Reconstructing features in the primordial power spectrum

Yuhao Li,1,* Hong-Ming Zhu2 and Baojiu Li3
1Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Falmer, Brighton, BN1 9RH, UK
2Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, Ontario M5S 3H8, Canada
3Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham, DH1 3LE, UK

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
Potential features in the primordial power spectrum, such as oscillatory patterns, have been searched for in galaxy surveys in recent years, since these features can assist in understanding the nature of inflation and distinguishing between different scenarios 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 weakened or even wiped out on small scales in the late Universe due to nonlinear 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 suite of large N-body simulations, we show that a recently-proposed nonlinear reconstruction algorithm can effectively retrieve lost oscillatory features from the mock galaxy catalogues and improve the accuracy of the measurement of feature parameters (assuming such primordial features do exist). We do a Fisher analysis to forecast how reconstruction affects the constraining power, and find that it can lead to significantly more robust constraints on the oscillation amplitude for a DESI-like survey. In particular, we compare the application of reconstruction with other ways of improving constraints, such as increasing the survey volume and range of scales, and show that it can achieve what the latter do, but at a much 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 simplest version of single-field inflation (Guth 1981; Linde 1982; Albrecht & Steinhardt 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 inflaton field that is believed to have driven inflation is not fully understood yet, and the fact that the very high energy in the early Universe makes it an ideal place to witness the consequences of the laws of fundamental physics offers hope that new physics could be revealed by cosmological observations of the large-scale structures. More sophisticated inflation models leading to primordial non-Gaussianity have been developed across the last decades (see Bartolo et al. 2004; Chen 2010; Celoria & Matarrese 2020, for some reviews). These models predict unique features deviating from those of the simple single-field inflation model, which can be classified into different types, each of which can be attributed to different physical mechanisms (see Chen 2010; Chluba et al. 2015, for some reviews). Chen et al. (2016) briefly reviews several representative feature models of inflation, such as the sharp feature model which oscillates in linear scale in the primordial power spectrum, can be described by the template of the sinusoidal wiggle in the power-law primordial spectrum, which is also a good approximation for the scalar power spectrum in the axion monodromy model (Flauger et al. 2017). This oscillatory feature could be generated by a localised sharp feature in inflationary potentials or internal field space, and the nature of the sharp feature could correspond to distinct mechanisms, such as a step or bump in the single-field inflationary potential (e.g., Adams et al. 2001; Adshead et al. 2012; Bartolo et al. 2013; Hazra et al. 2014), or the embedding of new physics (Bozza et al. 2003). Other typical feature models include the resonance model in which the feature oscillates in logarithmic scale (see e.g. Wang et al. 2005; Bean et al. 2008; Flauger et al. 2010, for the mechanisms behind the resonance model), and the standard clock model which combines the form of the previous two feature models (e.g., Chen & Namjoo 2014; Chen et al. 2015).

As a result, the (non)detecion of primordial features can be used to distinguish among different scenarios of inflation. A variety of classes of well-motivated inflationary models, such as the so-called general single-field inflation (e.g., Chen et al. 2007; Senatore et al. 2010), the multi-field inflation (see Byrnes & Choi 2010, for a review) and so on, are continuously tested with the updated release of data from the Planck mission (Ade et al. 2014b, 2016a; Akrami et al. 2020a),
though none of them is preferable to the simplest single-field inflation model so far, which suggests that such features should be fairly weak if they do exist. Since the primordial features are not only imprinted in the CMB, and some can also leave a signature in the matter power spectrum, some future large-scale structure (LSS) surveys with high sensitivity, such as Euclid (Racca et al. 2016), DESI (DESI Collaboration et al. 2016) and SPHEREx (Doré et al. 2014) surveys, will provide the opportunity to search for the primordial features as a complement to CMB data (e.g., Chen et al. 2016; Ballardini et al. 2016; Palma et al. 2018; L’Huillier et al. 2018; Zeng et al. 2019), or strengthen the constrains on the strength of such features.

However, any feature imprinted in the primordial density or curvature field by inflation is subject to the impact of cosmic evolution that leads to today. In particular, even if the primordial features once existed in the very early Universe, they would have been significantly weakened or even wiped out on small scales in the late-time Universe due to nonlinear structure formation (Beutler et al. 2019; Ballardini et al. 2020). Meanwhile, the available information on large scales, where the evolution can be described by linear theory, is limited due to the cosmic variance. This can affect the confidence level at which to measure or constrain these features. A potential method to address this issue is to undo the cosmological evolution in a process usually called reconstruction, which can partially recover the initial density field, unlocking the information that once existed on small scales. 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; Zhu et al. 2017; Wang et al. 2017; Shi et al. 2018; Sarpa et al. 2019; Mao et al. 2021). Similarly, reconstructing the primordial power spectrum from the observed galaxy catalogues might be beneficial for probing the primordial features by using future LSS surveys, which is what we set out to check here.

In this work, as a first step towards assessing the potential benefit of reconstruction, we assume additional simple oscillatory features in the power-law primordial power spectrum. By utilising a suite of large N-body simulations, we study the performance of a nonlinear reconstruction algorithm proposed recently by Shi et al. (2018); Birkin et al. (2019) to retrieve the lost primordial features from the mock galaxy catalogue. In particular, we carry out parameter fitting to the damped and reconstructed wigglest o assess whether reconstruction can lead to more robust constraints on the feature parameters. To investigate the impact of reconstruction in the real galaxy surveys, we also forecast the constraints on the feature parameters for the DESI survey using the Fisher matrix approach, and compare the cases with and without reconstruction.

This paper is organised 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 reconstruction method to retrieve the lost 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 survey. In Section 4 we show the results of 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 nonlinear reconstruction method which will be used to recover the small-scale oscillation features from evolved dark matter and halo fields. Finally, we describe the analytic 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 powerlaw-type primordial power spectrum to be our fiducial no-wiggle model, which is given by

$$P_{nw}(k) = A_s \left( \frac{k}{k_s} \right)^{n_s-1},$$

(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 $k_s = 0.05 \text{ Mpc}^{-1}$. Motivated by Ballardini et al. (2020), we consider three wiggled models which share the same form of oscillatory features, of which the featured primordial power spectrum is expressed as

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

(2)

where $A$, $\omega$ and $\phi$ are respectively the amplitude, frequency and phase of the oscillation, $m$ is the power of the comoving wavenumber $k$.

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 reconstruction can improve the accuracy of finding the feature parameters.

The oscillation parameters of the four models are listed in Table 1. The initial oscillations of the three wiggled models are shown in the red dashed lines in the right panel of Fig. 1, where we have presented the difference between $P_w$ and $P_{nw}$. Within our interested range of scales, $k = (0.05 - 0.5) \text{ hMpc}^{-1}$, Model 1 has three peaks, Model 2 and Model 3 have only two peaks. The first peak of Model 2 is chosen to be on a smaller scale than that of Model 1. The two peaks of Model 3 are at the same position as the first and third peaks of Model 1.

By comparing the reconstructed wiggles of the three wiggled models later, we would be able to comprehend the effect of the reconstruction method on different scales.

