

## Blinded challenge for precision cosmology with large-scale structure: Results from effective field theory for the redshift-space galaxy power spectrum

Takahiro Nishimichi<sup>1,2</sup> , Guido D'Amico<sup>3,4</sup>, Mikhail M. Ivanov<sup>5,6</sup>, Leonardo Senatore<sup>3,7</sup>,  
Marko Simonović<sup>8</sup> , Masahiro Takada<sup>2</sup> , Matias Zaldarriaga<sup>9</sup> and Pierre Zhang<sup>10,11,12</sup>

<sup>1</sup>*Center for Gravitational Physics, Yukawa Institute for Theoretical Physics,  
Kyoto University, Kyoto 606-8502, Japan*

<sup>2</sup>*Kavli Institute for the Physics and Mathematics of the Universe (WPI), UTIAS,  
The University of Tokyo, Kashiwa, Chiba 277-8583, Japan*

<sup>3</sup>*Stanford Institute for Theoretical Physics, Physics Department, Stanford University,  
Stanford, California 94306, USA*

<sup>4</sup>*Dipartimento di SMFI dell' Universita' di Parma & INFN Gruppo Collegato di Parma,  
Parma 43124, Italy*

<sup>5</sup>*Center for Cosmology and Particle Physics, Department of Physics, New York University,  
New York, New York 10003, USA*

<sup>6</sup>*Institute for Nuclear Research of the Russian Academy of Sciences,  
60th October Anniversary Prospect, 7a, 117312 Moscow, Russia*

<sup>7</sup>*Kavli Institute for Particle Astrophysics and Cosmology, SLAC and Stanford University,  
Menlo Park, California 94025, USA*

<sup>8</sup>*Theoretical Physics Department, CERN, 1 Esplanade des Particules, Geneva 23, CH-1211, Switzerland*

<sup>9</sup>*School of Natural Sciences, Institute for Advanced Study,  
1 Einstein Drive, Princeton, New Jersey 08540, USA*

<sup>10</sup>*Department of Astronomy, School of Physical Sciences, University of Science and Technology of China,  
Hefei, Anhui 230026, China*

<sup>11</sup>*CAS Key Laboratory for Research in Galaxies and Cosmology,  
University of Science and Technology of China, Hefei, Anhui 230026, China*

<sup>12</sup>*School of Astronomy and Space Science, University of Science and Technology of China,  
Hefei, Anhui 230026, China*



(Received 18 March 2020; accepted 8 December 2020; published 28 December 2020)

An accurate theoretical template for the galaxy power spectrum is key for the success of ongoing and future spectroscopic surveys. We examine to what extent the effective field theory (EFT) of large-scale structure is able to provide such a template and correctly estimate cosmological parameters. To that end, we initiate a blinded challenge to infer cosmological parameters from the redshift-space power spectrum of high-resolution mock catalogs mimicking the BOSS galaxy sample but covering a 100 times larger cumulative volume. This gigantic simulation volume allows us to separate systematic bias due to theoretical modeling from the statistical error due to sample variance. The challenge is to measure three unknown input parameters used in the simulation: the Hubble constant, the matter density fraction, and the clustering amplitude. We present analyses done by two independent teams, who have fitted the mock simulation data generated by yet another independent group. This allows us to avoid any confirmation bias by analyzers and to pin down possible tuning of the specific EFT implementations. Both independent teams have recovered the true values of the input parameters within subpercent statistical errors corresponding to the total simulation volume.

DOI: [10.1103/PhysRevD.102.123541](https://doi.org/10.1103/PhysRevD.102.123541)

### I. INTRODUCTION

Modern cosmology is becoming more mature as accumulating observational data become available. However, we have a fundamental lack of understanding of the physical nature of the dark components introduced to explain the dominant source of gravity that gathers material

---

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

to form rich structures in the late Universe (dark matter), as well as the accelerating cosmic expansion (dark energy), together filling the majority of the cosmological energy budget. To gain further insight, a growing number of large-scale observational programs are ongoing and planned (e.g., [1–4]).

Of crucial importance from a theoretical point of view is our ability to prepare an accurate model template with which one can confront such observational data for interpretation. Since a larger survey means a smaller statistical error, the relative contribution from the systematic error arising from the inaccuracy of the template is more important. Given the gigantic area coverage and depth of ambitious future programs, we need to find an accurate theoretical framework to predict the observed large-scale structure to attain their full potential to infer the underlying theory governing the Universe.

One of the most difficult aspects of the large-scale structure prediction is the complicated relationship between the matter density fluctuations dominated by invisible dark matter and visible structures such as galaxies [5]. The so-called galaxy bias cannot be predicted from first principles, unless one can model all the baryonic effects relevant for the formation and evolution of galaxies. While hydrodynamical simulations might be one way to proceed, the very large dynamical range, kpc to Gpc in length scale, is a big obstacle. Typically, one uses empirical subgrid models and calibrates them against the observed statistics of galaxies (see, e.g., [6–13] for recent attempts).

Alternatively, one can formulate the statistical properties of galaxies on large scales via a perturbative expansion in which poorly known galaxy physics is parametrized by a set of effective bias operators. The strength of these operators is controlled by free coefficients, which should be treated as nuisance parameters. The recently developed effective field theory of large-scale structure (EFTofLSS) provides a systematic way to derive all possible operators and corresponding bias coefficients that are allowed by symmetry [14–22] (see also [23] for a review). Since this approach, in principle, does not assume any specific model of galaxy formation, it provides us with a conservative theoretical model for the galaxy density and velocity fields on large scales. The generality of the effective field theory approach comes at the price of having to marginalize over many free coefficients, which can compromise cosmological constraints. These constraints can become weaker compared to other theoretical templates in which a specific bias prescription is employed, such as halo model approaches. The detailed balance between the robustness and the tightness of the cosmological constraints has been addressed in recent studies (e.g., [24–26]).

There are several nontrivial choices behind the application of the EFT to the data. First, one should determine the wave number up to which the EFT calculation up to a chosen perturbative order is reliable. This data cut should

be carefully tested to avoid biased parameter estimates. Then, one has to decide how many nuisance parameters to keep in the fit (there are about ten at the one-loop order) and what priors to use. Indeed, at the power-spectrum level many EFT operators are degenerate among each other. Thus, one has to accurately determine their principal components to make the cosmological analysis efficient. All these subtleties should be examined and validated in a transparent manner to convince the community of the robustness of the EFT approach.

To that end, in this paper, we conduct a first *blind* test of EFTofLSS for clustering of galaxies in redshift space. Two independent groups, which we refer to as “West Coast” (D’Amico, Senatore, and Zhang) and “East Coast” (Ivanov, Simonović, and Zaldarriaga), have analyzed the mock data generated by yet another group (Nishimichi and Takada, simply called “Japan team” hereafter). In this process, the true cosmological parameters used to generate the simulation mock data were known only to the Japan team. The two analyzing teams participated in the challenge under the condition that the results would be published regardless of the outcome, and the pipelines could not be modified after unblinding. We present these results in our paper in the original form. To complement the result of the blinded analysis and to get more insight on the origin of the cosmological information, we briefly discuss post-unblinding analyses.

The layout of this paper is as follows. We first describe the design of our mock challenge program in Sec. II. We then specify the mock simulations in Sec. III. The theoretical template and the method to conduct parameter inference are explained in Sec. IV. Then, the results of the blinded analysis are summarized in Sec. V. We conclude this study in Sec. VI.

## II. DESIGN OF BLINDED COSMOLOGY CHALLENGE

Throughout this paper, we consider a flat  $\Lambda$ CDM cosmology. This is motivated by the recently claimed tension in the values of the Hubble parameter, one from local measurements such as the distance ladder and the other from the cosmic microwave background (CMB) assuming a flat  $\Lambda$ CDM model (see [27] and references therein). In such a situation, a robust measurement from other independent observable channels would be important, and indeed, the galaxy clustering, when the full shape information of its spectra is analyzed, has been shown to serve as such a probe [28–31]. Also important might be a similar but weaker tension in the amplitude of the density fluctuations in the current Universe [32–34]. This is known to be degenerate with the matter density parameter from the late-time observables. We wish to demonstrate, through the challenge, the current status of the use of galaxy clustering, in particular, with an EFT approach to describe the non-linear nature of the cosmological large-scale structure.

### A. Cosmological parameters

To assess the reliability of the galaxy-clustering analyses within the flat  $\Lambda$ CDM model, three cosmological parameters,  $\ln(10^{10}A_s)$ ,  $\Omega_m$ , and  $H_0$ , are randomly drawn from independent normal distributions. These parameters are the logarithm of the amplitude of the primordial power spectrum at  $k_0 = 0.05 \text{ Mpc}^{-1}$ , the matter density parameter at present, and the current Hubble expansion rate in  $\text{km/s/Mpc}$ , respectively. While the mean values of the normal distributions are set to be the best-fit values determined by the Planck satellite [35], we consider the standard deviation 4 times larger than the same experiment to test the validity of the model in a broader parameter space. While all of the information above is shared among all the collaborators, the three random numbers drawn were kept only within the Japan team until we finally unblinded them.

On the other hand, we fix the baryon fraction  $f_b = 0.1571$  and the spectral index  $n_s = 0.9649$ . These values are shared with the two US teams. In typical current large-scale structure survey analyses, these two parameters are not very well determined due to the weak sensitivity of the target galaxy observables unless one adds priors motivated by CMB observations and/or big-bang nucleosynthesis, while it would be possible to constrain them from futuristic galaxy surveys. Therefore, letting the US analysis teams know their exact values loosely corresponds to adding CMB priors.<sup>1</sup> Further, for simplicity and to avoid the complication of dealing with massive neutrinos both in theory and in simulations, we set the neutrino masses to be exactly zero. Under the above settings, the linear matter-density transfer function is computed using the public Boltzmann solver CAMB [36]. The parameter file passed to this code by the Japan team is provided to the US teams after the values of  $\omega_b$ ,  $\omega_c$ ,  $H_0$ , and  $A_s$  are erased.

The main goal of the challenge is to infer the three cosmological parameters  $A_s$ ,  $\Omega_m$ , and  $H_0$ . It was agreed among all the teams that, once these cosmological parameters are unblinded, the results reported by that time may not be modified any more.

### B. Target observables

We focus on the galaxy clustering in redshift space in the initial challenge presented in this paper. More specifically, we work in Fourier space and analyze the multipole moments of the galaxy power spectrum. This includes physical and observational effects such as the baryon acoustic oscillations (BAOs; [37–41]), redshift-space distortions (RSD; [42,43]), and the Alcock-Paczynski (AP; [44]) effect, where the AP effect is induced artificially by distorting the simulation boxes (see the next section for

further details). In addition to these distinctive features, the mock data should contain the cosmological information through the overall shape of the power spectra, which might be hindered by the presence of various nonlinear effects. The aim of this challenge is to assess how robustly one may extract the fundamental cosmological parameters within the flat  $\Lambda$ CDM framework.

The Japan team constructs mock galaxy catalogs and measures the multipole moments of the power spectra. To discriminate the *systematic* error from the statistical error, this experiment is done in huge simulation volumes much larger than the current surveys. The galaxy catalogs are constructed to roughly mimic the CMASS and the LOWZ catalog from the 12th Data Release of Sloan Digital Sky Survey Ref. [45] (hereafter, SDSS DR12). The details of these simulations will follow in the next section. Since the galaxy bias is formulated to be as general as possible in the EFT, based only on symmetry considerations without assuming any specific model with which galaxies are defined, the details of the mock galaxies would not have a significant impact on the blinded analysis as long as one sticks to an EFT approach. However, other approaches, such as the halo model, would be directly impacted by the information on the exact procedure with which the mock galaxies are distributed within the simulation volume. Therefore, any further information on the mock galaxies detailed in the next section was not provided to the US teams before unblinding.

