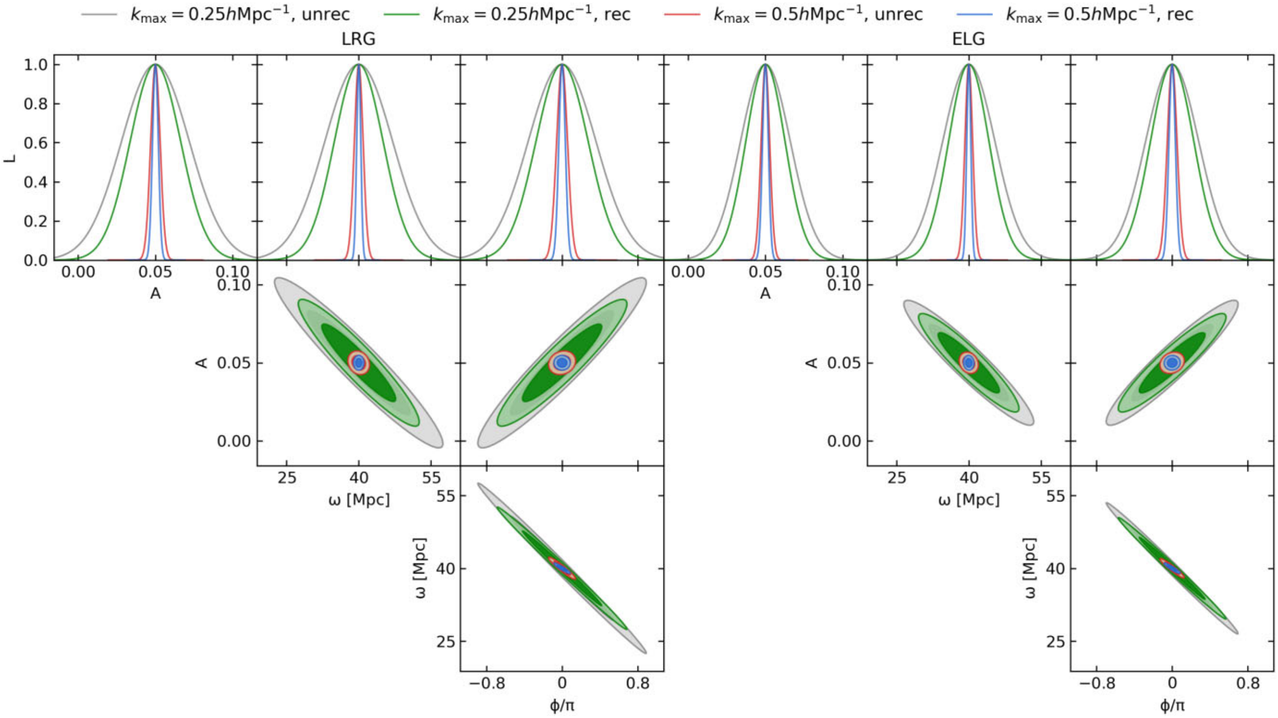

Figure 3. Forecasts of constraints on the oscillatory feature parameters for a DESI-like survey with a survey area of 14 000 deg^2, for the primordial oscillations of Model 1. The left side is for LRGs and the right side is for ELGs. The upper panels show the 1D marginalized posterior distributions. The middle and lower panels show the marginalized 68 and 95 per cent confidence contours for every two out of three feature parameters. The green and grey colours represent, respectively, the cases for $k_{\rm max} = 0.25h\,{\rm Mpc}^{-1}$ with and without reconstruction, while the blue and red colours represent the cases for $k_{\rm max} = 0.5h\,{\rm Mpc}^{-1}$ with and without reconstruction.

Lastly, similar to previous works (Beutler et al. 2019; Slosar et al. 2019; Ballardini et al. 2020), we show the marginalized uncertainties of feature amplitude as a function of oscillatory frequency for our feature models in Fig. 4 and discuss the implications of the results. Because Models 1, 2, and 3 have an identical form of oscillations and almost same damping parameters within the error bars, we only show the results of Models 1 and 4 here.

We consider Model 1 first. As expected, ELGs place slightly tighter constraints than LRGs due to their larger number densities and redshift range. The sharp peaks that appear at $\omega \simeq 150\,{\rm Mpc}$ are due to the degeneracy between the oscillatory features and the BAO wiggles. We have tested that for $\omega \gtrsim 200\,{\rm Mpc}$ the uncertainties almost stay as a constant, and so we have cut off the figure at $\omega^m = 300\,{\rm Mpc}^m$. For smaller ω, things are complicated and behave differently for different $k_{\rm max}$. For $k_{\rm max} = 0.25\,h\,{\rm Mpc}^{-1}$ we can see an increase in the uncertainties at $\omega \lesssim 70\,{\rm Mpc}$, while a similar increase starts to appear at even smaller ω–30 Mpc–for $k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$. Thus larger $k_{\rm max}$ has an extra advantage of significantly reducing the uncertainties for small ω, in addition to giving more stringent constraints (everything else the same) for all ω overall. By comparing the pairs of curves with the same colours, i.e. the same cases ($k_{\rm max}$ and reconstructed versus unreconstructed) but different survey areas, we find that, as expected, a larger survey area always gives better constraints.