2.2 N-body simulations

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

We have run four simulation runs including the no-wiggle model and three wiggled models. First we assume a flat universe and adopt Planck 2018 cosmology, with $h = 0.674$, $\Omega_m = 0.3135$, $\Omega_k h^2 = 0.120$, $\Omega_b h^2 = 0.0224$, $\Omega_A = 0.6865$, $n_s = 0.965$ and $A_s = 2 \times 10^{-9}$ (Aghanim et al. 2020). The value of $\sigma_8$ is approximately 0.79 though it varies a little bit across different models. We then customise the function of the primordial power spectrum in the Einstein-Boltzmann solver code camb (Lewis & Challinor 2011) to be Eq. (1) for the no-wiggle model and Eq. (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
Table 1. The oscillation parameters used for the no-wiggle model and three 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  | $\omega / \pi$ | $\phi / \pi$ |
|--------|----|----|----------------|--------------|
| Fiducial | 0  |    |                |              |
| Model 1 | 1  | 0.05 | 15.00         | 0            |
| Model 2 | 1  | 0.05 | 8.57          | 0            |
| Model 3 | 0.631 | 0.05 | 7.13          | 0            |

the publicly available code 2LPT ric (Croce et al. 2006) to generate the initial conditions used for the N-body simulations. In the left panel of Fig. 1 we compare the initial matter power spectrum given by CMB and the matter power spectrum measured from the initial conditions generated using 2LPT ric; 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

$$P_{w}(k) = \frac{P_{w}(k)}{P_{nw}(k)} - 1,$$

which are shown in the right panel of Fig. 1. This clearly shows that the oscillatory features are perfectly created in the initial conditions of the simulations.

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 $1024 \ h^{-1}\text{Mpc}$, and we output four snapshots at different redshifts, respectively as $z = 0, 0.5, 1, \text{ and } 1.5$. For each snapshot, we use the halo finder ROCKSTAR (Behroozi et al. 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 analytic formulae in Tinker et al. (2008). For each snapshot we establish one dark matter particle catalogue (hereafter DM) and two mock halo catalogues respectively with the number density of $1 \times 10^{-3} (h^{-1}\text{Mpc})^{-3}$ (hereafter H1) and $5 \times 10^{-4} (h^{-1}\text{Mpc})^{-3}$ (hereafter H2). Both host haloes and subhaloes are included in the halo catalogues. 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 nonlinear matter power spectrum from DM and nonlinear 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 $P_{tw}$ for all cases.

### 2.3 Reconstruction

In order to partially recover the primordial features lost in the structure formation, we perform reconstruction of the initial density field from the late-time density field using the nonlinear reconstruction algorithm described in Shi et al. (2018). This reconstruction method is based on mass conservation. Without assuming any cosmological model or having free parameters except the size of the mesh used to calculate the density field, it uses multigrid Gauss-Seidel relaxation to solve the nonlinear 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% correlated with the initial density field for $k \leq 0.6 h\text{Mpc}^{-1}$, if reconstruction is performed on the dark matter density field, which cover the scales of our interest, though the performance becomes poorer when the method is applied on density fields calculated from sparse tracers (Birkin et al. 2019; Wang et al. 2020; Liu et al. 2020). This method is implemented in a modified version of the EPOS 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. 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 EPOS 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 TSC 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 EPOS.

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

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

where $\xi_{hh}(r)$ is the auto-correlation function of haloes and $\xi_{hm}(r)$ is the cross-correlation function between the haloes and the dark matter particles. We use the publicly available MOCCA code (Alonso 2012) to measure $\xi_{hh}(r)$ and $\xi_{hm}(r)$ 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 $r \geq 10 h^{-1}\text{Mpc}$ 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, EPOS 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 TSC 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 subsection, we describe how this is achieved by using a fitting function.

Instead of fitting the absolute matter and halo power spectra, we propose an analytic formula to directly fit the quantity $P_{\text{rw}}$ obtained from the simulations and reconstructions, cf. Eq. (3), which combines the oscillatory feature model and a Gaussian damping function, given by

$$P_{\text{rw}}(k, z) = A \cos(\omega k^m + \phi) \exp\left[-\frac{k^2 \zeta(z)^2}{2}\right],$$  

(5)

where $\zeta(z)$ is the damping parameter that depends on the redshift $z$. For the fitting of each measured $P_{\text{rw}}(k)$, we let $\omega$, $\phi$, and $\zeta$ be the free parameters because $\omega$ and $\phi$ play an essential role in determining the position of the peaks, and $\zeta$ quantifies the extent of the damping effect. The parameters $A$ and $m$ are taken to be their theoretical values in Table 1. We apply the least-squares estimator to obtain the best-fit parameters by minimising

$$\chi^2 = \sum_{i=1}^{N} \left[ P_{\text{rw},i}(z) - P_{\text{rw}}(k, z; \omega, \phi, \zeta) \right]^2,$$  

(6)

where $P_{\text{rw},i}(z)$ are the data points of wiggle spectrum in the $i$th $k$ bin at redshift $z$. Since there is only one realisation of simulation for each model, we assume that the uncertainties of all data points $P_{\text{rw},i}(z)$ are the same and follow the same Gaussian distribution. Note that, as the quantity we fit is $P_{\text{rw}} = P_{\text{mod}}/P_{\text{rw}} - 1$, this is equivalent to doing the fitting of $P_{\text{rw}}$ with $\sqrt{P_{\text{rw}}}$ as uncertainty (e.g., Feldman et al. 1994).

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

### 3 FORECAST FOR THE DESI SURVEY

In order to investigate the impact of reconstruction, we will forecast the constraints on the feature parameters for the DESI 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 analytic marginalisation. Finally, we give the specifications of the DESI survey.

#### 3.1 Modelling the observed galaxy power spectrum

By combining the Eqs. (3) and (5), the featured galaxy power spectrum in real space can be modelled as,

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

(7)

where $P_{\text{nl}}(k, z)$ is the nonlinear matter power spectrum without the primordial oscillatory features at $z$, which includes the BAO wiggles and is equivalent to the nonlinear matter power spectrum of the no-wiggle model. However, since there is only one simulation realisation for a single no-wiggle model, which cannot provide a smooth nonlinear matter power spectrum, and since a fast method to get $P_{\text{mod}}$ is more convenient in the Fisher analysis, we use the halo fitting model in the camb code to calculate $P_{\text{mod}}(k, z)$ instead.