For completeness, the set of mock data, as well as information on the simulations provided to the US teams, is summarized at a dedicated website [46]. All the data and the information were shared through this website. Interested readers may download the same set of data and participate in the blinded challenge by analyzing the data using their own theoretical template, as the exact cosmological-parameter values are not shown in this paper nor on the website.

## III. GENERATING MOCK REDSHIFT-SPACE POWER SPECTRA OF BOSS-LIKE GALAXIES

The Japan team worked on the construction of mock galaxy catalogs and measurement of the power spectra. The settings for the numerical simulations, the prescription for the mock galaxies, and the analysis methods to determine their statistics are described in this section.

### A. Specification of simulations

We follow the gravitational dynamics of ten random realizations of the matter density field expressed by  $3072^3$  mass elements sampled in comoving periodic cubes with side length  $L = 3840 h^{-1} \text{ Mpc}$ . The total volume,  $566 (h^{-1} \text{ Gpc})^3$ , is about 100 times that of the CMASS and LOWZ samples from SDSS BOSS DR12, which together have a volume coverage of  $5.7 (h^{-1} \text{ Gpc})^3$  [47].

<sup>1</sup>It is not trivial how one can best arrange a challenge where external prior information is added. To keep the analysis fully blinded, the Japan team decided not to give any prior information to the analysis teams for the challenge presented in this paper.

The large volume of our simulations allows us to determine the statistics of the mock galaxies very precisely, with little sample-variance error. Therefore, we can conduct a fairly stringent test of the systematic error due to an imperfect modeling of the target statistics.

The initial conditions are generated with a code developed in [48] and then parallelized in [49] based on the second-order Lagrangian perturbation theory (2LPT; [50,51]). Following the result presented in [52], the starting redshifts of the simulations are set at  $z = 29$  to roughly optimize the total systematic error arising from the artificial growing mode due to the grid preinitial condition [53–55] and the truncation of the LPT at the second order, given the mean interparticle distance of the simulations. We prepare ten independent random realizations, each of which is then evolved by a public tree-particle mesh code GADGET2 [56] with  $6144^3$  grid points for fast Fourier transform (FFT) and a tree softening length of  $62.5 h^{-1}$  kpc. The other simulation parameters to control the force accuracy as well as the time-stepping criteria are the same as in [52]. We store the particle snapshots at  $z = 3, 2, 1, 0.61, 0.51$ , and  $0.38$ . We populate galaxies to the lowest three redshifts and conventionally call the catalogs as CMASS2 ( $z = 0.61$ ), CMASS1 ( $z = 0.51$ ), and LOWZ ( $z = 0.38$ ) in what follows.

## B. Mock galaxy identification

After obtaining the particle snapshots, we run the ROCKSTAR halo finder [57], which is based on the six-dimensional phase-space friends-of-friends algorithm. This code identifies not only isolated “central” halos but also “satellite” halos existing as substructures of more massive halos without any distinction at first in the primary output files. For simplicity, we treat each of them irrespectively of whether it is a central or a satellite halo and populate a galaxy only according to the virial mass assigned by ROCKSTAR. We impose a soft cutoff to the virial mass to select massive halos to populate galaxies randomly with the probability

$$P(M_{\text{vir}}) = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log_{10} M_{\text{vir}} - \log_{10} M_{\text{min}}}{\sigma_{\log_{10} M}} \right) \right], \quad (1)$$

where  $\text{erf}(x)$  is the error function. We have two parameters,  $\log_{10} M_{\text{min}}$  and  $\sigma_{\log_{10} M}$ , which determine the typical minimum mass and the profile of the soft mass cutoff, respectively. We set  $\log_{10} M_{\text{min}} = 13.08, 12.97$ , and  $12.95$  for LOWZ, CMASS1, and CMASS2 ( $M_{\text{min}}$  is given in units of  $h^{-1} M_{\odot}$ ), respectively, while the value of  $\sigma_{\log_{10} M}$  is fixed to 0.35 for all of the samples. These choices are made such that the resultant clustering signal of the mock galaxies, especially the amplitude of the power spectra at small  $k$ , becomes roughly consistent with the observation (see the next subsection for more details). We assume that

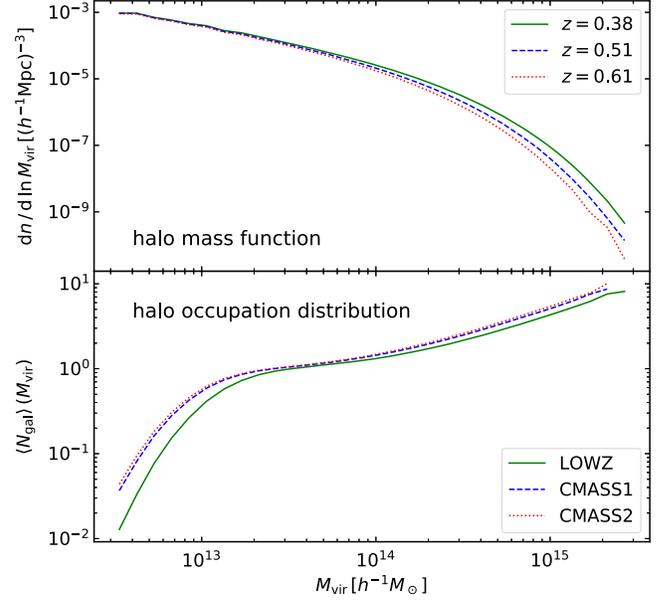


FIG. 1. Abundance of halos per unit logarithmic mass interval (upper) and mean number of mock galaxies per halo (lower) as a function of the virial mass of halos. The mean of the ten random realizations is shown at three output redshifts of the simulations as indicated by the figure legend.

the populated mock galaxies are located at the center-of-mass position of the core particles determined by ROCKSTAR. Similarly, we assign the center-of-mass velocities of the same core particles to the mock galaxies, which are used when we displace the positions of mock galaxies to redshift space [25].

We show the abundance of (central) halos as well as the mean number of galaxies per central halo as a function of the virial mass in Fig. 1. Here, we define “central” halos from the ROCKSTAR catalog as those satisfying the condition that any other halo is not more massive than the halo of interest to within a sphere of radius  $R_{\text{vir}}^{\text{cen}}$ , where  $R_{\text{vir}}^{\text{cen}}$  is the virial radius of the central halo. Note that an isolated halo is also identified as a central halo according to this definition. On the other hand, the halos which reside around a more massive neighbor to within the neighbor’s virial radius are identified as “satellite” (sub)halos. The particular definition does not really affect our mock galaxy catalog due to our recipe [Eq. (1)] for populating galaxies. The lower panel of Fig. 1 shows the average number of mock galaxies in central halos, i.e., the halo occupation distribution (HOD), as a function of central halo mass. Note that unlike the standard HOD prescription, the HOD of our mock catalog is not given *a priori* and, rather, is measured from the mocks with the central/satellite split. Nevertheless, the shape of HOD in our mock catalogs looks similar to what can be found in the literature, e.g., [58,59]. There are two regimes: halos around the soft cutoff near  $M_{\text{vir}} = 10^{13} h^{-1} M_{\odot}$  host only one galaxy (i.e., a central galaxy), while massive halos above  $10^{14} h^{-1} M_{\odot}$  receive a

significant contribution from satellite galaxies, displaying a power-law-like form in the HOD.

### C. Measurement of the mock signal and error

Here, we describe the method to measure the power spectra and estimate the data covariance from the mock galaxy catalogs.

The measurement is done based on FFT of the density field. We first assign the mock galaxies in redshift space to  $n_g^3 = 2048^3$  grid points using the cloud in cell (CIC) interpolation scheme. We employ the distant observer approximation in the mapping to the redshift space. We follow Ref. [60] to correct for the aliasing effect [61], by the so-called interlacing method. To do this, we prepare another density grid but with mass assignment done after shifting the galaxy positions by half the grid size along all three Cartesian axes and then corrected for the phase shift by multiplying an appropriate factor to the field in Fourier space. By taking the average of the two density grids, the original and the interlaced, we can get rid of the aliasing effect due to the odd images, which would give the dominant aliasing source to standard cosmological power spectra with decaying amplitude toward higher wave numbers. The effect of the CIC window function will eventually be removed in Eq. (3).

We then reinterpret the wave numbers by taking account of the AP effect. Namely, we rescale the fundamental modes along each of the three axes as

$$\begin{aligned}\tilde{k}_{f,x} &= \tilde{k}_{f,y} = \frac{D_A^{(\text{true})}(z)}{D_A^{(\text{fid})}(z)} k_f, \\ \tilde{k}_{f,z} &= \frac{H^{(\text{fid})}(z)}{H^{(\text{true})}(z)} k_f,\end{aligned}\quad (2)$$

where  $k_f = 2\pi/L$  is the original fundamental mode in the absence of the AP effect. In the above, we take the  $z$  direction in the simulation box as the line-of-sight direction, and the upper scripts, (true) and (fid), indicate that the comoving angular diameter distance  $D_A(z)$  or the Hubble expansion rate  $H(z)$  is calculated assuming the correct, blinded cosmological parameters and fiducial cosmological parameters, respectively. Here, we adopt a flat  $\Lambda$ CDM cosmology with  $\Omega_m^{(\text{fid})} = 0.3$  as the fiducial cosmology, and this information is shared with the two US analysis teams. Note that the  $\Omega_m^{(\text{fid})}$  used to create the mock catalogs should not be confused with the true cosmological parameter  $\Omega_m$ , which was used in the simulations and which was kept from the analyzing teams.

The Japan team then estimates the first three nonzero multipole moments—monopole ( $\ell = 0$ ), quadrupole ( $\ell = 2$ ), and hexadecapole ( $\ell = 4$ )—by taking weighted averages of the squared Fourier modes:

$$\hat{P}_\ell(k_i) = \frac{2\ell + 1}{N_i} \sum_{\tilde{\mathbf{k}} \in \text{bini}} \mathcal{P}_\ell(\mu_{\tilde{\mathbf{k}}}) \hat{P}(\tilde{\mathbf{k}}), \quad (3)$$

with

$$\hat{P}(\tilde{\mathbf{k}}) = \frac{\tilde{V} |\delta_{\tilde{\mathbf{k}}}|^2 - \tilde{P}_{\text{shot}}(\tilde{\mathbf{k}})}{W_{\text{CIC}}^2(\tilde{\mathbf{k}})}, \quad (4)$$

where the distorted volume  $\tilde{V}$  is given by

$$\tilde{V} = \left( \frac{D_A^{(\text{fid})}(z)}{D_A^{(\text{true})}(z)} \right)^2 \frac{H^{(\text{true})}(z)}{H^{(\text{fid})}(z)} L^3. \quad (5)$$