Most interestingly, everything else equal, performing the non-linear reconstruction can significantly reduce the uncertainties of A. As an example, for large values of ω, in the case of $k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$ and a survey area equal to 14 000 deg^2, reconstruction reduces $\sigma(A)$ from ∼0.003 to ∼0.002, and this improvement is stronger than not performing reconstruction, but instead going from 9000 to 14 000 deg^2 with $k_{\rm max}$ fixed to 0.25 or 0.5 $h\,{\rm Mpc}^{-1}$, or increasing $k_{\rm max}$ from 0.25 to 0.5 $h\,{\rm Mpc}^{-1}$ keeping the survey area fixed to either 9000 or 14 000 deg^2. A similarly good improvement can be seen with $k_{\rm max} = 0.25\ h\,{\rm Mpc}^{-1}$ or survey area equal to 9000 deg^2, when doing reconstruction. In certain cases, e.g. the large-ω regime of the lower panels of Fig. 4, reconstruction with $k_{\rm max} = 0.25\ h\,{\rm Mpc}^{-1}$ and a survey area equal to 9000 deg^2 (the thin green dashed line) can lead to comparable constraints to not doing reconstruction but with $k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$ and a survey area equal to 14 000 deg^2 (the thick orange dot–dashed line). Given that increasing survey area is not always possible due to the finite sky area, but increasing $k_{\rm max}$ in analyses for these primordial feature models is comparably more straightforward (Beutler et al. 2019), combining an increase in $k_{\rm max}$ with non-linear reconstruction can be a potentially promising way to obtain even stronger constraints on the feature parameters, and help to maximize the scientific return of future survey data.

The behaviour of Model 4 is similar to that of Model 1, e.g. both the absolute and the relative heights of the different curves, as well as their shapes are the same as before. There are, however, some notable differences, e.g. the main peaks in $\sigma(A)$ in Model 4 are at slightly different values of ω from the other models, and the curves are also less smooth. As mentioned above, the bump (which has the structure of a double peak) of $\sigma(A)$ for Model 1 is related to the BAO peak in the matter/galaxy correlation function, which is at $\simeq 150\,{\rm Mpc}$. The primordial wiggles of Model 1, in configuration space, correspond to a spike at matter or halo separation $r = \omega$. When $\omega \gg 150\,{\rm Mpc}$, the BAO and primordial peaks are separated afar and thus the former does not affect the accuracy of the measurement for the latter. As ω approaches 150 Mpc from above, the BAO and primordial peaks start to 'interfere', leading to changes of both the amplitude and shape of the latter, making it harder to measure its parameters accurately. We