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 $P_{\text{obs}}(k, z)$ to the modellled galaxy power spectrum $P_{\text{mod}}(k, z)$, the standard practice is to project the galaxies to their comoving positions assuming some reference cosmology via the coordinate transformation based on the relations

$$k_{\parallel} = \frac{D_A(z)}{D_A(z)^{\text{ref}}(z)} k_{\parallel}^{\text{ref}}, \quad k_{\perp} = \frac{H(z)}{H(z)^{\text{ref}}} k_{\perp},$$  

(8)

where $k_{\parallel}$ and $k_{\perp}$ are respectively the light-of-sight and transverse components of the wavevector $k$, i.e., $k^2 = k_{\parallel}^2 + k_{\perp}^2$, the superscript $\text{ref}$ denotes the reference cosmology, note that the reference cosmology hereafter is the same one used in the simulations unless otherwise stated; $D_A(z) = a(z)/(1 + z)$ is the angular diameter distance at $z$ with the comoving distance $r(z)$; under the assumption of

---

**Figure 1.** [Colour Online] The left 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 bottom up they are respectively the fiducial model, Model 1, Model 2 and Model 3, each model is shifted upwards by a factor of 10 successively to avoid the clutter of all curves. The right panel shows the $P_{\text{rw}}$ results, cf., Eq. (3), obtained from the left panel for the three wiggled models, for instance, the bottom curves show the ratio of Model 1 to the fiducial model, followed by the ones for Model 2 and Model 3 upwards; each model is shifted upwards by a constant of 0.15 successively for the same reason as above.
flat universe it is given by
\[ r(z) = \frac{c}{H_0} \int_0^z \frac{dz'}{H(z')} \left[ \Omega_m (1 + z')^3 + \Omega_r + \Omega_L \right]^{-\frac{1}{2}}, \] (9)
where \( \Omega_r = 1 - \Omega_m - \Omega_L \) is the current density parameter of the cosmological constant, and the Hubble parameter \( H(z) \) is given by
\[ H(z) = H_0 \left[ \Omega_m (1 + z)^3 + \Omega_L \right]^{\frac{1}{2}}. \] (10)
Along with several main factors being considered, i.e., the redshift-space distortions (RSD) and shot noise, we can model the observed galaxy power spectrum as
\[ P_{\text{mod}}(k, \mu, z) = \frac{\Delta^2}{2 \pi^2} P_{\text{mod}}(k, \mu, z) + N_{\text{gal}}(z), \] (11)
where \( \sigma_8(z) \) is the R.M.S. linear density fluctuations on the scale of \( 8h^{-1}\text{Mpc} \), \( N_{\text{gal}}(z) \) is the shot noise with \( \bar{n}_g(z) \) being the galaxy number density, and the Finger-of-God factor \( F_{\text{FOG}}(k, \mu, z) \) describing the effect of RSD is modelled as Ballardini et al. (2020)
 \[ F_{\text{FOG}}(k, \mu, z) = \frac{[b(z) \sigma_8(z) + f(z) \sigma_8(z) \mu^2]^2}{1 + k^2 \mu^2 \sigma_8^2 / 2} \exp \left( -k^2 \mu^2 \sigma_8^2 / 2 \right), \] (12)
where \( b(z) \) is the linear halo bias at \( z \),
\[ f(z) = \frac{\ln D(a)}{\ln a}. \] (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 normalise \( D(a) \) so that \( D(a = 1) = 1 \) in this work), \( \mu = \cos \theta \) with \( \theta \) being the angle between the wavevector \( k \) and the line of sight, i.e., \( \mu = k \cos \theta / k \), \( \sigma_r(p) = \sigma_p / [\Omega(z)a] \) is the distance dispersion corresponding to the physical velocity dispersion \( \sigma_p \) whose fiducial value is taken to be \( 290 \, \text{km} \, \text{s}^{-1} \), and the last exponential damping factor accounts for the redshift effect \( \sigma(z) \) with \( \sigma_{r, z} = c \sigma(z) / H(z) \).

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 \( z = z_c \), can be approximated as
\[ F_{ij}(z_c) = \frac{V_{\text{eff}}(z_c)}{4\pi^2} \int_0^1 \frac{d\mu}{2} \int_{k_{\text{min}}}^{k_{\text{max}}} \frac{dk}{k} \frac{\partial \ln P_{\text{obs}}(k, \mu, z_c)}{\partial \theta_i} \frac{\partial \ln P_{\text{obs}}(k, \mu, z_c)}{\partial \theta_j}, \] (14)
where \( k_{\text{min}}, k_{\text{max}} \) are respectively the minimum and maximum values of \( k \) used for the forecast. We set \( k_{\text{min}} = 0.05 \, h^{-1}\text{Mpc} \) and adopt two values of \( k_{\text{max}} \), respectively 0.25 \, h^{-1}\text{Mpc} \) and 0.5 \, h^{-1}\text{Mpc} \), to compare the constraints for different range of scales. The effective volume of the redshift bin \( V_{\text{eff}}(z_c) \) is expressed as
\[ V_{\text{eff}}(z_c) = \left[ 1 + \frac{1}{\bar{n}_g(z) P_{\text{obs}}(k, \mu, z)} \right]^{-2} V_{\text{surv}}(z_c), \] (15)
where \( \bar{n}_g(z) P_{\text{obs}}(k, \mu, z) \) is the signal-to-noise, the comoving survey volume \( V_{\text{surv}}(z_c) \) with the redshift bin width \( \Delta z \) is given by
\[ V_{\text{surv}}(z_c) = \frac{4\pi}{3} \left[ r(z_c + \frac{\Delta z}{2})^3 - r(z_c - \frac{\Delta z}{2})^3 \right] \Omega_{\text{surv}} / \Omega_{\text{sky}}, \] (16)
where \( \Omega_{\text{surv}} \) and \( \Omega_{\text{sky}} \) are respectively the survey area and the area of the full sky. Additionally, \( \theta \) is the 8-dimensional parameter vector which consists of five cosmological parameters and three oscillation parameters,
\[ \omega_c = \Omega_h h^2, \omega_b = \Omega_b h^2, n_s, A_s, A_s, A_s, \omega, \phi. \] (17)
The partial derivatives of \( P_{\text{obs}}(k, \mu, z_c) \) w.r.t. the cosmological parameters are calculated numerically using the finite difference,
\[ \frac{\partial P_{\text{obs}}(k, \mu, z_c)}{\partial \theta_i} = \frac{P_{\text{obs}}(\theta_i^{\text{fid}} + \Delta \theta_i) - P_{\text{obs}}(\theta_i^{\text{fid}} - \Delta \theta_i)}{2\Delta \theta_i}, \] (18)
where \( \Delta \theta_i \) is taken to be 10% of the fiducial value of \( \theta_i^{\text{fid}} \), 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 analytic form of the oscillations.

The Fisher matrices of all redshift bins are then summed up to return a \( 8 \times 8 \) Fisher matrix, the inverse of which gives the uncertainties of all parameters. As we mainly aim at the constraints on the oscillation parameters, we marginalise the cosmological parameters using the analytic marginalisation method given by Taylor & Kitching (2010), which can marginalise the nuisance parameters and preserve the information about the target parameters. The marginalised Fisher matrix is identified as
\[ F_{ij}^{\text{M}} = F_{ij} - F_{am} F_{mn}^{-1} F_{nb}, \] (19)
where the subscripts \( a \) and \( b \) denote the target parameters, while \( m \) and \( n \) denote the nuisance parameters. Finally, we get the uncertainties of the oscillation parameters from the covariance matrix, i.e., the inverse of the marginalised 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 survey (DESI Collaboration et al. 2016) are discussed here.