Analogously to Eq. (2), to account for the AP effect, the summation runs over wave vectors  $\tilde{\mathbf{k}}^T = (\tilde{k}_{f,x} i_x, \tilde{k}_{f,y} i_y, \tilde{k}_{f,z} i_z)$  specified by an integer vector  $(i_x, i_y, i_z)$ ,  $\mathcal{P}_\ell$  denotes the  $\ell$ th-order Legendre polynomial,  $\mu_{\tilde{\mathbf{k}}}$  is the cosine between the wave vector  $\tilde{\mathbf{k}}$  and the  $z$  direction, and  $N_i$  stands for the number of Fourier modes contained in the  $i$ th wave-number bin. In the above, we have subtracted the shot noise  $\tilde{P}_{\text{shot}}$  from the measured power spectrum.<sup>2</sup> We evaluate the shot noise, taking into account the interlacing technique for the aliasing correction and the CIC window function. Denoting

$$\tilde{\kappa}_a = \frac{\pi \tilde{k}_a}{2 \tilde{k}_{\text{Ny},a}}, \quad (6)$$

with  $\tilde{k}_{\text{Ny},a} = \tilde{k}_{f,a} n_g / 2$  being the direction-dependent Nyquist frequency ( $a = x, y, \text{ or } z$ ), the resultant expression for the wave-vector-dependent shot-noise contribution is given as

$$\begin{aligned}\tilde{P}_{\text{shot}}(\tilde{\mathbf{k}}) &= \sum_{n_x, n_y, n_z : \text{even}} W_{\text{CIC}}^2(\tilde{\mathbf{k}} + 2\tilde{\mathbf{k}}_{\text{Ny}} \mathbf{n}^T) \frac{\tilde{V}}{N_{\text{gal}}} \\ &= \left[ \prod_{a=x,y,z} C_a(\tilde{\kappa}_a) \right] \frac{\tilde{V}}{N_{\text{gal}}},\end{aligned}\quad (7)$$

with  $W_{\text{CIC}}$  being the CIC window function,

<sup>2</sup>Notice that this contribution is coming from the zero-lag correlator inherent in point processes and is thus exactly  $1/n_g$  for any tracers with a number density  $n_g$ . However, on the modeling side, the stochastic contribution in galaxy spectra uncorrelated with large-scale density fluctuations is sometimes also referred to as the shot noise. In this definition, it is well known that the shot noise, i.e., the level of stochasticity, can deviate from the  $1/n_g$  Poissonian noise. While we omit this in the analyses shown in the main text, the possible impact of treating this as an additional free parameter is discussed in the Appendix B 1.

$$W_{\text{CIC}}(\tilde{\mathbf{k}}) = \prod_{a=x,y,z} \text{sinc}^2 \tilde{\kappa}_a, \quad (8)$$

and the final shot-noise correction factor  $C_a$ , given as the infinite summation over even integers, can be computed analytically as

$$C_a(\tilde{\kappa}_a) = \frac{1}{12} (1 + \cos \tilde{\kappa}_a)^2 (2 + \cos \tilde{\kappa}_a). \quad (9)$$

See Ref. [61] for a similar expression but without the interlacing correction that erases the odd images.

The estimator, Eq. (3), is computed at 100 wave-number bins between the first bin edge taken at zero to the final bin edge at  $1 h \text{Mpc}^{-1}$ , evenly spaced by  $0.01 h \text{Mpc}^{-1}$ . The representative wave number of each bin,  $k_i$  in Eq. (3), is computed as the average of the norm of the wave vectors that actually enter the bin:

$$k_i = \frac{1}{N_i} \sum_{\tilde{\mathbf{k}} \in \text{bini}} |\tilde{\mathbf{k}}|. \quad (10)$$

The pairs of numbers,  $(k_i, \hat{P}_\ell(k_i))$ , are provided to the analysis teams as the mock measurements, and the above way to compute the representative number of each  $k$ -bin is given to the analysis teams. The data files also contain estimates of the covariance matrix, which is obtained assuming Gaussianity [25]:

$$\begin{aligned} \text{Cov}_{ij}^{\ell\ell'} &= \langle (\hat{P}_\ell(k_i) - \langle \hat{P}_\ell(k_i) \rangle) (\hat{P}_{\ell'}(k_j) - \langle \hat{P}_{\ell'}(k_j) \rangle) \rangle, \\ &= \delta_{ij}^K \frac{(2\ell + 1)(2\ell' + 1)}{N_i^2} \\ &\quad \times \sum_{\tilde{\mathbf{k}} \in \text{bin } i} \mathcal{P}_\ell(\mu_{\tilde{\mathbf{k}}}) \mathcal{P}_{\ell'}(\mu_{\tilde{\mathbf{k}}}) [P(\tilde{\mathbf{k}}) + P_{\text{shot}}]^2, \end{aligned} \quad (11)$$

where  $P(\tilde{\mathbf{k}})$  is the expectation value of  $\hat{P}(\tilde{\mathbf{k}})$ . The expression reduces to the real-space formula of Ref. [62] when  $\ell = \ell' = 0$ . In reality, however, we have to make use of a noisy estimate of the power spectrum  $\hat{P}(\tilde{\mathbf{k}})$  for each wave vector  $\tilde{\mathbf{k}}$  instead of  $P(\tilde{\mathbf{k}})$ , and this can impact the estimation of the covariance matrix significantly. Therefore, instead of computing Eq. (11), we first bin the Fourier modes in ten evenly spaced  $|\mu_{\tilde{\mathbf{k}}}|$  bins and take the average of  $\hat{P}(\tilde{\mathbf{k}})$  within each bin to suppress the noise. The binned estimates are then used in Eq. (11), but the summation now runs over bins instead of individual wave vectors to obtain our estimate of the covariance matrix.

The Japan team considers two settings for the covariance matrix. The first is to use the volume and the shot noise, consistent with the mock simulations. In addition, they provide another estimate scaled to the BOSS DR12 catalogs by substituting the number density from the observation and then scaling the number of Fourier modes

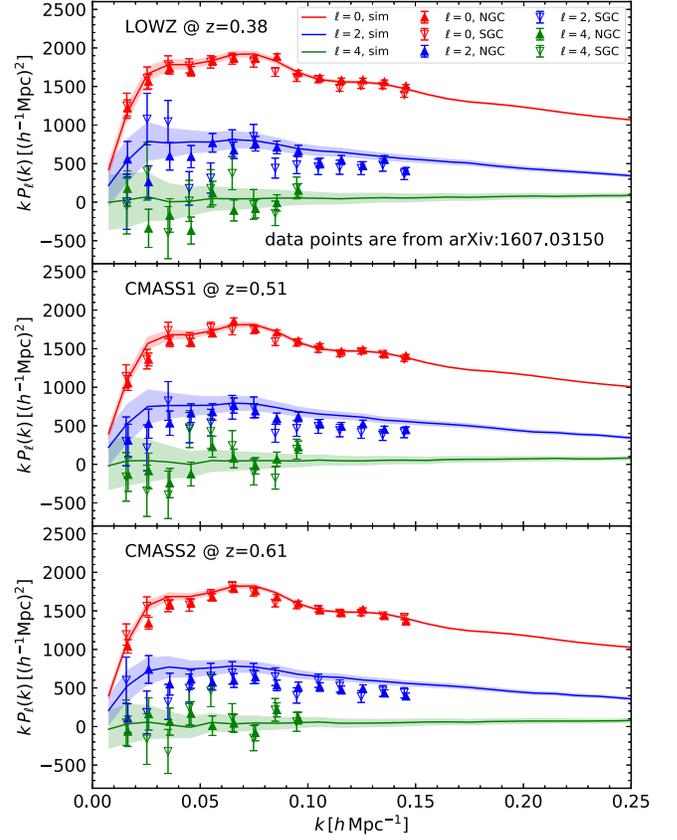


FIG. 2. First three multipole moments (monopole, quadrupole, and hexadecapole) of the power spectrum in redshift space measured from our mock galaxy catalogs at three redshifts (solid lines). The  $1\sigma$  uncertainty intervals, assuming the survey parameter of SDSS data release 12, are shown by the shaded regions. The error bars are taken from Ref. [63] based on SDSS DR12. For these data points, the measurements from the sample in the North Galactic Cap (NGC) and the South Galactic Cap (SGC) are shown separately by different symbols as indicated by the figure legend. Note that the Alcock-Paczynski effect is artificially induced assuming  $\Omega_m = 0.3$  in the redshift-distance conversion. The analysis teams can only access exactly the data vector shown in this figure. The analyses presented in this paper are based on the monopole and the quadrupole moments from the catalog at  $z = 0.61$ .

according to the ratio of the surveyed and the simulated volume. The set of estimates,  $\hat{P}(k_i)$  and  $\text{Cov}_{ij}^{\ell\ell'}$ , with the latter now has only diagonal entries with respect to the subscripts,  $i$  and  $j$ , due to the Gaussian approximation, and are tabulated for each of the ten random realizations and provided on the website. The Japan team leaves the decision to the US teams on how to exactly use these estimates: which survey specification for the estimation of the covariance matrix to adopt, to combine the ten realization and analyze the averaged spectra just once or to analyze each realization one by one, or to further

estimate the non-Gaussian error from the realization-to-realization scatter.

We show in Fig. 2 the average multipole moments of the power spectra at the three redshifts corresponding to LOWZ, CMASS1, and CMASS2. The solid lines show the mock measurements, where the shaded region around each line denotes the  $1\sigma$  error scaled to the SDSS BOSS DR12 survey parameters. The three lines in each panel depict the monopole, quadrupole, and hexadecapole from top to bottom. Also shown by the symbols with error bars are the actual measurements from the BOSS data from Ref. [63]. The measurements from the North and South Galactic Caps are, respectively, plotted by the upward and the downward triangles.

Overall, the mock data follow the observed spectra. The monopole moment especially exhibits an excellent agreement because the model parameters used to distribute the mock galaxies are chosen to match this moment. There is, however, a small mismatch in the quadrupole moment: the observed data show a stronger damping behavior for the higher wave numbers. It is out of the scope of the current investigation to see if this can be alleviated by further tuning the model parameters without spoiling the success in the monopole. This is nontrivial since the cosmological parameters adopted in the mock simulations could depart from the true unknown parameters governing our Universe, or the recipe for populating mock galaxies might not be flexible enough to meet the reality.

#### IV. THEORETICAL TEMPLATE

In this section we describe the implementation of the theoretical model by the two teams participating in the cosmological analysis challenge. The employed methodologies are almost identical to the ones used in the analysis of the actual BOSS data by the same teams [28–30].

Both teams participating in the PT challenge use essentially the same theoretical template. However, there are differences in the implementation of IR resummation, the choice of nuisance parameters, and their priors. In addition, the two teams use absolutely independent pipelines based on different software. This section describes in detail the pipelines used by the two teams and focuses on methodological differences.

##### A. Common basis for the EFT formulation

On general grounds, it is believed that any physical system has a unique and correct description at long wavelengths where the microscopical details of the physical system under consideration can be encoded in just a few coefficients of the terms in the equations of motion. In the context of the long-distance universe, this description is believed to be the EFTofLSS [15,16]. The originality of the EFTofLSS with respect to other preexisting perturbative

methods that were applied in the context of LSS is twofold. First is the presence of suitable terms in the equations of motion that encode the effect of short-distance nonlinearities and galaxies at long distances and that cannot be predicted without detailed knowledge of galaxy physics and therefore are generically fit to observations. Second, the equations of motion in the EFTofLSS have nonlinear terms that are proportional to some parameters. Due to the many phenomena that control the evolution of our Universe, there are several of these parameters, such as the size of the density perturbation or the ratio of a given wavelength with respect to the size of the displacements induced by short-distance modes [18]. For all of these parameters but one, an iterative solution is performed. For the parameter encoding the effect of long-wavelength displacements, a nonlinear solution is performed, which is called IR resummation [18,64–67]. Different incarnations of the EFTofLSS make this expansion more or less manifest. For example, the Lagrangian-space EFTofLSS [68] automatically solves the effect of long displacements nonlinearly, so it is identical to the Eulerian EFTofLSS that we use here after it has been IR resummed [18].