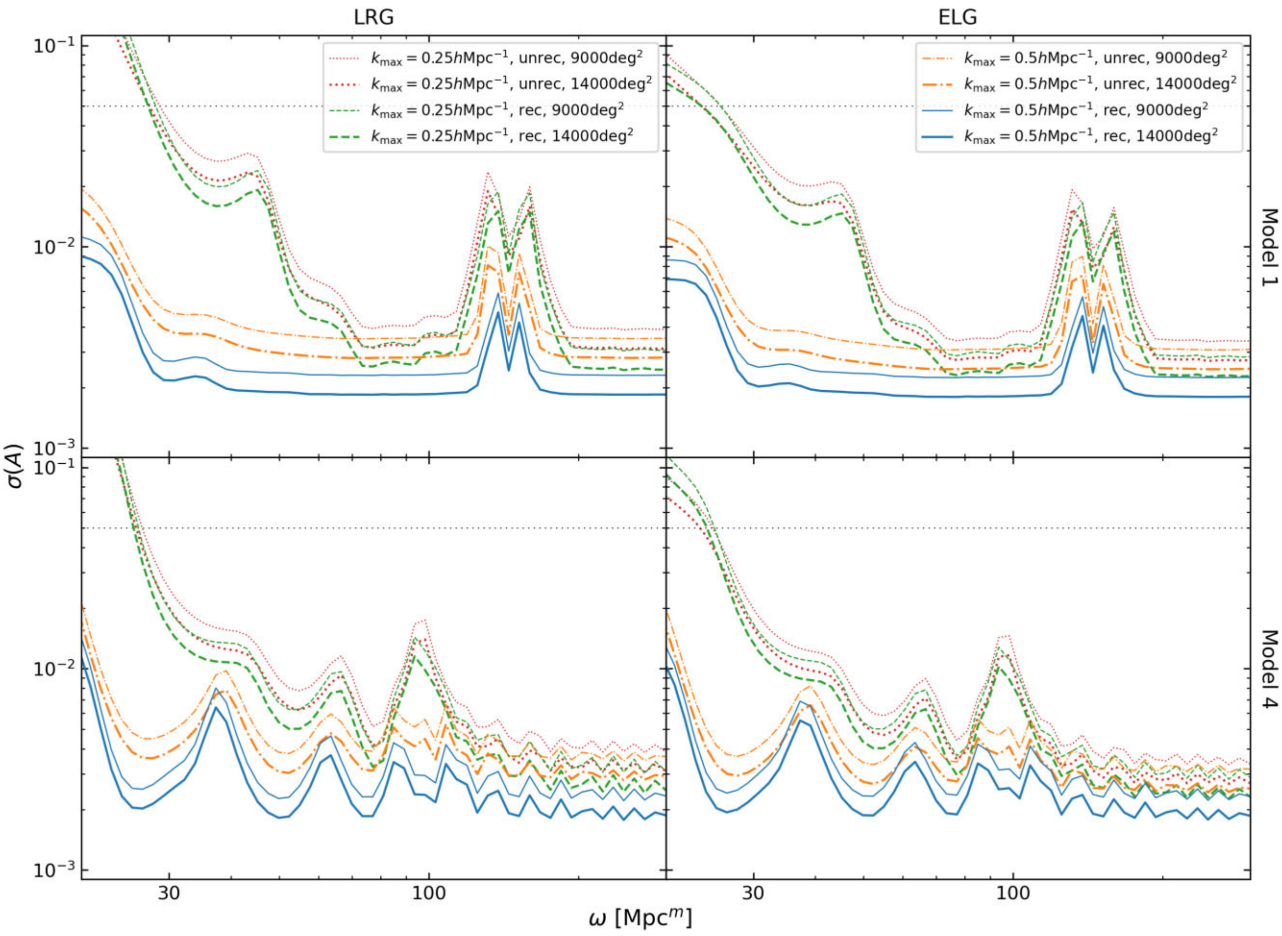

Figure 4. Forecasts of the marginalized uncertainties of the oscillation amplitude A as a function of the frequency ω, for the two models, Model 1 (top row) and Model 4 (bottom row); the result for Models 2 and 3 are not shown here since the two models have the identical form of oscillations to that of Model 1. The first column is for LRGs and the second column is for ELGs. The dotted black lines mark the theoretical amplitudes of the oscillations, $A = 0.05$, used in the forecasts. The meanings of the different colours and line styles are indicated in the legends. The same colours represent the cases with same $k_{\rm max}$ and same situation of reconstruction but different survey areas; the thick lines are for the survey area of 14 000 deg^2 and the thin lines are for 9000 deg^2.

speculate that the dip – which causes the double-peak structure in $\sigma(A)$ for Model 1 – is due to the fact that, when the primordial peak does not coincide well with the centre of the (rather wide) BAO peak, its shape can be affected in an asymmetric manner, making the measurement of its parameters even more inaccurate. In contrast, the structure of the primordial wiggles in Model 4 is more complicated in configuration space, because $m \neq 1$ in equation (2), which can cause the differences in the units of ω and other fine details of $\sigma(A)$ between this and the other models.

5 CONCLUSIONS

In this paper, we have investigated the effect of a non-linear density reconstruction method on retrieving hypothetical oscillatory features in the primordial power spectrum which are significantly damped on small scales in the late-time Universe due to cosmological structure formation.

We considered four different oscillatory features which are added to a simple power-law type primordial power spectrum, for which we ran *N*-body simulations and identified dark matter halo catalogues at a number of redshifts. We reconstructed the initial density fields from the particle data and halo catalogues with two different number densities. Finally, we compared the fitted feature parameters from the power spectra of the unreconstructed and reconstructed density fields, to identify the improvement by reconstruction. We showed that non-linear reconstruction can effectively help to retrieve the damped wiggles with a range of frequencies between 40 and 150 Mpc – not only does it lead to less biased best-fitting values of the feature parameters, but it also substantially shrinks the measurement uncertainty. The improvement was especially strong where the primordial features have been less well preserved pre-reconstruction to start with, such as at $z < 1$.

In order to forecast the constraints on the feature parameters from a DESI-like galaxy survey, we modelled the observed broad-band galaxy power spectrum based on the HALOFIT prediction of the non-linear matter power spectrum with the addition of oscillatory features studied in this work, and then used the analytical marginalized Fisher matrix to calculate the expected constraints on the oscillation parameters using the specifications of DESI LRGs and ELGs. We found that non-linear reconstruction led to more robust constraints on the oscillation parameters, with the equivalent effects of enlarging the survey area (but at a much smaller cost) and/or increasing the k range.