We start with the most crucial parameter, the damping parameter \( \zeta \) 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 \( \zeta \) for four redshifts, i.e., \( z = 0, 0.5, 1, 1.5 \), and two different halo number densities, i.e., \( n_{\text{halo}} = 1 \times 10^{-3} (h^{-1}\text{Mpc})^{-3} \) and \( 5 \times 10^{-4} (h^{-1}\text{Mpc})^{-3} \), but the forecasted number density achievable in the DESI survey varies over the redshift range, so the values of \( \zeta \) 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 \( 5 \times 10^{-4} (h^{-1}\text{Mpc})^{-3} \). We use a bilinear interpolation between the redshift and the number density to estimate an appropriate value of \( \zeta \) for a given combination of the redshift and number density. For those the number density is larger than \( 1 \times 10^{-3} (h^{-1}\text{Mpc})^{-3} \) or smaller than \( 5 \times 10^{-4} (h^{-1}\text{Mpc})^{-3} \), we simply adopt the values of \( \zeta \) for \( n_{\text{halo}} = 5 \times 10^{-4} (h^{-1}\text{Mpc})^{-3} \) instead. In this work, we use different values of \( \zeta \) 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 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 $z = (0.6 - 1.3)$ for ELGs and $z = (0.6 - 0.9)$ 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 survey, the fixed values of $\pi_0(z)P_{\text{DM}}(0.14, 0.6, z)$ are used for the signal-to-noise, two survey areas are considered including the expected survey area of $14,000\, \text{deg}^2$ and $9,000\, \text{deg}^2$ as the pessimistic case (DESI Collaboration et al. 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 et al. 2016)

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

The DESI survey defines the redshift error as $\sigma(z) = 0.0005/(1+z)$, in this case, the exponential damping factor is very close to 1 for our intended range of scales, so we neglect it in the calculation.

4 RESULTS AND DISCUSSION

In this section, we will first compare the linear, nonlinear and reconstructed $P_{\text{rw}}$ measured for all models and redshifts. Then we present the results of the analytic 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 survey.

4.1 Comparisons among wiggle spectra

In Fig. 2 we compare the results of the linear, nonlinear and reconstructed $P_{\text{rw}}(k)$ obtained from DM, H1 and H2 at the four redshifts for the three wiggled models. The black solid lines represent the linear $P_{\text{rw}}(k)$ obtained from the initial conditions of the simulations, which are equivalent to the primordial oscillatory features. The blue dashed lines represent the nonlinear $P_{\text{rw}}(k)$ obtained from the output snapshots of the simulations, which are also referred to as the unreconstructed $P_{\text{rw}}(k)$ 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 $P_{\text{rw}}(k)$ obtained from the reconstructed density field, which helps to partially retrieve the lost wiggles.

The $P_{\text{rw}}(k)$ 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 seems that the scale at which the wiggles start to be weakened becomes larger as time progresses, though the specific range differs a bit for different models due to the discrepancies in their original shape of the oscillations. For instance, deviation from linear theory at $z = 0$ is at $k \geq 0.09\, \text{hMpc}^{-1}$ for Model 1, $k \geq 0.07\, \text{hMpc}^{-1}$ for Model 2 and $k \geq 0.08\, \text{hMpc}^{-1}$ for Model 3. Furthermore, the wiggles on scales $k \geq 0.3\, \text{hMpc}^{-1}$ are almost totally lost at $z = 0$, and so the recovery of the wiggles on scales $0.3 \leq k \leq 0.5\, \text{hMpc}^{-1}$ would be an important objective of reconstruction. By comparing the reconstructed $P_{\text{rw}}$ results with the linear-theory prediction, it can be seen that, despite some imperfection, the reconstruction still to a large extent achieves this by retrieving the initial oscillations on our interested scales, i.e., $0.05 \leq k \leq 0.5\, \text{hMpc}^{-1}$.

The success of the reconstruction from the dark matter particles is largely thanks to their high number density, which allows the late-time nonlinear density field to be accurately produced: in this sense, reconstruction from DM can be considered as an idealised case or an upper limit, which is difficult to achieve in real observations. For a rough comparison, we have shown, in the middle and right columns of Fig. 2, the $P_{\text{rw}}$ 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 realisations 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 $P_{\text{rw}}(k)$ 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 $P_{\text{rw}}$ results of H1 seem slightly better compared to those of H2, as a result of the slightly larger number densities in H1, though the difference is again insignificant visually; we will revisit this point when discussing the analytic fit in the next subsection. Comparing the results with and without reconstruction, it is clear that the former does give less damped and sharper oscillation features, confirming that reconstruction can help to partially recover the lost wiggles. This recovery is more substantial at lower redshifts than at higher redshifts, since at higher redshifts there is little damping to start with. At lower redshifts, on the other hand, reconstruction can even recover some of the wiggles at $k \sim 0.5\, \text{hMpc}^{-1}$ where there is literally nothing in the unreconstructed case. We expect that this will greatly help in the accurate measurements of wiggle parameters, especially in models with few wiggles at $k \lesssim 0.3\, \text{hMpc}^{-1}$ – we will discuss this in the parameter fittings next.

4.2 Wiggle parameter fitting

Figs. 3, 4 and 5 show, respectively, the results of the analytic fit to the unreconstructed and reconstructed $P_{\text{rw}}$ results for the three models studied in this work. It can be seen that, in most cases, the analytic model Eq. (5), with a Gaussian damping function characterised by the parameter $\xi(z)$, fits the data very well.

The corresponding best-fit parameters of $\omega, \phi$ and $\xi(z)$, as well as their uncertainties, are displayed in Table 2, which assist the understanding from a quantitative perspective. As mentioned before, we mainly focus on the results of H1 and H2, and thus 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 $\xi$ effectively describes the extent of the nonlinear effects in structure formation and characterises the suppression of the primordial oscillations. It is zero in the linear regime, such as at $z = 49$, and gradually increases as the redshift decreases because the structures become progressively more nonlinear. Thus reconstruction aims to reduce $\xi$ and retrieve the primordial oscillations. In Table 2 it shows that, despite the reconstructed values of $\xi$ are not reduced to zero due to the imperfection of reconstruction, they are evidently smaller than the unreconstructed values in all cases, and the uncertainties of $\xi$ are also reduced after reconstruction in most low-redshift ($z < 1$) cases, which confirms that the reconstruction succeeds in retrieving the lost wiggles to a great extent. Specifically, by comparing the cases among different models but 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 depend on the shape of
The primordial oscillations\footnote{This makes sense given that the amplitude of the primordial oscillations is small in this work, and so the wiggles can be considered as small perturbations to the primordial density field. Reconstruction, on the other hand, is sensitive to the overall distribution of matter.}. This is similar to the unreconstructed $\zeta$, which supports the qualitative inference in the previous subsection that the nonlinear regime is similar at the same redshift for different models, although the unreconstructed values of $\zeta$ in Model 1 are commonly a bit larger than those in Model 2 and Model 3. Additionally, for each model, the reconstructed values of $\zeta$ in H1 are smaller than those in H2 at the same redshift, and the same trend can be seen in the unreconstructed values as well, which is attributed to that H1 has twice the halo number density as H2. Furthermore, for each catalogue, it appears that the reconstructed $\zeta$ is only reduced with increasing redshift at low redshift ($z < 1$), it cannot be further reduced at high redshift ($z > 1$). Taking Model 1 as an example, it could also be found in Fig. 3 that the fitting curves of reconstructed wiggles at high redshift show very little difference compared with the low-redshift cases. By contrast, the unreconstructed $\zeta$ decreases with increasing redshift, thus the difference between the reconstructed and unreconstructed $\zeta$ seems to be large at low redshift and small at high redshift. In other words, the improvement given by the reconstruction is effective at low redshift but relatively limited at high redshift.