In the EFTofLSS, the description of the clustering of galaxies in redshift space is performed in the following way. First, the dark matter and baryonic fields are described in terms of fluids with a nontrivial stress tensor. Galaxies are biased tracers in the sense that, if  $\delta_g$  is the galaxy overdensity, we have [19]

$$\begin{aligned}\delta_g(x, t) &= \sum_n \int dt' K_n(t, t') \tilde{\mathcal{O}}_n(x_{\Pi}, t') \\ &= \sum_{n,m} b_{n,m}(t) \mathcal{O}_{n,m}(x, t)\end{aligned}\quad (12)$$

where  $\tilde{\mathcal{O}}_n$  are all possible fields—such as, for example, the dark matter density—that, by general relativity, can affect the formation of galaxies. Here,  $K_n(t, t')$  are some kernels that relate how a field at a certain time affects the galaxies at later times, and  $x_{\Pi}$  is the location at time  $t'$  of the fluid element that is at  $x$  at time  $t$ . The last step of the above equation can be performed using the perturbative expression for the matter and baryonic fields. In fact, in perturbation theory the time- and space-dependent parts factorize in a form schematically given by  $\delta(\vec{k}, t) \sim \sum_n f_n(t) \delta^{(n)}(\vec{k})$ , where  $\delta^{(n)}$  is order  $n$  in the expansion parameters. This allows us to define the biases  $b$  as  $b_{n,m}(t) \sim \int dt' K_n(t, t') f_m(t')$ . This provides the first complete parametrization of the bias expansion, though many earlier attempts were made and substantial but partial successes were obtained.

Next, we describe the observed density field in redshift space. This is a combination of the density field in configuration space and density times powers of the velocity field of galaxies, such as  $\rho(\vec{x}, t) v(\vec{x}, t)^i, \rho(\vec{x}, t) v^i(\vec{x}, t) v_j(\vec{x}, t), \dots$ . Again, these short-distance-dependent terms are described

as above as biased tracers of the density and baryonic fields [20].

Because of what we just discussed, the range over which different implementations of the EFTofLSS can differ is extremely limited: one may choose a different basis for the EFT parameters; they may add an incomplete, and therefore different, set of higher-order counterterms to partially include the effect of some higher-order calculation that was not performed; or they may have different implementations or approximations for the IR resummation. We list these differences in detail next.

## B. Group-dependent implementation

Although both teams use the same theoretical model, there are several important methodological differences. Moreover, the two groups have made very different choices in the model implementation and numerical algorithms. This section describes in detail the pipelines used by the two teams.

### 1. East Coast team

The East Coast team used only the monopole and the quadrupole in the analysis. The East Coast team analyzed the challenge data with and without the hexadecapole moment and found identical constraints.<sup>3</sup> Thus, the East Coast team refrained from using the hexadecapole moment in the baseline analysis.

The theoretical model used by the East Coast team for these two multipoles can be written schematically as

$$P_\ell(k) = P_\ell^{\text{tree}}(k) + P_\ell^{\text{loop}}(k) + P_\ell^{\text{ctr}}(k) + P_\ell^{\nabla^4\delta}(k). \quad (13)$$

The tree-level contribution is given by the Kaiser formula [43]. The loop corrections are calculated using the standard one-loop power spectra for dark matter and biased tracers (see, e.g., [23,69,70] and references therein). The bias model consists of the following bias operators [17,19,71],

$$\delta_g(\mathbf{k}) = b_1\delta(\mathbf{k}) + \frac{b_2}{2}\delta^2(\mathbf{k}) + b_{\mathcal{G}_2}\mathcal{G}_2(\mathbf{k}), \quad (14)$$

where the momentum-space representation of the  $\mathcal{G}_2$  operator is given by

<sup>3</sup>On the scales of interest, the hexadecapole signal is dominated by leakage contributions from the monopole and quadrupole. These contributions appear due to discreteness effects, i.e., because the monopole and quadrupole are not exactly orthogonal to the hexadecapole on a finite grid. Even with the gigantic volume of the challenge simulation and the wide binning, the hexadecapole moment is dominated by the systematic leakage from lower multipole moments.

$$\mathcal{G}_2(\mathbf{k}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \left[ \frac{(\mathbf{p} \cdot (\mathbf{k} - \mathbf{p}))^2}{p^2|\mathbf{k} - \mathbf{p}|^2} - 1 \right] \delta(\mathbf{p})\delta(\mathbf{k} - \mathbf{p}). \quad (15)$$

The one-loop power spectrum has one extra bias operator multiplied by an additional parameter  $b_{\Gamma_3}$ . However, this contribution is almost fully degenerate with the counterterms and the  $\mathcal{G}_2$  operator on the scales of interest. Given this strong degeneracy, the East Coast team set  $b_{\Gamma_3} = 0$  in the baseline analysis. Running the Markov chain Monte Carlo (MCMC) chains with and without  $b_{\Gamma_3}$ , it was checked that this choice does not affect constraints on cosmological parameters.

The standard one-loop counterterms for the monopole and the quadrupole are [20]

$$P_0^{\text{ctr}}(k) = -2c_0^2k^2P_{11}(k), \quad P_2^{\text{ctr}}(k) = -\frac{4f}{3}c_2^2k^2P_{11}(k), \quad (16)$$

where  $f = d \ln D_+ / d \ln a$  is the logarithmic growth rate,  $D_+$  denotes the linear growth factor, and  $P_{11}(k)$  is the linear power spectrum. The purpose of these counterterms is to fix the UV dependence of the loops and to partly take into account the effects of the fingers of God [42]. The East Coast team also added an extra  $k^4$  term shared between the multipoles,

$$P_{\nabla^4\delta}^{\text{ctr}}(k, \mu) = -c(\mu k f)^4(b_1 + f\mu)^2P_{11}(k). \quad (17)$$

This new counterterm takes into account the next-to-leading order of the fingers of God. Note that on general grounds, one also expects the presence of the stochastic contribution of the form [20,72]

$$P_{\text{RSD,stoch}} = -c_e k^2 \mu^2. \quad (18)$$

This contribution is very degenerate with the counterterm (17) on the scales of interest for the analysis, and it was not included in the model by the East Coast team.

The East Coast team implemented IR resummation and the Alcock-Paczynski effect as explained in detail in Refs. [73,74]. Importantly, the East Coast team used the IR-resummation algorithm based on the wiggly smooth decomposition directly in Fourier space [64,67,75], which allowed for a significant boost of computational speed. This scheme is efficient and numerically stable. Moreover, it is based on solid systematic parametric expansion that guarantees that the error is under control at every order of IR resummation. It was explicitly checked that the residuals introduced by our procedure are much smaller than the two-loop contributions which are not included in the model, in full agreement with theoretical expectations [67,75]. The labels that indicate IR resummation and the AP effect were omitted in all equations in this section to avoid clutter. However, the reader should keep in mind that they are always included in the model.

The total number of nuisance parameters used in the blinded analysis of the East Coast team is six: three counterterms ( $c_0^2$ ,  $c_2^2$ ,  $c$ ) and three bias parameters ( $b_1$ ,  $b_2$ ,  $b_{G_2}$ ). Since the shot-noise contribution has been subtracted from the measured spectra, the corresponding parameter is not fitted, in contrast to Ref. [74]. As far as the cosmological parameters are concerned, the basis that was used consists of the dimensionless Hubble constant  $h$  ( $H_0 = h \cdot 100$  km/s/Mpc), the physical matter density  $\omega_m$ , and the normalization  $A^{1/2}$  defined with respect to the best-fit Planck value for the base  $\Lambda$ CDM cosmology,

$$A^{1/2} \equiv \left( \frac{A_s}{A_{s,\text{Planck}}} \right)^{1/2},$$

where  $A_{s,\text{Planck}} = 2.0989 \times 10^{-9}$ . (19)

All varied cosmological and nuisance parameters were assigned flat priors without boundaries, i.e.,  $(-\infty, \infty)$ .

The evaluation of perturbation theory integrals was performed using the FFTLog method of [76] implemented as a module in the CLASS Boltzmann solver [77,78]. Using the IR resummation based on wiggly smooth decomposition, a single evaluation of a theoretical model is of the order  $\mathcal{O}(1)$  sec for high precision settings. This allows for a new evaluation of the nonlinear power spectra at every step of the MCMC chain, which is what is done in the East Coast team analysis. The MCMC analysis was performed using the Montepython v3.0 [79,80] sampler interfaced with the modified version of the CLASS code. The nuisance parameters were sampled in the “fast mode” [81] at a negligible computational cost.

Since the  $k$ -binning of the challenge spectra is very wide ( $\Delta k = 0.01 h \text{ Mpc}^{-1}$ ) compared to the fundamental mode of the box, the theoretical predictions have to be properly averaged over each bin. The boundaries of the bins were estimated using the simulation volume, which was known to both teams. The East Coast team checked that the estimated boundaries allow one to accurately reproduce the provided weighted means of the  $k$ -bins and found that averaging the theory over the bin versus evaluating it in the mean can induce roughly  $\mathcal{O}(0.5)\sigma$  shifts in cosmological parameters.

## 2. West Coast team

The implementation of the West Coast team is the result of a long journey where each of the ingredients of the EFTofLSS necessary to apply it to data was subsequently developed one by one, tested on simulations, and shown to be successful. Though not all those results are directly used in the analysis, the West Coast team would have never applied the model to the data without those intermediate successes. We therefore find it nice to add, in each instance where the EFTofLSS is applied to data, the following

footnote where we acknowledge at least a fraction of those important developments.<sup>4</sup>

The model for the West Coast team and the analysis techniques are the same as the ones used in [28,30], to which we refer for details. The one-loop redshift-space galaxy power spectrum reads

$$\begin{aligned} P_g(k, \mu) = & Z_1(\mu)^2 P_{11}(k) \\ & + 2 \int \frac{d^3 q}{(2\pi)^3} Z_2(\mathbf{q}, \mathbf{k} - \mathbf{q}, \mu)^2 P_{11}(|\mathbf{k} - \mathbf{q}|) P_{11}(q) \\ & + 6 Z_1(\mu) P_{11}(k) \int \frac{d^3 q}{(2\pi)^3} Z_3(\mathbf{q}, -\mathbf{q}, \mathbf{k}, \mu) P_{11}(q) \\ & + 2 Z_1(\mu) P_{11}(k) \left( c_{\text{ct}} \frac{k^2}{k_M^2} + c_{r,1} \mu^2 \frac{k^2}{k_M^2} + c_{r,2} \mu^4 \frac{k^2}{k_M^2} \right) \\ & + \frac{1}{\bar{n}_g} \left( c_{\epsilon,1} + c_{\epsilon,2} \frac{k^2}{k_M^2} + c_{\epsilon,3} f \mu^2 \frac{k^2}{k_M^2} \right). \end{aligned} \quad (20)$$

Here,  $k_M^{-1}$  controls the bias derivative expansion, and we set it to be  $\simeq k_{\text{NL}}^{-1}$ , which is the scale controlling the expansion of the dark matter derivative expansion. We set  $k_M = 0.7 h \text{ Mpc}^{-1}$ , and  $\bar{n}_g$  is the mean galaxy density.