While non-linear reconstruction has been proposed to be used in improving the measurement of the BAO scale (e.g. Wang et al. 2017), and hence the determination of the expansion rate of the Universe and hence the properties of dark energy, this work has demonstrated that similar applications are possible in other cases where certain features in matter clustering are present, following the spirit of earlier works such as Beutler et al. (2019). This is particularly true if these features are in the mildly non-linear regime, $0.1 \lesssim k/(h\,\mathrm{Mpc}^{-1}) \lesssim 0.5$, since this range of scales is what the non-linear reconstruction method used here helps most: on even larger scales the benefit of reconstruction is insignificant, while on further smaller scales reconstruction will not help much.

The methodology exemplified in this paper assumes that we know the functional form of the primordial features *a priori* – this is how we forecasted constraints on the oscillation amplitude A. However, the reconstruction step is completely independent of any assumption of a particular primordial feature, and hence any method developed for detecting general features from the matter clustering should apply to and benefit from the reconstructed density field.

As a first step, the present study is based on various simplifications, and we discuss a couple here which can be improved in the future. The first is related to the post-reconstruction damping parameter ζ. As we have discussed, ζ characterizes the damping of the primordial features, and a smaller ζ means that the reconstruction has done a better job. Due to the limited number of simulations carried out in this work (one realization per model), shot noise will impact the estimated reconstruction efficiency. This could be improved by increasing the number of simulations and more studies are needed in the future.

The second is related to the modelling of redshift-space distortions (RSD), for which we have adopted a simplistic prescription and well pushed beyond the limit (e.g. $k \simeq 0.5\,h\,\mathrm{Mpc}^{-1}$) where it is expected to work. This is not an issue for a forecast work, but for constraints using real data it should be treated more carefully. The reconstruction method here has been extended to remove RSD from observed galaxy catalogues (Wang et al. 2020), though that is unlikely to work reliably at k as large as $\simeq 0.5\,h\,\mathrm{Mpc}^{-1}$. Of course, we can always cut $k_{\max}$ to something that we are comfortable with. However, as mentioned above, if we would like to take maximum benefit from reconstruction, it is likely that we need to go substantially beyond $k \simeq 0.1\,h\,\mathrm{Mpc}^{-1}$. This can be achieved, for example, by using emulators of redshift-space galaxy or halo clustering (see e.g. Zhai et al. 2019; Kobayashi et al. 2020); actually, as long as the primordial oscillations are weak (as implied by current null detections), one might assume that their presence has little or negligible impact on RSD.

The ultimate objective, of course, is to apply this method to real observation data from future galaxy surveys such as *Euclid* and DESI. For this, the above-mentioned improvements, amongst many others, would need to be done properly. These will be left for future works, in which we plan to carry out updated forecasts for these surveys and eventually real constraints.

ACKNOWLEDGEMENTS

We thank collaborators within Euclid and DESI for various discussions while this project was going on. YL thanks Robert Smith for his support during this project. BL is supported by the European Research Council through ERC Starting Grant ERC-StG-716532-PUNCA, and the Science Technology Facilities Council (STFC) through ST/T000244/1 and ST/P000541/1. HMZ is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) [funding reference number CITA 490888-16]. This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1, and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

DATA AVAILABILITY

Simulation data used in this work can be made available upon request to the authors.