Next, we shall focus on whether the improved wiggles after reconstruction can lead to more accurate measurements of the oscillation parameters $\omega$ and $\phi$. Regarding the oscillation frequency $\omega$, the reconstructed values of $\omega$ 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. Specifically, when comparing amongst different models, it appears that the reconstructed $\omega$ values of Model 1 show slightly better performances over those of Model 2 and Model 3, which is because the reconstructed wiggles of Model 1 have four evident peaks within the fitting range of scales at all redshifts, while Model 2 and Model 3 only have two; clearly more peaks can enhance the accuracy of the fit. Similarly, the unreconstructed values of Model 1 seem to be closer to the theoretical values than those of Model 2 and Model 3; more details will be discussed later. For each model, the reconstructed values of H1 are a little bit better than those of H2 at the same redshift in most low-redshift cases, which can be explained by the larger halo number density of H1. By contrast, there is no evident distinction of

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{[Colour Online] Comparisons among the linear (black solid line), nonlinear (blue dash-dotted line) and reconstructed (red dashed line) $P_{\text{rw}}$. The linear $P_{\text{rw}}$ is measured from the initial conditions generated using 2LPTic, the nonlinear $P_{\text{rw}}$ is measured from the output snapshots of the simulations, and the reconstructed $P_{\text{rw}}$ 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 bottom up respectively belong to Model 1, Model 2 and Model 3.}
\end{figure}
Figure 3. [Colour Online] The analytic fit to the unreconstructed and reconstructed $P_{rw}$ for Model 1. The black solid lines represent the measured $P_{rw}$ and the red dashed lines represent the fitting curves given by our analytic model, Eq. (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 bottom up 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 4. [Colour Online] The same as Fig. 3 but for Model 2.
Reconstruction of primordial features

Figure 5. [Colour Online] The same as Fig. 3 but for Model 3.

Table 2. The best-fit parameters of $\omega$, $\phi$ and $\zeta$ and their 95\% uncertainties for the three wiggled models studied in this work. The values of $\omega$ and $\phi$ in the table are respectively in the units of $\pi h^2$ Mpc and $\pi$, and their corresponding 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 $n_{\text{halo}} = 1 \times 10^{13} (h^{-1} \text{Mpc})^{-3}$ and H2 the halo catalogue with $n_{\text{halo}} = 5 \times 10^{12} (h^{-1} \text{Mpc})^{-3}$. Each group of six rows includes the unreconstructed and reconstructed cases for the same redshift.

| $z$ | para | DM | 0.0 | 0.5 | 1.0 | 1.5 |
|-----|------|----|-----|-----|-----|-----|
|     |      |     |     |     |     |     |
| $\omega$ | 14.70±0.24 | 14.81±0.36 | 14.67±0.37 | 7.07±0.74 | 7.84±0.34 | 7.76±0.41 | 6.47±0.16 | 6.74±0.25 | 6.43±0.24 |
| unrec $\phi$ | 0.00±0.03 | 0.01±0.05 | 0.00±0.04 | 0.09±0.08 | 0.02±0.04 | 0.04±0.05 | 0.11±0.04 | 0.06±0.06 | 0.14±0.06 |
| $\zeta$ | 7.43±0.26 | 6.95±0.40 | 7.27±0.41 | 7.23±0.95 | 6.21±0.44 | 6.90±0.53 | 6.96±0.23 | 6.62±0.35 | 6.68±0.36 |
| $\omega$ | 15.05±0.02 | 15.04±0.21 | 14.94±0.20 | 8.58±0.03 | 8.65±0.09 | 8.76±0.23 | 7.19±0.02 | 7.32±0.09 | 7.12±0.14 |
| rec $\phi$ | -0.01±0.01 | 0.00±0.05 | 0.01±0.04 | 0.00±0.01 | -0.02±0.02 | -0.04±0.04 | -0.01±0.01 | -0.05±0.03 | 0.03±0.05 |
| $\zeta$ | 2.06±0.03 | 3.58±0.28 | 4.00±0.25 | 2.05±0.05 | 3.41±0.13 | 4.07±0.32 | 2.11±0.04 | 3.41±0.14 | 3.75±0.22 |
| $\omega$ | 14.83±0.09 | 14.58±0.20 | 14.64±0.26 | 8.06±0.29 | 8.01±0.23 | 8.08±0.22 | 6.75±0.06 | 6.74±0.15 | 6.51±0.17 |
| unrec $\phi$ | 0.00±0.01 | 0.05±0.03 | 0.02±0.04 | 0.02±0.04 | 0.03±0.03 | 0.02±0.03 | 0.07±0.02 | 0.06±0.05 | 0.12±0.05 |
| $\zeta$ | 5.88±0.10 | 5.72±0.23 | 6.18±0.30 | 5.53±0.40 | 5.27±0.32 | 5.71±0.29 | 5.33±0.08 | 4.96±0.23 | 5.58±0.26 |
| $\omega$ | 15.04±0.02 | 15.03±0.08 | 14.95±0.13 | 8.58±0.02 | 8.61±0.07 | 8.64±0.15 | 7.19±0.02 | 7.23±0.07 | 7.26±0.11 |
| rec $\phi$ | -0.01±0.01 | 0.00±0.02 | 0.01±0.03 | 0.00±0.01 | -0.01±0.02 | -0.01±0.03 | -0.01±0.01 | -0.03±0.03 | -0.03±0.04 |
| $\zeta$ | 1.51±0.04 | 3.12±0.10 | 3.83±0.17 | 1.52±0.04 | 3.17±0.10 | 3.75±0.20 | 1.53±0.04 | 3.04±0.11 | 3.46±0.16 |
| $\omega$ | 14.90±0.04 | 14.73±0.17 | 14.80±0.17 | 8.29±0.13 | 8.15±0.16 | 8.05±0.19 | 6.88±0.02 | 6.72±0.13 | 6.63±0.12 |
| unrec $\phi$ | 0.00±0.01 | 0.02±0.03 | 0.00±0.03 | 0.01±0.02 | 0.01±0.03 | 0.01±0.03 | 0.05±0.01 | 0.08±0.04 | 0.10±0.03 |
| $\zeta$ | 4.73±0.05 | 4.90±0.21 | 5.48±0.20 | 4.39±0.18 | 4.38±0.22 | 5.08±0.26 | 4.19±0.03 | 4.30±0.20 | 4.99±0.17 |
| $\omega$ | 15.01±0.01 | 15.02±0.10 | 15.02±0.15 | 8.58±0.01 | 8.60±0.07 | 8.52±0.14 | 7.19±0.01 | 7.12±0.09 | 7.03±0.10 |
| rec $\phi$ | 0.00±0.01 | 0.00±0.02 | -0.01±0.03 | 0.00±0.01 | -0.01±0.02 | -0.01±0.03 | -0.01±0.01 | -0.02±0.03 | -0.04±0.04 |
| $\zeta$ | 1.11±0.03 | 3.23±0.14 | 3.81±0.20 | 1.09±0.04 | 3.09±0.11 | 3.56±0.20 | 1.12±0.03 | 3.04±0.13 | 3.55±0.15 |
| $\omega$ | 14.94±0.02 | 14.82±0.17 | 14.64±0.22 | 8.43±0.07 | 8.17±0.14 | 7.84±0.18 | 6.95±0.01 | 6.81±0.15 | 6.70±0.14 |
| unrec $\phi$ | 0.00±0.01 | 0.01±0.03 | 0.03±0.03 | 0.00±0.01 | 0.02±0.03 | 0.07±0.03 | 0.04±0.01 | 0.05±0.05 | 0.08±0.04 |
| $\zeta$ | 3.90±0.02 | 4.43±0.21 | 5.17±0.26 | 3.60±0.10 | 4.07±0.20 | 4.92±0.25 | 3.41±0.02 | 3.92±0.22 | 4.73±0.21 |
| $\omega$ | 15.01±0.01 | 15.07±0.09 | 14.94±0.23 | 8.58±0.01 | 8.49±0.10 | 8.60±0.23 | 7.17±0.01 | 7.19±0.11 | 7.18±0.12 |
| rec $\phi$ | 0.00±0.01 | -0.01±0.02 | 0.01±0.05 | 0.00±0.01 | 0.02±0.02 | -0.01±0.05 | 0.00±0.01 | -0.02±0.04 | -0.01±0.04 |
| $\zeta$ | 0.88±0.03 | 3.24±0.12 | 3.90±0.29 | 0.84±0.03 | 3.08±0.15 | 3.66±0.32 | 0.83±0.02 | 3.19±0.16 | 3.65±0.17 |
the unreconstructed values at the same redshift between H1 and H2, which could result from the non-negligible noise in the wiggle spectra. It can be checked by having more realisations of the simulations in the future. Moreover, for each catalogue, the reconstructed values become closer to the theoretical values as the redshift increases in the low-redshift range, but it cannot be further improved at high redshift, which is similar to how the redshift affects the reconstructed $\zeta$ as we found above. By contrast, it is shown that the unreconstructed values of DM become more accurate with increasing redshift, but those of H1 and H2 do not show the same trend, which could be caused by the evident noise in the wiggle spectra due to the wide gap between the number densities of dark matter particles and haloes. Therefore, the reconstruction mainly improves the prediction of $\omega$ at low redshift, and the predictions for Model 2 and Model 3 are improved more than those for Model 1.