In the next-to-last line of Eq. (20), the term in  $c_{\text{ct}}$  represents a linear combination of a higher derivative bias [19] that appears in Eq. (12) and the speed of sound of dark matter [15,16]:  $\delta(\vec{k}, t) \supset k^2 \delta_{\text{lin}}(\vec{k}, t)$ . The terms in  $c_{r,1}$  and  $c_{r,2}$  represent the redshift-space counterterms [20]:  $\delta_{\text{redshift}}(\vec{k}, t) \supset k^2 \mu^2 \delta(k, t)$ ,  $k^2 \mu^4 \delta(k, t)$ . In the last line of

<sup>4</sup>The initial formulation of the EFTofLSS was performed in Eulerian space in [15,16] and then extended to Lagrangian space in [68]. The dark matter power spectrum was computed at one-, two-, and three-loop orders in [16,18,82–90]. Some additional theoretical developments of the EFTofLSS that accompanied these calculations were a careful understanding of renormalization [16,91,92] (including rather subtle aspects such as lattice running [16] and a better understanding of the velocity field [82,93]), of the several ways for extracting the value of the counterterms from simulations [16,94], and of the nonlocality in time of the EFTofLSS [19,82,84]. These theoretical explorations also include an instructive study in  $1+1$  dimensions [94]. In order to correctly describe the BAO peak, an IR resummation of the long displacement fields had to be performed. This led to the so-called IR-resummed EFTofLSS [18,64–67]. A method to account for baryonic effects was presented in [21]. The dark matter bispectrum was computed at one loop in [95,96], the one-loop trispectrum in [97], and the displacement field in [98]. The lensing power spectrum was computed at two loops in [99]. Biased tracers, such as halos and galaxies, were studied in the context of the EFTofLSS in [19,71,72,100–102] (see also [14]), and the halo and matter power spectra and bispectra (including all cross-correlations) in [19,100]. Redshift-space distortions were developed in [20,22,72]. Clustering dark energy was included in the formalism in [89,103–105], primordial non-Gaussianities in [22,100,106–109], and neutrinos in [110,111]. Faster evaluation schemes for evaluation for some of the loop integrals were developed in [76].

Eq. (20), we have the stochastic counterterms:  $c_{e,1}$  and  $c_{e,2}$  originate from Taylor expansion of Eq. (12) [19], while  $c_{e,3}$  originates from the redshift-space expressions [20].

The redshift-space galaxy density kernels  $Z_1$ ,  $Z_2$ , and  $Z_3$  are given in Appendix A. These kernels depend on the bias coefficients that we define as explained below Eq. (12). By choosing only the linearly independent ones, this gives rise to the so-called base of descendants. While up to cubic order this base is equivalent to more standard bases, already at quartic perturbative order new terms appear.

The IR resummation is performed in a numerically efficient way using the original method for configuration and redshift space developed in [18,65,66], where all the errors are parametrically controlled by the perturbative order of the calculation (i.e., no uncontrolled approximations are present).<sup>5</sup>

We define the following combination of parameters:  $c_2 = (b_2 + b_4)/\sqrt{2}$ ,  $c_4 = (b_2 - b_4)/\sqrt{2}$ ,  $c_{e,\text{mono}} = c_{e,2} + fc_{e,3}/3$ , and  $c_{e,\text{quad}} = 2fc_{e,3}/3$ . As we analyze only the monopole and the quadrupole, we set  $c_{r,2} = 0$  since the two redshift-space counterterms are degenerate in this case, but we allow a larger prior on  $c_{r,1}$  to absorb the contribution of  $c_{r,2}$  in the quadrupole. Additionally, since the shot noise is known and has been subtracted from the data, we set  $c_{e,1} = 0$ . This leaves us with the set  $(b_1, c_2, b_3, c_4, c_{ct}, c_{r,1}, c_{e,\text{mono}}, c_{e,\text{quad}})$  of eight parameters. The PT challenge data are precise enough to determine all EFT parameters with no priors. However, we impose the following priors motivated by the fact that all EFT parameters are expected to be  $\mathcal{O}(1)$ <sup>6</sup>:

$$\begin{aligned} b_1 &\in [0, 4]_{\text{flat}}, & c_2 &\in [-4, 4]_{\text{flat}}, & b_3 &\in 10_{\text{gauss}}, \\ c_4 &\in 2_{\text{gauss}}, & c_{ct} &\in 4_{\text{gauss}}, & c_{r,1} &\in 8_{\text{gauss}}, \\ c_{e,\text{mono}} &\in 2_{\text{gauss}}, & c_{e,\text{quad}} &\in 4_{\text{gauss}}. \end{aligned} \quad (21)$$

As is evident from Eqs. (20) and (A1), some EFT parameters appear linearly in the model power spectrum and therefore appear quadratically in the likelihood. If we

<sup>5</sup>Especially within the observational community, a nonlinear treatment of the BAO based on the decomposition of the wiggle and smooth parts of the power spectrum has been popular for a long time (see for example [112]). However, this team does not find this decomposition to be under parametric control (i.e., there is no small parameter controlling its correctness). It is possible to go from the original IR resummation to the simplified ones based on the decomposition by performing a series of approximations (see the Appendix of [66]). Of course, this does not mean that the errors which are introduced are large or significant, as can be *a posteriori* checked in numerical simulations.

<sup>6</sup>Notice that the consistency of the EFTofLSS is based on a power-counting argument that assumes that the subsequent terms of the perturbative expansion are much smaller than the ones that are kept. In order for this to be the case, it is essential that the physical nuisance parameters are kept,  $\mathcal{O}(1)$ , once the relevant physical scales have been factorized.

are not interested in the actual value of these parameters, as is our case, we can marginalize over these parameters analytically, obtaining a marginalized likelihood that is a function of only three parameters:  $b_1$ ,  $c_2$ , and  $c_4$ .

Given that the  $k$ -bins ( $\Delta k = 0.01 h/\text{Mpc}$ ) contain many fundamental modes, the West Coast team averages the predictions of the model over each bin. As a check, the team verifies that the provided effective  $k$  of the bin is correctly reproduced.

In terms of the cosmological parameters, the West Coast team parametrizes their analysis in terms of the dimensionless Hubble constant  $h$  ( $H_0 = h \cdot 100 \text{ km/s/Mpc}$ ), the present-day matter density fraction  $\Omega_m$ , and the normalization of the power spectrum  $A_s$ . The evaluation of the perturbation theory integrals is performed either by direct numerical integration or by the FFTLog method of [76], obtaining the same result.

## V. RESULTS OF BLINDED ANALYSIS

In this section we display the results obtained by the two teams. The input values of the cosmological parameters were unblinded after each team submitted its results for consensus data cuts. We present these results in the original form prepared by either team independently. Both teams have chosen to analyze the mean power spectrum (at  $z = 0.61$ ) over ten realizations, with the covariance estimated from the inverse sum of covariances for ten single boxes,

$$\bar{C} = \left( \sum_i C_i^{-1} \right)^{-1}, \quad \bar{P} = \bar{C} * \sum_i C_i^{-1} P_i, \quad (22)$$

where  $P_i, C_i$  are the power spectrum and covariance of the  $i$ th box and  $\bar{P}, \bar{C}$  are the final mean and covariance that were analyzed.

This procedure ensures that the analysis is approximately equivalent to fitting the spectrum from a single simulation box of  $566 (h^{-1} \text{ Gpc})^3$  volume. We stress that the obtained statistical errors on cosmological parameters correspond to the total volume of ten simulation boxes, i.e.,  $566 (h^{-1} \text{ Gpc})^3$ .

### A. East Coast team

Although the East Coast team submitted its baseline results for the average over ten challenge boxes at  $z = 0.61$ , they also analyzed the data for other redshifts and found consistent results across all challenge spectra. Prior to unblinding, the East Coast team submitted results for eight different evenly spaced values of  $k_{\text{max}}$  in the range  $(0.08 - 0.2) h \text{ Mpc}^{-1}$ .

The marginalized posteriors for the three cosmological parameters are shown in Fig. 3 for several choices of  $k_{\text{max}}$ . Between  $k_{\text{max}} = 0.08 h \text{ Mpc}^{-1}$  and  $k_{\text{max}} = 0.14 h \text{ Mpc}^{-1}$  the different contours are compatible within  $1\sigma$ . When pushing to higher values of  $k_{\text{max}}$ , the shifts in the central

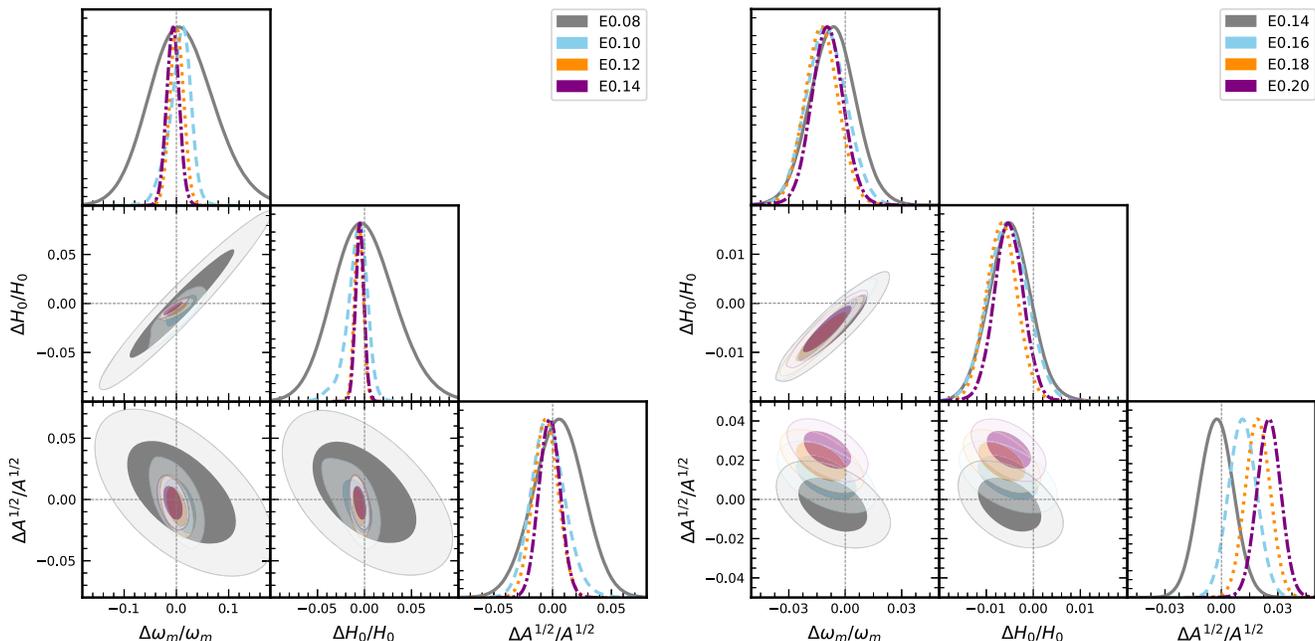


FIG. 3. Marginalized posteriors for the three varied cosmological parameters as a function of  $k_{\max}$  (quoted in  $h\text{Mpc}^{-1}$  in the figure legend) obtained by the East Coast team. Dashed lines mark the input parameters, which were revealed once the team submitted its final result.

values of the posterior distributions become significant. Note that for  $k_{\max} > 0.14 h\text{Mpc}^{-1}$  the contours of  $h$  and  $\omega_m$  remain consistent even though the other parameter exhibits clear shifts. The East Coast team quoted its final results for a conservative choice of  $k_{\max} = 0.12 h\text{Mpc}^{-1}$  because this is the scale up to which the team believed the theoretical modeling is sufficiently accurate given subpercent statistical error bars and the size of neglected nonlinear corrections (see Fig. 4, in which we display an estimate of the two-loop correction from [113]). The 1D marginalized limits for the cosmological parameters and the linear bias  $b_1$  are given in Table I. After the true parameters were unblinded, the values obtained by the East Coast team were replaced by relative differences. For convenience, the values of  $\sigma_8$ ,  $\Omega_m$ , and  $\ln(10^{10}A_s)$  derived from the East Coast team MCMC chains are also quoted. As we have seen after unblinding, the true values of  $\omega_m$  and  $h$  reside within  $2\sigma$  posterior regions even at  $k_{\max} = 0.2 h\text{Mpc}^{-1}$ , while the clustering amplitude measurement is consistent up to  $k_{\max} = 0.14 h\text{Mpc}^{-1}$ . Importantly, the team also inferred a correct value of the linear bias<sup>7</sup> coefficient  $b_1$ .