REFERENCES

Achucarro A., Gong J.-O., Hardeman S., Palma G. A., Patil S. P., 2011, J. Cosmol. Astropart. Phys., 01, 030
Adams J. A., Cresswell B., Easther R., 2001, Phys. Rev. D, 64, 123514
Ade P. A. R. et al., 2014a, A&A, 571, A22
Ade P. A. R. et al., 2014b, A&A, 571, A24
Ade P. A. R. et al., 2016a, A&A, 594, A17
Ade P. A. R. et al., 2016b, A&A, 594, A20
Adshead P., Dvorkin C., Hu W., Lim E. A., 2012, Phys. Rev. D, 85, 023531
Aghanim N. et al., 2020, A&A, 641, A6
Akrami Y. et al., 2020a, A&A, 641, A9
Akrami Y. et al., 2020b, A&A, 641, A10
Albrecht A., Steinhardt P. J., 1982, Phys. Rev. Lett., 48, 1220
Alonso D., 2012, preprint (arXiv:1210.1833)
Ballardini M., Finelli F., Fedeli C., Moscardini L., 2016, J. Cosmol. Astropart. Phys., 10, 041
Ballardini M., Finelli F., Maartens R., Moscardini L., 2018, J. Cosmol. Astropart. Phys., 04, 044
Ballardini M., Murgia R., Baldi M., Finelli F., Viel M., 2020, J. Cosmol. Astropart. Phys., 04, 030
Bartolo N., Komatsu E., Matarrese S., Riotto A., 2004, Phys. Rep., 402, 103
Bean R., Chen X., Hailu G., Tye S. H. H., Xu J., 2008, J. Cosmol. Astropart. Phys., 03, 026
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, ApJ, 762, 109
Beutler F., Biagetti M., Green D., Slosar A., Wallisch B., 2019, Phys. Rev. Res., 1, 033209
Birkin J., Li B., Cautun M., Shi Y., 2019, MNRAS, 483, 5267
Cautun M. C., van de Weygaert R., 2011, Astrophysics Source Code Library, record ascl:1105.003
Chen X., 2010, Adv. Astron., 2010, 638979
Chen X., 2012, J. Cosmol. Astropart. Phys., 01, 038
Chen X., Namjoo M. H., 2014, Phys. Lett. B, 739, 285
Chen X., Easther R., Lim E. A., 2007, J.Cosmol. Astropart. Phys., 06, 023
Chen X., Namjoo M. H., Wang Y., 2015, J. Cosmol. Astropart. Phys., 02, 027
Chen X., Dvorkin C., Huang Z., Namjoo M. H., Verde L., 2016, J. Cosmol. Astropart. Phys., 11, 014
Chen S.-F., Vlah Z., White M., 2020, J. Cosmol. Astropart. Phys., 11, 035
Chluba J., Hamann J., Patil S. P., 2015, Int. J. Mod. Phys. D, 24, 1530023
Colombi S., Novikov D., 2011, Astrophysics Source Code Library, record ascl:1110.017
Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369
Debono I., Hazra D. K., Shafieloo A., Smoot G. F., Starobinsky A. A., 2020, MNRAS, 496, 3448
DESI Collaboration, 2016, preprint (arXiv:1611.00036)
Doré O. et al., 2014, preprint (arXiv:1412.4872)
Eisenstein D. J., Seo H.-j., Sirko E., Spergel D., 2007, ApJ, 664, 675
Feldman H. A., Kaiser N., Peacock J. A., 1994, ApJ, 426, 23
Flauger R., Pajer E., 2011, J. Cosmol. Astropart. Phys., 01, 017
Flauger R., McAllister L., Pajer E., Westphal A., Xu G., 2010, J. Cosmol. Astropart. Phys., 06, 009
Gao X., Langlois D., Mizuno S., 2012, J. Cosmol. Astropart. Phys., 10, 040
Guth A. H., 1981, Phys. Rev. D, 23, 347
Hazra D. K., Aich M., Jain R. K., Sriramkumar L., Souradeep T., 2010, J. Cosmol. Astropart. Phys., 10, 008
Hazra D. K., Shafieloo A., Smoot G. F., Starobinsky A. A., 2014, J. Cosmol. Astropart. Phys., 08, 048
Hinshaw G. et al., 2013, ApJS, 208, 19

Huang Z., Verde L., Vernizzi F., 2012, J. Cosmol. Astropart. Phys., 04, 005

Ivezić v. et al., 2019, ApJ, 873, 111

Kazin E. A. et al., 2014, MNRAS, 441, 3524

Kobayashi Y., Nishimichi T., Takada M., Takahashi R., Osato K., 2020, Phys. Rev. D, 102, 063504

Komatsu E. et al., 2009, ApJ Suppl., 180, 330

L'Huillier B., Shafieloo A., Hazra D. K., Smoot G. F., Starobinsky A. A., 2018, MNRAS, 477, 2503

Lewis A., Challinor A., 2011, CAMB: Code for Anisotropies in the Microwave Background

Li B., Zhao G.-B., Teyssier R., Koyama K., 2012, J. Cosmol. Astropart. Phys., 01, 051

Li B., Barreira A., Baugh C. M., Hellwing W. A., Koyama K., Pascoli S., Zhao G.-B., 2013, J. Cosmol. Astropart. Phys., 11, 012

Linde A. D., 1982, Phys. Lett. B, 108, 389

Liu Y., Yu Y., Li B., 2021, ApJS, 254, 4

Mao T.-X., Wang J., Li B., Cai Y.-C., Falck B., Neyrinck M., Szalay A., 2021, MNRAS, 501, 1499

Palma G. A., Sapone D., Sypsas S., 2018, J. Cosmol. Astropart. Phys., 06, 004

Peiris H. V. et al., 2003, ApJS, 148, 213

Racca G. D. et al., 2016, in MacEwen H. A., Fazio G. G., Lystrup M., eds, Proc. SPIE Conf. Ser. Vol. 9904, Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 0O

Sarpa E., Schimd C., Branchini E., Matarrese S., 2019, MNRAS, 484, 3818

Schaye J. et al., 2015, MNRAS, 446, 521

Schmittfull M., Feng Y., Beutler F., Sherwin B., Chu M. Y., 2015, Phys. Rev. D, 92, 123522

Seo H.-J., Eisenstein D. J., 2003, ApJ, 598, 720

Shi Y., Cautun M., Li B., 2018, Phys. Rev. D, 97, 023505

Slosar A., Chen X., Dvorkin C., Meerburg D., Wallisch B., Green D., Silverstein E., 2019, BAAS, 51, 98