The situation is quite different in the case of the oscillation phase $\phi$. The unreconstructed values of Model 1 and Model 2 are determined very well in most cases, so the reconstructed values only show a little bit improvement on the unreconstructed $\phi$ even for low-redshift cases. However, for Model 3 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. Therefore the reconstruction once again shows its advantage of allowing more accurate measurement of $\phi$, especially for H2 at low redshift.

When combining the results of all three parameters, it seems that the reconstruction is most useful at low redshifts, $z < 1$, and Model 2 and Model 3 benefit more from it than Model 1. Although the remaining peaks of Model 1 are kept very well so that its reconstructed results are better than those of the other two models, the improvement is relatively limited for it. Thus the improvement depends less on how great the reconstructed wiggles are, and more on how poor the primordial wiggles are kept before reconstruction; in other words, the reconstruction would be more useful if the primordial wiggles are lost to a greater extent. As we mentioned before, the wiggles on scales $k \gtrsim 0.3$ $h$ Mpc$^{-1}$ are totally lost at $z = 0$. Model 1 has exactly the first two original peaks outside this range of scales, so these two peaks are effectively preserved at low redshift. By contrast, Model 3 has one original peak at the same position of the first peak of Model 1 which is effectively preserved, and its second peak is at the same position of the third peak of Model 1, which is almost lost. Hence the primordial wiggles of Model 3 are preserved less well than those of Model 1, and Model 3 would benefit more from the reconstruction. Similarly, Model 2 has two original peaks in the range $k \lesssim 0.5$ $h$ Mpc$^{-1}$: the first is on a smaller scale compared with the first peak of other two models, and so it is not preserved as well as the first peak of the other two models due to the larger damping effect, while the second peak is totally wiped out. Thus the primordial wiggles of Model 2 are kept even worse than those of Model 3. However, since the fitting range of scales includes the first trough to the left of the first peak for Model 2 but not for Model 3, this partially balances the accuracy of the fit for Model 2. Therefore the improvement by the reconstruction is similar for Model 2 and Model 3, and both benefit from reconstruction substantially more than Model 1.

### 4.3 Constraints on oscillation parameters for DESI-like survey

Since the results of the constraints on the oscillation parameters are similar between Model 2 and Model 3, we take Model 1 and Model 2 as two examples to illustrate and discuss how the reconstruction could improve the constraints in a real galaxy survey. We shall first talk about some common features exhibited in both models, and then clarify the distinctions between them. Finally, we also forecast how much the uncertainties of the feature amplitude can be improved after reconstruction for three models.

Figs. 6 and 7 show, respectively, the 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 and Model 2. The marginalised likelihoods of the oscillation parameters shown in the upper panels of both figures indicate that the unreconstructed cases with $k_{\text{max}} = 0.5$ $h$ Mpc$^{-1}$ (red lines) give better constraints than the unreconstructed cases with $k_{\text{max}} = 0.25$ $h$ Mpc$^{-1}$ (grey), because in the former case more $k$ modes are included in the Fisher matrix which increase the accuracy of the constraints. Additionally, by comparing the likelihoods of the same $k_{\text{max}}$, we find that reconstruction leads to more robust constraints on the parameters, because it successfully recovers some of the lost peaks in the nonlinear regime and provides more effective $k$ modes. Furthermore, stronger constraints are shown for ELGs (right panels) compared with LRGs (left panels), since the former has more available redshift bins and larger halo number density for the same redshift bins.

The above trends are also shown in the marginalised 2D confidence contours in the lower part of the corner plots in Fig. 6 and 7. In particular, it can be seen in most contours that the cases for $k_{\text{max}} = 0.5$ $h$ Mpc$^{-1}$ are largely improved by reconstruction as the boundary of 68 % limits shrinks to be around the boundary of 95 % limits after reconstruction. It is because the peaks on scales $k \gtrsim 0.2$ $h$ Mpc$^{-1}$ are heavily damped at low redshift, and the recovered wiggles at $k = (0.2 - 0.5)$ $h$ Mpc$^{-1}$ significantly contribute to the constraints. By contrast, since the peaks on scales $k \lesssim 0.2$ $h$ Mpc$^{-1}$ are preserved very well, the recovery of wiggles for $k_{\text{max}} = 0.25$ $h$ Mpc$^{-1}$ is not as important as in the case for $k_{\text{max}} = 0.5$ $h$ Mpc$^{-1}$. Besides, the $A$-$\omega$ and $A$-$\phi$ contours for $k_{\text{max}} = 0.25$ $h$ Mpc$^{-1}$ show that these parameters are degenerate with each other, and these degeneracies are broken and replaced with stronger constraints when $k_{\text{max}} = 0.5$ $h$ Mpc$^{-1}$, which include more $k$ modes.