Figure 4 shows the comparison of the best-fit model at  $k_{\max} = 0.12 h\text{Mpc}^{-1}$  to the data and the residuals. The quality of the fit is quite good,  $\chi^2/\text{dof} = 12/(24 - 9)$ . It is

<sup>7</sup>The true value of the linear bias was estimated as follows. The Japan team measured the real-space matter-matter autospectra along with the galaxy-matter cross spectrum. Then, we took the ratio  $b_1 = P_{gm}/P_{mm}$  evaluated in the very first  $k$ -bin averaged over the ten realizations as an estimate of the bias parameter.

consistent with the hypothesis that the data follow the  $\chi^2$  distribution with 15 degrees of freedom. The lower panel of Fig. 4 displays a breakdown of different contributions to the best-fit model. The linear theory contribution dominates on all scales, which is consistent with the applicability of perturbation theory. Towards  $k_{\max} = 0.12 h\text{Mpc}^{-1}$  the loop corrections (including the  $k^2$  counterterms) become progressively important. Note that the one-loop corrections are detectable already on very large scales,  $\sim 0.02 h\text{Mpc}^{-1}$ . The  $k^4$  counterterm is important only for the quadrupole around  $k_{\max} = 0.12 h\text{Mpc}^{-1}$ , where it dominates over the other loop corrections.

## B. West Coast team

As specified before, the West Coast team analyzed the mean over the ten boxes in the high redshift bin at  $z = 0.61$ , using the covariance on the mean. Originally, for the purpose of parameter estimation, the team presented the results up to  $k_{\max} = 0.12 h\text{Mpc}^{-1}$  since this is the  $k_{\max}$  at which the team predicted the estimates to still be unbiased. The marginalized posteriors for the cosmological parameters are shown in Fig. 5, and the best fits and means are listed in Table II. When the true results are revealed, it is found that  $A_s$  and  $H_0$  lie within the  $1\sigma$  region of the estimates of the West Coast team, and  $\Omega_m$  within the  $1.5\sigma$  region. Note that  $b_1$  is also correctly reproduced within the  $1\sigma$  interval. Additionally, one can see that the pre-unblinding results at  $k_{\max} = 0.14 h\text{Mpc}^{-1}$ , which, however, was not the  $k_{\max}$  at which the team anticipated to be most accurate, are even closer to the true values.

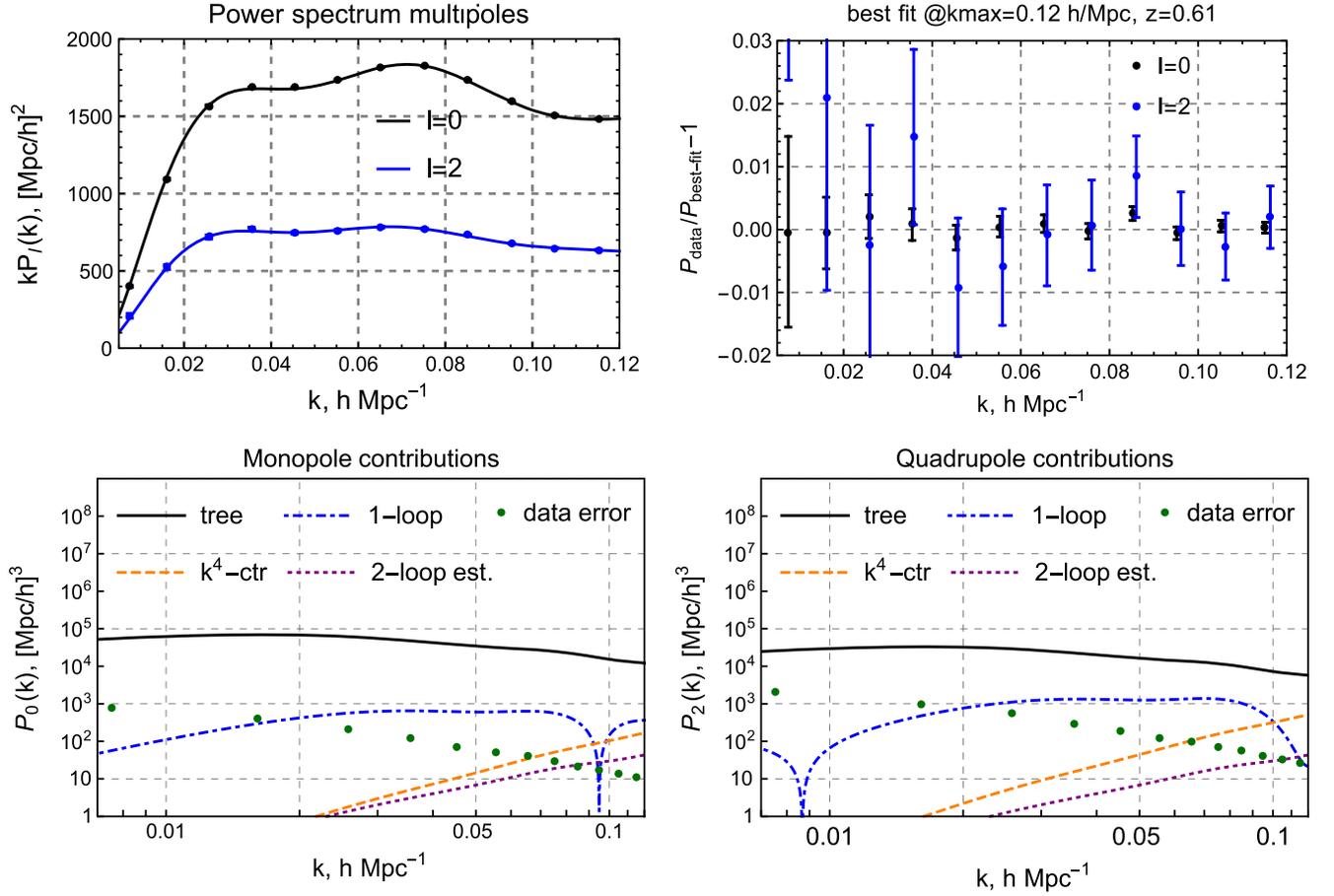


FIG. 4. Upper panel: comparison of the data for the monopole and the quadrupole (with error bars, albeit they are barely visible) with the best-fit model (left panel) obtained by the East Coast team. The residuals for the monopole and the quadrupole for the best-fit model with  $\chi^2/\text{dof} = 12/(24 - 9)$  are shown in the right panel. Note that the quadrupole data points are slightly shifted for better visibility. Lower panel: different contributions to the monopole (left panel) and quadrupole (right panel) power spectra. The data errors and the two-loop estimate are also displayed. We plot the absolute values; some terms are negative.

In Fig. 6 the team shows that the data are well fitted by the theoretical model with the best-fit parameters, with  $-2 \log \mathcal{L}/\text{dof} = 16/(24 - 6)$ , corresponding to a very good  $p$ -value.<sup>8</sup> In the lower panel, different contributions to the best-fit power spectra are shown to check the self-consistency of the perturbative expansion. It is apparent that the one-loop term is safely less than 10% of the linear one at all  $k$ 's. In addition to the one-loop term, an estimate of the two-loop contribution, i.e.,  $P_{1\text{-loop}}^2/P_{\text{lin}}$ , is shown: clearly, at least for the quadrupole, this estimate is of the order of the error on the data at the highest  $k$ . This is an additional indication that, for roughly  $k_{\text{max}} \gtrsim 0.12 - 0.14 h \text{ Mpc}^{-1}$ , the one-loop model will not be an accurate description of the data, and parameter estimation will suffer from theory systematics.

After unblinding, the West Coast team submitted additional results at  $k_{\text{max}} = 0.14, 0.16, 0.18, 0.20 h \text{ Mpc}^{-1}$

because it was subsequently decided that it was interesting to explore the  $k_{\text{max}}$  dependence of the theory-systematic error. In fact, though this was already analyzed by the team in both their original papers [28,30], the challenge simulation is different and its volume is larger. At the higher

TABLE I. Baseline results obtained by the East Coast team for  $k_{\text{max}} = 0.12 h \text{ Mpc}^{-1}$  at  $z = 0.61$ . Only the cosmological parameters and  $b_1$  are shown. Note that  $\Omega_m$ ,  $\ln(10^{10} A_s)$ , and  $\sigma_8$  in the lower disjoint table show the results for the derived parameters.

$k_{\text{max}} = 0.12 h \text{ Mpc}^{-1}$	Best fit	Mean $\pm 1\sigma$
$\Delta A^{1/2}/A^{1/2} \times 10^2$	-0.15	$-0.16 \pm 1.0$
$\Delta h/h \times 10^2$	-0.55	$-0.59 \pm 0.46$
$\Delta \omega_m/\omega_m \times 10^2$	0.2	$0.15 \pm 1.4$
$\Delta b_1/b_1 \times 10^2$	0.20	$0.22 \pm 1.2$
$\Delta \Omega_m/\Omega_m \times 10^2$	1.3	$1.2 \pm 0.9$
$\Delta \ln(10^{10} A_s)/\ln(10^{10} A_s) \times 10^2$	-0.098	$-0.11 \pm 0.69$
$\Delta \sigma_8/\sigma_8 \times 10^2$	-0.094	$-0.022 \pm 0.92$

<sup>8</sup>Notice that the likelihood of this team is not Gaussian.

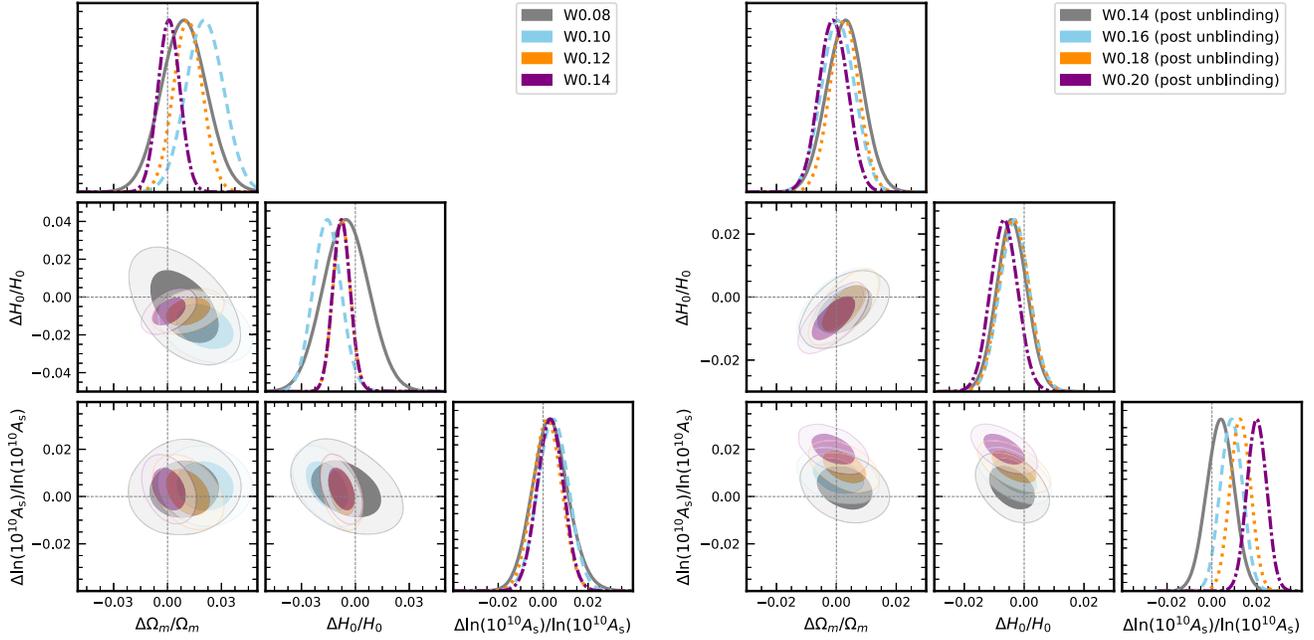


FIG. 5. Marginalized posteriors for the three varied cosmological parameters as a function of  $k_{\max}$  (quoted in  $h\text{Mpc}^{-1}$  in the figure legend) obtained by the West Coast team. Dashed lines mark the input parameters which were revealed once the team submitted its final result, similarly to Fig. 3.