Spergel D. N. et al., 2007, ApJS, 170, 377

Springel V. et al., 2018, MNRAS, 475, 676

Starobinsky A. A., 1992, Sov. J. Exp. Theor. Phys. Lett., 55, 489

Taylor A. N., Kitching T. D., 2010, MNRAS, 408, 865

Tegmark M., 1997, Phys. Rev. Lett., 79, 3806

Teyssier R., 2002, A&A, 385, 337

Tinker J. L., Kravtsov A. V., Klypin A., Abazajian K., Warren M. S., Yepes G., Gottlober S., Holz D. E., 2008, ApJ, 688, 709

Vasudevan A., Ivanov M. M., Sibiryakov S., Lesgourgues J., 2019, J. Cosmol. Astropart. Phys., 09, 037

Vlah Z., Seljak U., Chu M. Y., Feng Y., 2016, J. Cosmol. Astropart. Phys., 03, 057

Wang X., Yu H.-R., Zhu H.-M., Yu Y., Pan Q., Pen U.-L., 2017, ApJ, 841, L29

Wang Y., Li B., Cautun M., 2020, MNRAS, 497, 3451

Zeng C., Kovetz E. D., Chen X., Gong Y., Muñoz J. B., Kamionkowski M., 2019, Phys. Rev. D, 99, 043517

Zhai Z. et al., 2019, ApJ, 874, 95

Zhu H.-M., Yu Y., Pen U.-L., Chen X., Yu H.-R., 2017, Phys. Rev. D, 96, 123502

APPENDIX: RESULTS OF WIGGLE FITTING

Figs A1, A2, A3, and A4 show, respectively, the results of the analytic fit to the unreconstructed and reconstructed $O_{\rm rw}$ results for the four models studied in this work. It can be seen that, in most cases, the analytic model equation (5), with a Gaussian damping function characterized by the parameter $\zeta(z)$, fits the pre- and post-reconstruction data very well.

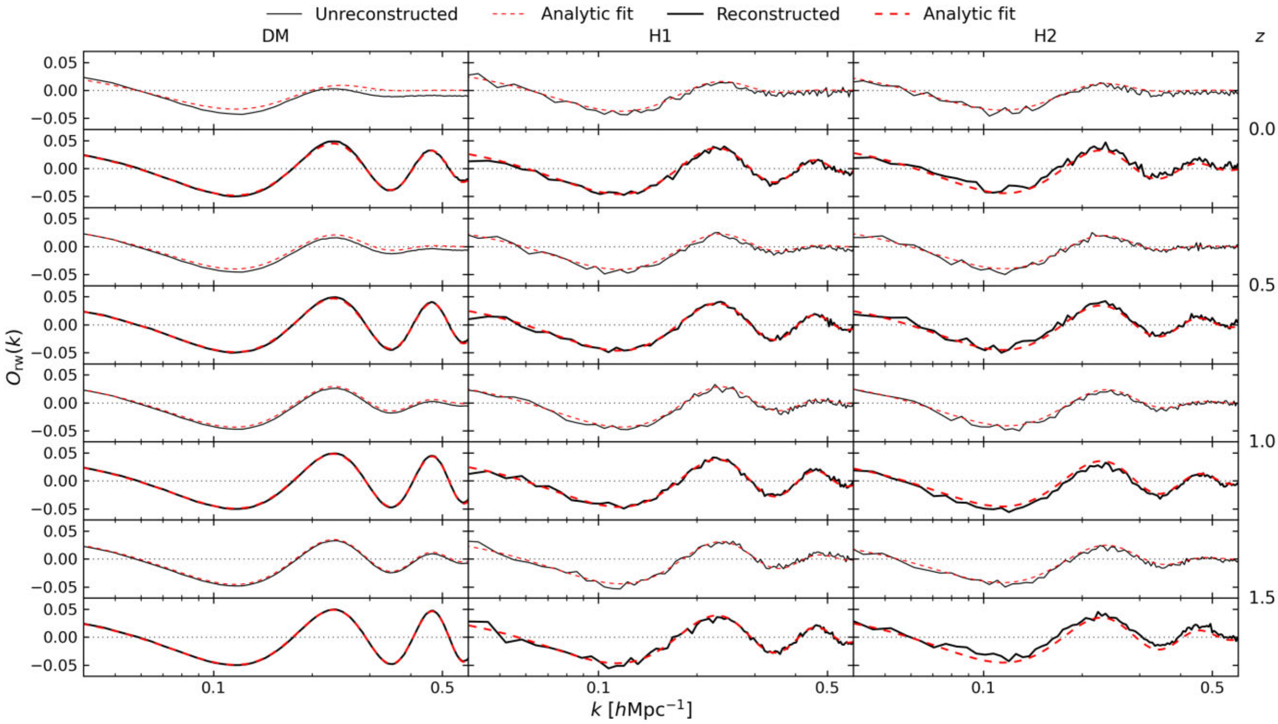

Figure A1. The analytical fit to the unreconstructed and reconstructed $O_{\rm rw}$ for Model 1. The black solid lines represent the measured $O_{\rm rw}$ and the red dashed lines represent the fitting curves given by the analytical model, equation (5). The thin lines are for the unreconstructed cases and the thick lines are for the reconstructed cases. The three columns from left to right, respectively, denote the dark matter particle catalogue DM, and the halo catalogues H1 and H2. Every two rows from the top down represent the same redshift shown on the right side. In each group of two rows, the upper one is for the unreconstructed, and the lower one for the reconstructed, case.