However, in all cases the $\omega$-$\phi$ contours show that the two parameters are strongly degenerate. It is understood from the $P_{\text{rw}}$ results that the region $k = (0.1 - 0.3)$ $h$ Mpc$^{-1}$ dominates the results of constraints, because on larger scales ($k \lesssim 0.1$ $h$ Mpc$^{-1}$) the uncertainty is large due to cosmic variance, while on smaller scales ($k \gtrsim 0.3$ $h$ Mpc$^{-1}$) the wiggles are damped which have less influence on the constraints. We have checked explicitly that combinations of $\omega$ and $\phi$ along the degeneracy direction shown in Figs. 6 and 7 lead to little change to the $P_{\text{rw}}$ curves in the range of $k = (0.1 - 0.3)$ $h$ Mpc$^{-1}$. This is a feature of the oscillation model itself, which is why a strong degeneracy is still present even in the cases of reconstruction (albeit significantly less strong than the unreconstructed cases).

When comparing the constraints between Model 1 and Model 2, it appears that the choice of $k_{\text{max}}$ has a different influence on the constraints. For both models, the constraints are similar for $k_{\text{max}} = 0.5$ $h$ Mpc$^{-1}$ but significantly different for $k_{\text{max}} = 0.25$ $h$ Mpc$^{-1}$. For Model 2 there is a wide gap between the marginalised 1D distributions of both the unreconstructed and reconstructed cases, when increasing $k_{\text{max}}$ from 0.25 $h$ Mpc$^{-1}$ to 0.5 $h$ Mpc$^{-1}$, which could be attributed to the fact that Model 2 has few peaks at $k \lesssim 0.25$ $h$ Mpc$^{-1}$. Thus for the primordial oscillations with relatively low frequency, increasing $k_{\text{max}}$ would be very beneficial to improve the constraints. Furthermore, the improvement on the constraints by the reconstruction shows that Model 2 benefits more from the reconstruction than Model 1, which is consistent with the analysis in Section 4.2.

Lastly, as Ballardini et al. (2020), we show the marginalised uncertainties of $A$ as a function of $\omega$ for the three models in Fig. 8 and discuss the implications of the results.

First, we focus on Model 1 and Model 2 which show almost iden-
Figure 6. [Colour Online] 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 marginalised likelihoods. The middle and lower panels show the marginalised 68% and 95% confidence contours for every two out of three feature parameters. The green and grey colours represent, respectively, the cases for $k_{\text{max}} = 0.25 h\text{Mpc}^{-1}$ with and without reconstruction, while the blue and red colours represent the cases for $k_{\text{max}} = 0.5 h\text{Mpc}^{-1}$ with and without reconstruction.

Figure 7. [Colour Online] The same as Fig. 6 but for Model 2.
Figure 8. [Colour Online] Forecasts of the marginalised uncertainties of the oscillation amplitude $A$ as a function of the frequency $\omega$, for the three models. The first column is for LRGs and the second column is for ELGs. The bottom panels are for Model 1, followed by Model 2 and Model 3 upwards. 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 $k_{\text{max}}$ and same situation of reconstruction but different survey areas; the thick lines are for the survey area of $14,000 \text{ deg}^2$ and the thin lines are for $9,000 \text{ deg}^2$.

Variations of the uncertainties w.r.t. $\omega$, due to their identical form of the oscillations — we note that these are the same oscillation model with different choices of $\omega$, and thus we should have expected exactly identical constraints in Fig. 8. However, as we have discussed above, the best-fit values of $\zeta$ are slightly different (cf. Table 2), even though we expect any model dependence of $\zeta$ to be weak (this suggests that more realisations of simulations are needed to measure $\zeta$ as functions of halo number density, redshift, and model (or model parameter), etc., more accurately, which will be left for future work).

As expected, ELGs place slightly tighter constraints than LRGs due to their larger number densities and redshift range. The sharp peaks appeared at $\omega \approx 100 \text{ h}^{-1}\text{Mpc}$ are due to the degeneracy between the oscillatory features and the BAO wiggles over the signal-dominated range of scales. For $\omega \gtrsim 150 \text{ h}^{-1}\text{Mpc}$ the uncertainties do not vary with $\omega$, but for smaller $\omega$ things are complicated for different $k_{\text{max}}$. For $k_{\text{max}} = 0.25 \text{ hMpc}^{-1}$ we can see an increase in the uncertainties at $\omega \lesssim 50 \text{ h}^{-1}\text{Mpc}$, while a similar increase starts to appear at even smaller $\omega - 20 \text{ h}^{-1}\text{Mpc}$ for $k_{\text{max}} = 0.5 \text{ hMpc}^{-1}$. Thus larger $k_{\text{max}}$
has an extra advantage of significantly reducing the uncertainties for small $\omega$, in addition to giving more stringent constraints (everything else the same) for all $\omega$ overall. By comparing the pairs of curves with the same colours, i.e., the same cases ($k_{\text{max}}$ and reconstructed vs. unreconstructed) but different survey areas, we find that, as expected, a larger survey area always gives better constraints.

Most interestingly, everything else equal, doing reconstruction can significantly reduce the uncertainties of $A$. As an example, for large $\omega$ values, in the case of $k_{\text{max}} = 0.5 \ h\text{Mpc}^{-1}$ and survey area equal to $14,000 \text{deg}^2$, reconstruction reduces $\sigma(A)$ from $\sim 0.003$ to $\sim 0.002$, and this improvement is stronger than not performing reconstruction, but instead going from $9,000$ to $14,000 \text{deg}^2$ with $k_{\text{max}}$ fixed to $0.25$ or $0.5 \ h\text{Mpc}^{-1}$, or increasing $k_{\text{max}}$ from $0.25$ to $0.5 \ h\text{Mpc}^{-1}$ keeping the survey area fixed to either $9,000$ or $14,000 \text{deg}^2$. A similarly good improvement can be seen with $k_{\text{max}} = 0.25 \ h\text{Mpc}^{-1}$ or survey area equal to $9,000 \text{deg}^2$, when doing reconstruction. In certain cases, e.g., the large-$\omega$ regime of the lower panels of Fig. 8, reconstruction with $k_{\text{max}} = 0.25 \ h\text{Mpc}^{-1}$ and a survey area equal to $9,000 \text{deg}^2$ (the thin green dashed line) can lead to comparable constraints to not doing reconstruction but with $k_{\text{max}} = 0.5 \ h\text{Mpc}^{-1}$ and a survey area equal to $14,000 \text{deg}^2$ (the thick orange dot-dashed line). Given that increasing survey area is usually not possible, and increasing $k_{\text{max}}$ is also not straightforward given the effort required to model the non-linear regime of matter clustering, especially with RSD, performing reconstruction seems to be a cheap way to maximise the exploitation and scientific return of survey data.