$k_{\max}$ 's, the team performs the (analytical) marginalization over the additional  $c_{e,\text{mono}}$  parameter, with a Gaussian prior with  $\sigma_{c_{e,\text{mono}}} = 2$ . The effect of adding this parameter is completely negligible at low  $k_{\max}$ : in fact, the team chose to safely set it to zero for the original chains. Indeed, one can check that the results are unchanged at low  $k_{\max}$  when adding this parameter. However, because of the small error bars of the simulation data, at higher  $k_{\max}$  this parameter has to be added to the model.

The trend as a function of  $k_{\max}$  is apparent from Fig. 5. Note that  $\Omega_m$  and  $H_0$  are well recovered up to  $k_{\max} = 0.18 h\text{Mpc}^{-1}$ , approximately within the  $1\sigma$  region; the estimate of clustering amplitude  $A_s$  starts to deviate significantly from the true value after  $k_{\max} \gtrsim 0.14 h\text{Mpc}^{-1}$ .

### C. Comparison of the two analyses

So far we have presented the analyses done by two teams. We now compare the two and discuss how different

TABLE II. Similar to Table I, but with the results obtained by the West Coast team for  $k_{\max} = 0.12 h\text{Mpc}^{-1}$  at  $z = 0.61$ . Only cosmological parameters and  $b_1$  are shown.

Parameters	Best fit	Mean $\pm \sigma$
$\Delta\Omega_m/\Omega_m \times 10^2$	1.3	$1.2^{+0.8}_{-0.8}$
$\Delta h/h \times 10^2$	-0.7	$-0.6^{+0.6}_{-0.6}$
$\Delta \ln(10^{10}A_s)/\ln(10^{10}A_s) \times 10^2$	0.1	$0.1^{+0.7}_{-0.7}$
$\Delta b_1/b_1 \times 10^2$	0.8	$0.7^{+1.0}_{-1.1}$

model assumptions lead to different cosmological-parameter constraints.

First, since the two teams employ different sets of cosmological parameters as the varied parameters, a direct comparison between Figs. 3 and 5 is not very clear. Here, instead, we focus on the parameter space  $(\Omega_m, H_0, A_s)$  to find the constraints. We first show in Fig. 7 the one-dimensional marginalized error on these parameters as a function of the maximum wave number  $k_{\max}$  used in the analysis. The  $1\sigma$  credible intervals by the East (West) Coast team are shown by the upward (downward) triangles with error bars. Also shown by the shaded areas are the same intervals but scaled for the SDSS BOSS DR12 according to the ratio of the simulated and the observed volume.<sup>9</sup>

Overall, the ground truth values of the three cosmological parameters stay within or slightly off from the  $1\sigma$  interval up to  $k_{\max} = 0.14 h\text{Mpc}^{-1}$ . The inferred primordial scalar amplitude  $A_s$ , in particular, is always within the interval up to this  $k_{\max}$  from both teams. On the one hand,  $A_s$  starts to deviate from the ground truth in a systematic way with statistical significance above this  $k_{\max}$ . This is consistent with the expectation that two-loop corrections become important at these scales. On the other hand,  $H_0$  and  $\Omega_m$  stay roughly within  $1\sigma$  from the true value all the way up to  $k_{\max} = 0.2 h\text{Mpc}^{-1}$ . However, if one focuses on the shaded regions corresponding to the statistical error from the actual BOSS survey, the ground truth values

<sup>9</sup>We adopt the total volume of SDSS BOSS DR12,  $5.7(h^{-1}\text{Gpc})^3$ , instead of that of CMASS2.

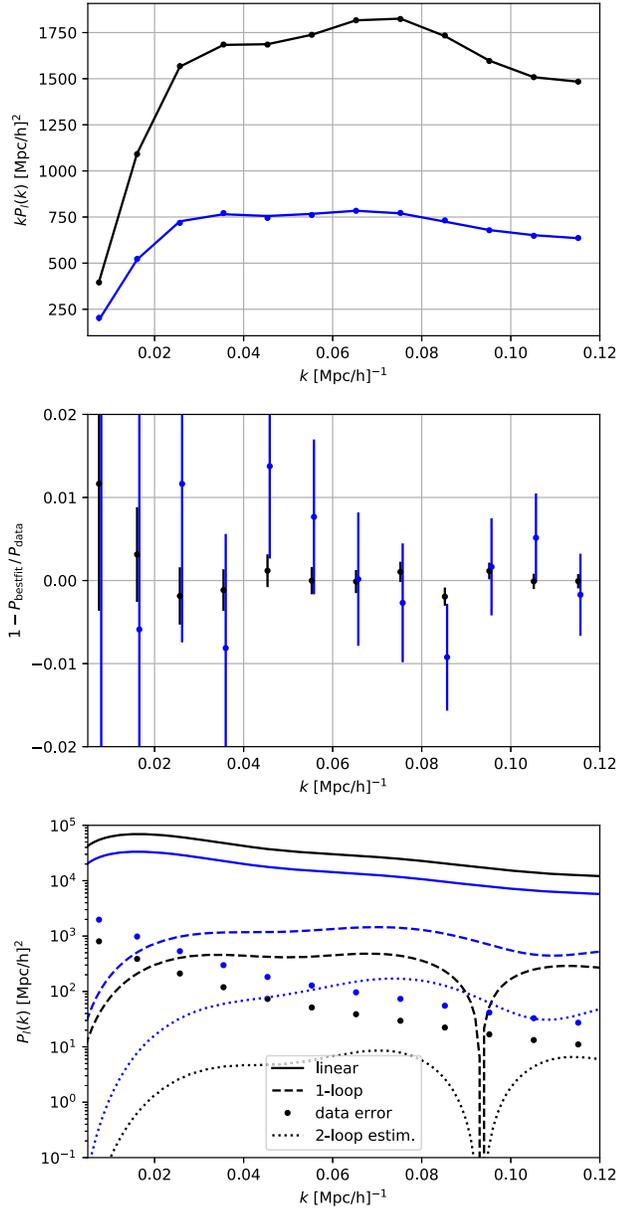


FIG. 6. Upper panel: comparison of the data for the monopole (black) and the quadrupole (blue) with the best-fit model obtained by the West Coast team. Middle panel: residuals for the monopole and the quadrupole for the best-fit model with the partially marginalized likelihood giving  $-2 \log \mathcal{L}/\text{dof} = 16/(24 - 6)$  for  $k_{\text{max}} = 0.12 \ h \text{ Mpc}^{-1}$ . Lower panel: different contributions to the monopole and quadrupole power spectra. We only plot the absolute values; some terms are negative.

are always well within the  $1\sigma$  interval, which justifies the  $k_{\text{max}}$  choice of the analyses from the same teams in Refs. [28–30].

While the size of the error bars shrinks towards higher  $k_{\text{max}}$ , the gain is small after  $k_{\text{max}} \gtrsim 0.14 \ h \text{ Mpc}^{-1}$ . This could be caused by the combination of two effects. First, the relative contribution of the shot noise in the data covariance becomes

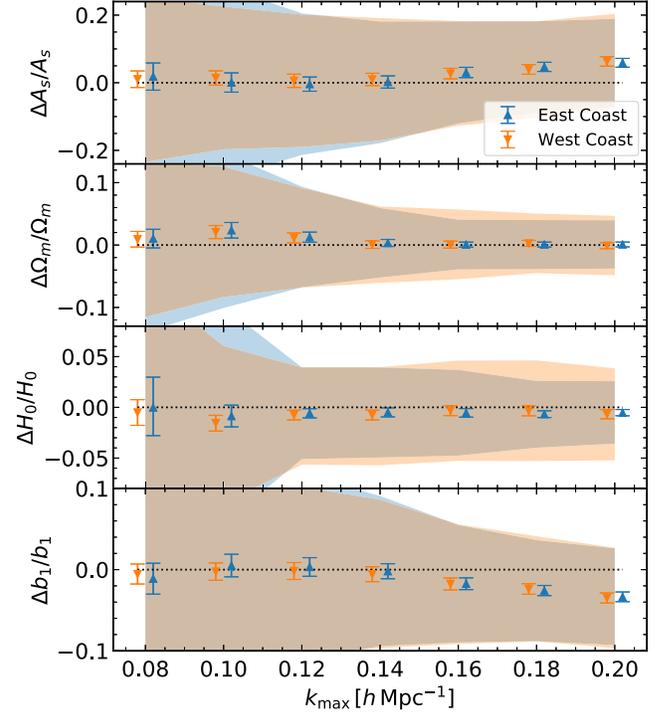


FIG. 7. One-dimensional marginalized posterior distributions for the three main cosmological parameters as well as the linear bias parameter as a function of the maximum wave number  $k_{\text{max}}$  considered in the analysis. The 68% credible intervals derived by the East and West Coast teams are shown, respectively, by the blue and red error bars, with the mean marked by the upward and downward triangles. Also plotted by the shaded regions are those scaled to the volume of SDSS DR12. The error bars are slightly shifted horizontally to avoid a heavy overlap.

important. Second, the EFT parameters controlling the nonlinear corrections become important in such a way that the additional information coming from small-scale modes mainly determines these parameters rather than the cosmological parameters. If one looks into the trend in the error bars more closely, the results from the two teams are clearly different, especially when  $k_{\text{max}} \lesssim 0.1 \ h \text{ Mpc}^{-1}$ , up to a factor  $\sim 2$  smaller by the West Coast team. This difference is driven by the prior treatment. The East Coast team had no priors on the chosen set of nuisance parameters, whereas the West Coast team always kept the nuisance parameters within physically motivated bounds. Thus, the observed difference of the results between the two teams implies that on scales larger than  $0.1 \ h \text{ Mpc}^{-1}$ , the data are not good enough to break degeneracies between the cosmological and nuisance parameters. These degeneracies are broken at larger wave numbers, where the results of the two teams agree regardless of the nuisance parameters' priors.