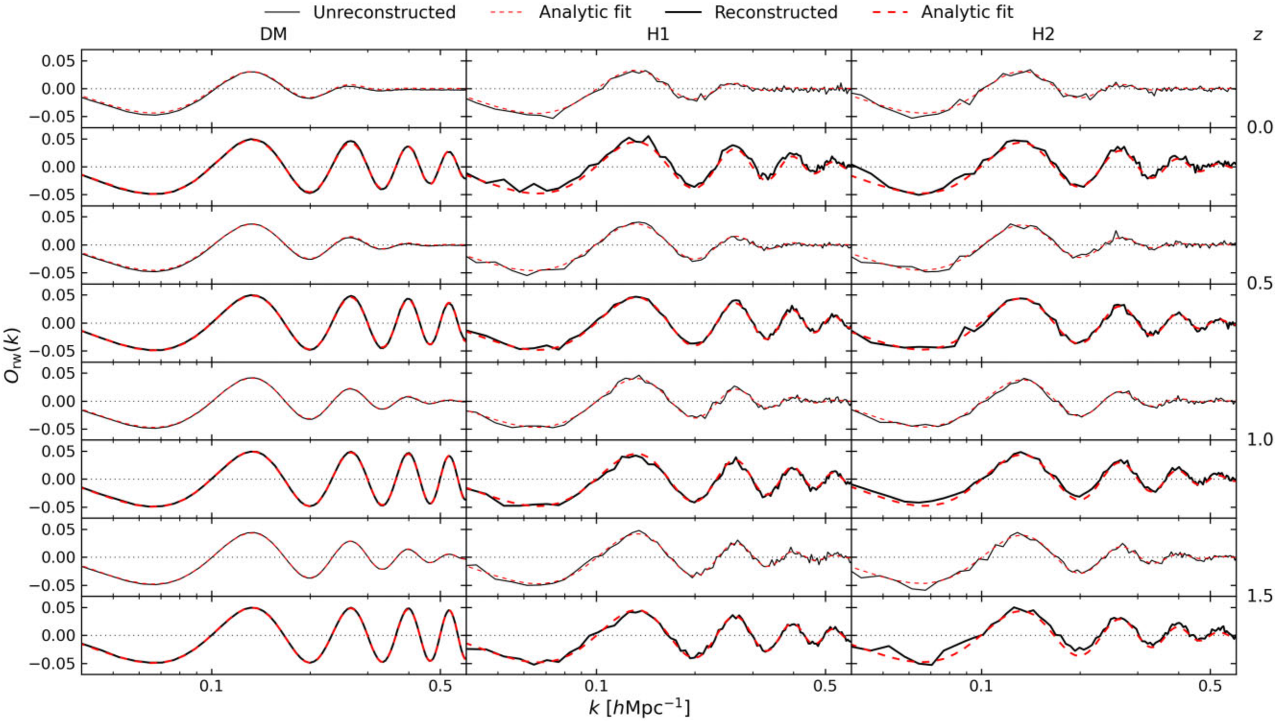

Figure A2. The same as Fig. A1 but for Model 2.