The behaviours of Model 3 are broadly similar to those of Model 1 and Model 2, 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 3 are at slightly different values of $\omega$ 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 and Model 2 is related to the BAO peak in the matter/galaxy correlation function, which is at $\sim 100h^{-1}\text{Mpc}$. The primordial wiggles for those two models, in configuration space, correspond to a spike at matter or halo separation $r = \omega$. In those two models, when $\omega \gg 100h^{-1}\text{Mpc}$, the BAO and primordial peaks are separated afar and thus the former does not affect the accuracy of the measurement for the latter. As $\omega$ approaches $100h^{-1}\text{Mpc}$ 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 speculate that the dip — which causes the double-peak structure in $\sigma(A)$ for Model 1 and Model 2 — 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 3 is much more complicated in configuration space, because $m \neq 1$ in Eq. (2): this can cause the differences in the fine details of $\sigma(A)$ between this and the other models.

5 CONCLUSIONS

In this paper, we have investigated the effect of density reconstruction on retrieving hypothetical oscillatory features in the primordial power spectrum which are erased on small scales in the late-time Universe due to the nonlinear cosmological evolution.

We considered three different oscillatory features that are added to a simple power-law primordial power spectrum, for which we ran N-body simulations and identified dark matter halo catalogues from the output snapshots at a number of redshifts. We reconstructed the initial density field from the particle data, and halo catalogues with different number densities. Finally, we compared the fitted feature parameters from the unreconstructed and reconstructed density fields, to identify the improvement by reconstruction. We showed that reconstruction can be very effective in helping retrieve the lost wiggles — with the finite volume of our simulations, not only does it lead to much less biased best-fit values of the feature parameters, but it also substantially shrinks the measurement uncertainty. The improvement was especially strong where the primordial features have been more severely erased to start with, such as at $z < 1$.

In order to forecast the constraints on the feature parameters for a typical DESI-like galaxy survey, we modelled the observed broad-band galaxy power spectrum based on the HALOFIT model with the addition of our oscillatory features, then used the analytic marginalised Fisher matrix to calculate the constraints on the oscillation parameters regarding the specifications of DESI LRGs and ELGs. We found that 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 reconstruction is commonly used in improving the measurement of the BAO scale, and hence the determination of the expansion rate and properties of dark energy, this work has demonstrated that similar applications are possible in other cases where certain features in matter clustering are present. This is particularly true if these features are in the mildly nonlinear regime, $0.1 \leq k/(h\text{Mpc}^{-1}) \leq 0.5$, since this range of scales is what the nonlinear reconstruction method used here helps most: on even larger scales the benefit of reconstruction is insignificant, while on even smaller scales reconstruction won’t help much. Indeed, while a comparison with other, e.g., the standard, reconstruction methods is beyond the scope of this paper, based on experience we would expect those latter methods should also improve the prospective of constraining potential primordial features.

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 damping parameter $\zeta$ post-reconstruction. As we have seen, $\zeta$ controls the improvement over the unreconstructed case. The simulations carried out for this work — due to their limited resolution, box size, output snapshots numbers, coverage of models and number of realisations — did not allow us to more accurately quantify how $\zeta$ depends on the oscillation model (though we suspect that any dependence on parameters $A, \omega$ should be weak as long as $A$ is small), redshift, or the tracer type or number density. It certainly would be great if better simulations will become available, allowing further improvements on these aspects.

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., $k \approx 0.5 \ h\text{Mpc}^{-1}$) 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 $\approx 0.5 \ h\text{Mpc}^{-1}$. Of course, we can always cut $k_{\text{max}}$ 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 $k = 0.1 \, \text{hMpc}^{-1}$.
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

Adams J. A., Cresswell B., Easther R., 2001, Phys. Rev. D, 64, 123514
Ade P. A. R., et al., 2014a, Astron. Astrophys., 571, A22
Ade P. A. R., et al., 2014b, Astron. Astrophys., 571, A24
Ade P. A. R., et al., 2016a, Astron. Astrophys., 594, A17
Ade P. A. R., et al., 2016b, Astron. Astrophys., 594, A20
Ashdown P., Dvorkin C., Hu W., Lim E. A., 2012, Phys. Rev. D, 85, 023531
Aghanim N., et al., 2020a, JCAP, 11, 010
Aghanim N., et al., 2020b, Phys. Rev. D, 102, 063504
Komatsu E., et al., 2009, Astrophys. J. Suppl., 180, 330
LHuillier B., Shafieloo A., Hazra D. K., Smoot G. F., Starobinsky A. A., 2018, Mon. Not. Roy. Astron. Soc., 477, 2503
Lewis A., Challinor A., 2011, CAMB: Code for Anisotropies in the Microwave Background (ascl:1102.026)
Li B., Zhao G.-B., Teissier R., Koyama K., 2012, JCAP, 10, 051
Li B., Barreira A., Baugh C. M., Hellwing W. A., Koyama K., Pascoli S., Zhao G.-B., 2013, JCAP, 11, 012
Linde A. D., 1982, Phys. Lett. B, 108, 389
Liu Y., Yu Y., Li B., 2020, arXiv e-prints, p. arXiv:2012.11251
Mao T.-X., Wang J., Li B., Cai Y.-C., Falck B., Neyrinck M., Szalay A., 2021, Mon. Not. Roy. Astron. Soc., 484, 3818
Schmittfull M., Feng Y., Beutler F., Sherwin B., Chu M. Y., 2015, Phys. Rev. D, 92, 123522
Senatore L., Smith K. M., Zaldarriaga M., 2010, JCAP, 10, 028
Seo H.-J., Eisenstein D. J., 2003, Astrophys. J., 598, 720
Shi Y., Cautun M., Li B., 2018, Phys. Rev. D, 97, 023505
Spergel D. N., et al., 2007, Astrophys. J. Suppl., 170, 377
Taylor A. N., Kitching T. D., 2010, Mon. Not. Roy. Astron. Soc., 408, 865
Tegmark M., 1997, Phys. Rev. Lett., 79, 3806
Teyssier R., 2002, Astron. Astrophys., 385, 337
Tinker J. L., Kravtsov A. V., Klypin A., Abazajian K., Warren M. S., Yepes G., Gottlober S., Holz D. E., 2008, Astrophys. J., 688, 709
Wang X., Feng B., Li M., Chen X.-L., Zhang X., 2005, Int. J. Mod. Phys. D, 14, 1347
Xu J., Dvorkin C., Huang Z., Namjoo M. H., Verde L., 2019, Mon. Not. Roy. Astron. Soc., 483, 5267
Zhai G.-B., 2013, JCAP, 11, 012
Zielinska A., Harnack M., 2018, Mon. Not. Roy. Astron. Soc., 477, 2503
Zhou H.-H., Yu Y., Pen U.-L., Chen X., Yu H.-R., 2017, Phys. Rev. D, 96, 123502