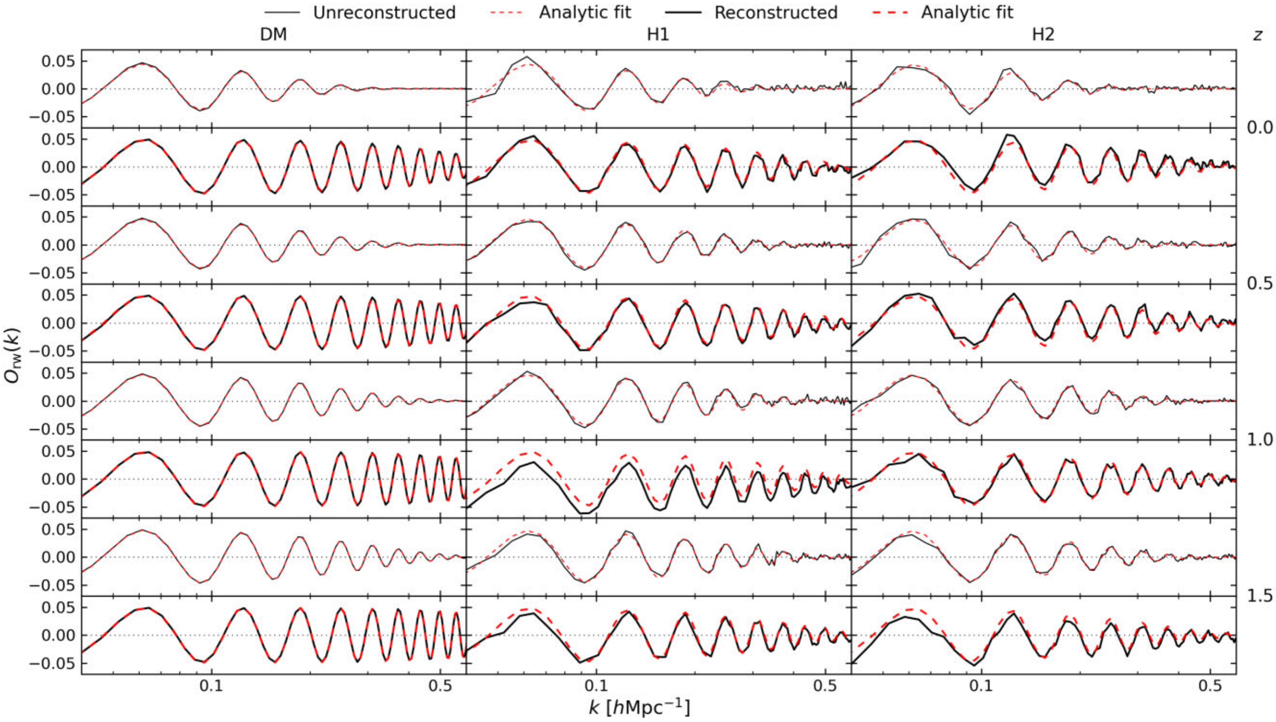

Figure A3. The same as Fig. A1 but for Model 3.

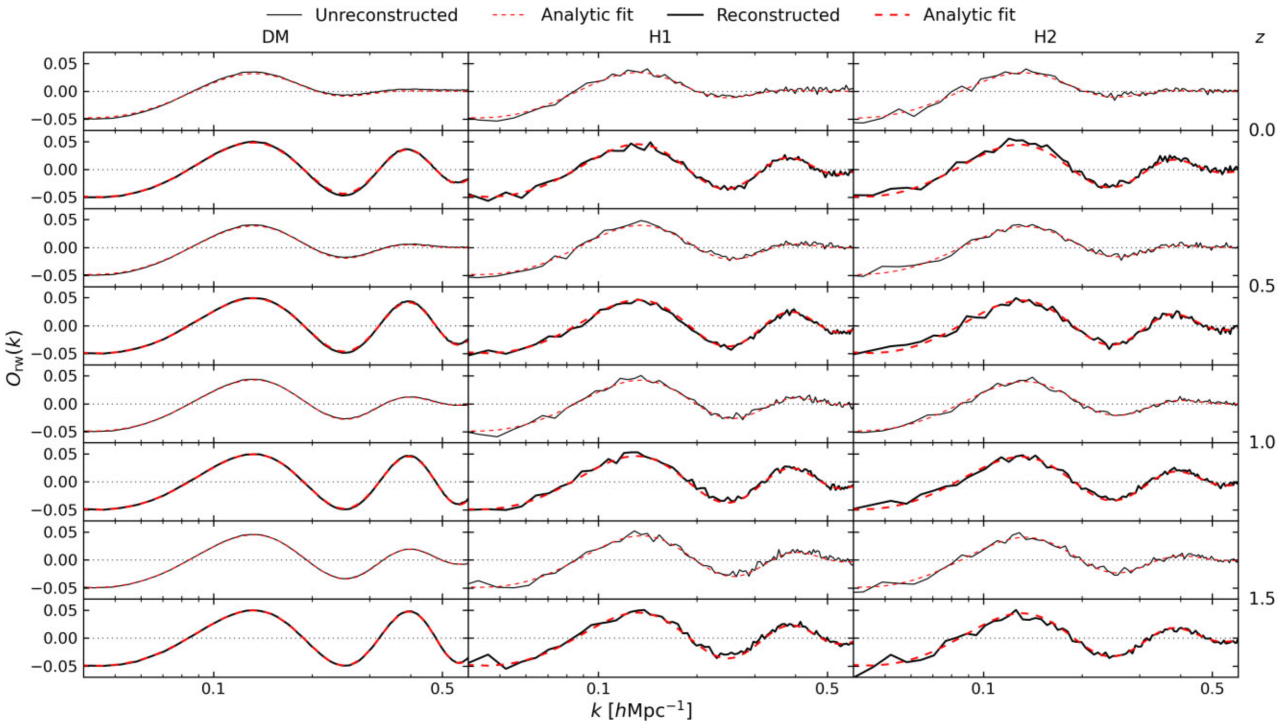

Figure A4. The same as Fig. A1 but for Model 4.

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