Let us briefly discuss some cosmological implications of our blinded analysis. The cosmological information probed by redshift galaxy surveys can be crudely divided into four different categories:

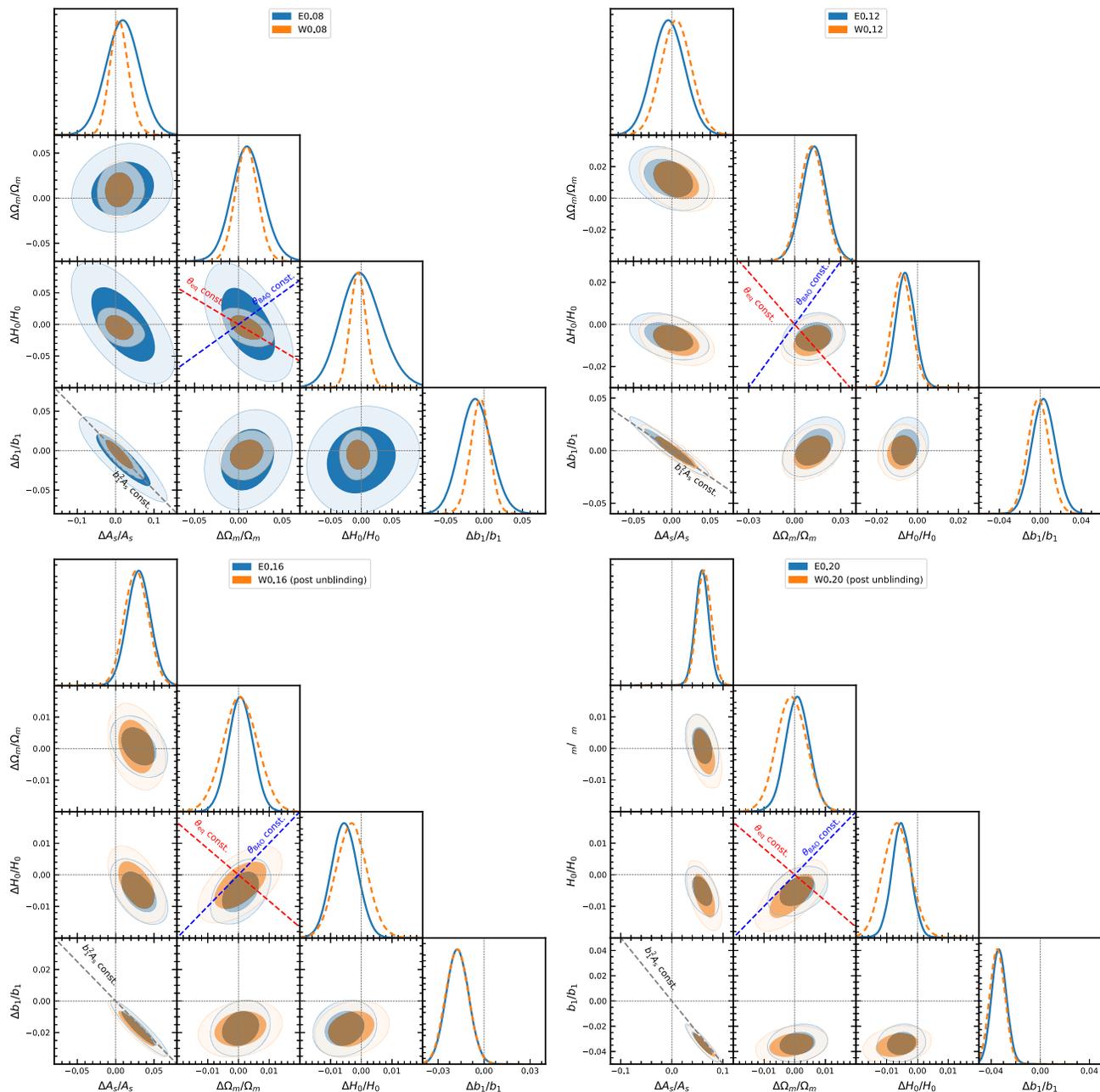


FIG. 8. Two-dimensional marginal posterior distributions for the three main cosmological parameters and the linear bias parameter. The 68% and 95% credible intervals derived by the East and West Coast teams are shown, respectively, by the cyan and orange contours. The corresponding one-dimensional marginal distributions are shown in the diagonal panels by the solid and dashed lines. The maximum wave numbers included in this analysis are  $k_{\max} = 0.08$  (upper left), 0.12 (upper right), 0.16 (lower left), and 0.2  $h \text{ Mpc}^{-1}$  (lower right). Three degeneracy directions for some parameter combinations are also displayed in the contour panels by the thick dashed lines (see text for more details).

(i) Shape information. The shape of the galaxy power spectrum is controlled by the physical matter density  $\omega_m$ . This parameter is measured from the data regardless of the choice of rulers such as  $H_0$ . Note that  $\omega_m$  is extracted from the features of the power spectrum, such as the form of the BAO peaks, the baryonic suppression, the turnover, and the overall slope.

(ii) Distance information, mainly encoded through the volume-average distance<sup>10</sup>  $D_V(z)$ . This parameter

<sup>10</sup>It is defined as  $D_V(z) = (z(1+z)^2 D_A^2(z)/H(z))^{1/3}$ , where  $D_A(z) = \frac{1}{1+z} \int_0^z \frac{dz'}{H(z')}$  and  $H^2(z) = H_0^2(\Omega_m(1+z)^3 + 1 - \Omega_m)$  in flat  $\Lambda$ CDM.

essentially controls the freedom to shift the power spectrum along the  $k$  axis. In the flat  $\Lambda$ CDM framework this distance depends only on two cosmological parameters,  $\omega_m$  and  $H_0$ . Since  $\omega_m$  is measured from the shape, the constraint on  $D_V$  translates directly into a constraint on  $H_0$ . Note that  $\Omega_m$  in this picture can be seen as a parameter derived from a combination of the shape and distance parameters.

- (iii) Redshift-space distortions. Observing galaxies in redshift space allows one to measure unbiased rms velocity fluctuations  $f\sigma_8(z) = f(z)D_+(z)\sigma_8$ . In  $\Lambda$ CDM  $D_+$  and  $f$  depend only on  $\Omega_m$ , which is constrained from the shape and the distance. Thus, the RSD measurements directly constrain  $A_s$ .
- (iv) Alcock-Paczynski geometric distance information. The AP effect allows one to measure the combination  $H(z)D_A(z)$ . However, in  $\Lambda$ CDM this combination is a slow function of cosmological parameters at small redshifts. Thus, it does not contribute significantly to the overall constraints on  $\Omega_m$ ; see Ref. [29] for more details.

We can see in Fig. 8 that, indeed, our results are fully in line with these theoretical expectations. First, let us focus on the two-dimensional posterior in the  $(\Omega_m-H_0)$  plane. We observe that the change in the degeneracy direction rotates by increasing the maximum wave number. When  $k_{\max} = 0.08 h \text{Mpc}^{-1}$ ,  $\Omega_m$  and  $H_0$  are negatively correlated. On the other hand, the correlation becomes positive for  $k_{\max} = 0.16$  and  $0.2 h \text{Mpc}^{-1}$ . We can interpret this as the outcome of the change in the relative importance of the BAO feature. Although the first BAO peak is already included at  $k_{\max} = 0.08 h \text{Mpc}^{-1}$ , the dominant constraint comes from the overall shape information, e.g., the matter-radiation equality scale [ $\theta_{\text{eq}} = 1/(k_{\text{eq}}D_V) \propto \Omega_m^{-0.83} h^{-1}$ , where  $k_{\text{eq}}$  denotes the equality wave number] at this maximum wave number. Indeed, the contours from the two teams are roughly oriented along this direction, depicted by the red dashed line. At  $k_{\max} = 0.2 h \text{Mpc}^{-1}$ , as we can clearly see the BAO feature up to the third peak (see Fig. 2), the BAO scale (the blue dashed line in Fig. 8:  $\theta_{\text{BAO}} = r_s/D_V$  with the sound horizon scale  $r_s$ ) plays a more significant role. The measurement of the relative location of these two characteristic scales allows us to determine the physical density  $\omega_m = \Omega_m h^2$ , and together with the distance measurement through the cosmology dependence of the redshift-distance conversion (i.e., a measurement of  $D_V$ ), we can break the degeneracy between  $\Omega_m$  and  $H_0$ .

Once  $D_V(z)$  and  $\omega_m$  are fixed, the other parameters—such as the distance parameters,  $H(z)$ ,  $D_A(z)$  (with  $h$  kept in the units as  $h \text{Mpc}^{-1}$  or  $h^{-1} \text{Mpc}$ ) or the growth parameter,  $f(z)$ —are merely dependent parameters fully determined by  $\Omega_m$ , given that we focus on the flat  $\Lambda$ CDM cosmology. Had we fitted the data with a more general

expansion model, e.g., dynamical dark energy or modified gravity models, the posterior distribution of these parameters would have been different. The parameters extracted

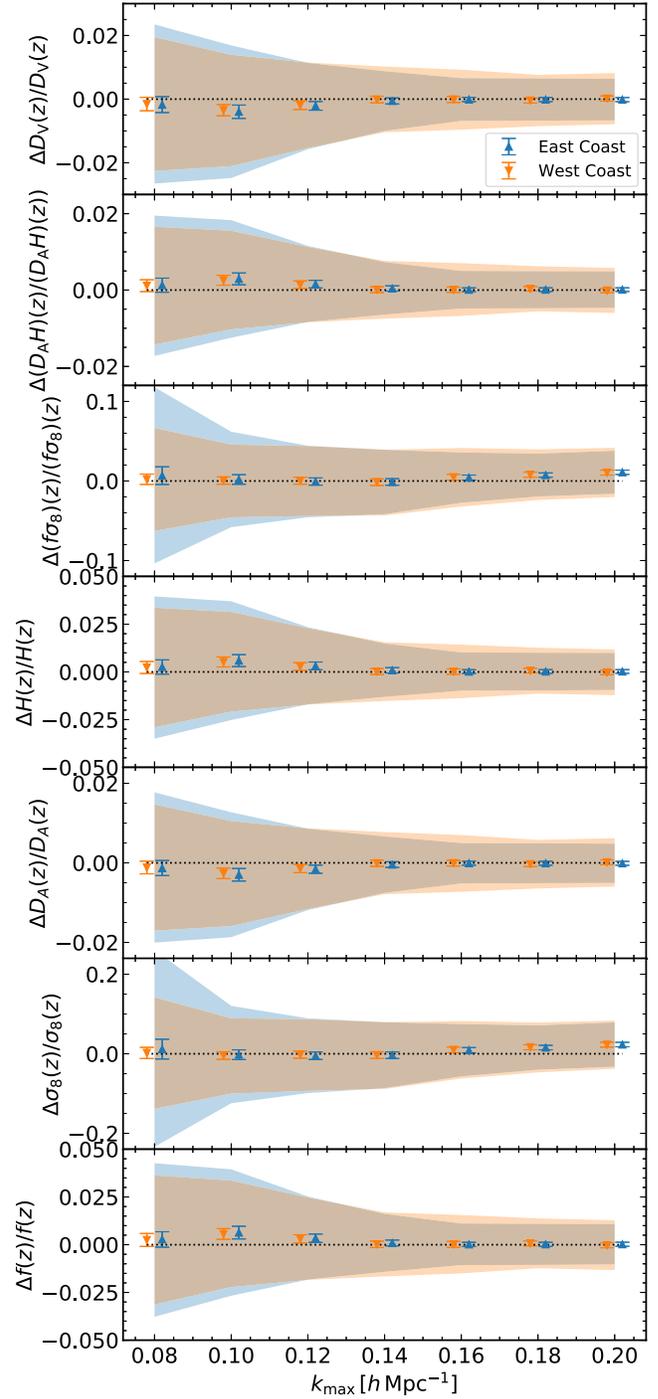


FIG. 9. One-dimensional marginalized posterior distributions of *derived* parameters for the flat  $\Lambda$ CDM model as a function of the maximum wave number included in the analysis,  $k_{\max}$ . The fractional error is shown with the uncertainty in  $H_0$  that is kept in the unit for the distance parameters (i.e.,  $D_A$  is expressed in  $h^{-1} \text{Mpc}$  and  $H$  is in  $h \text{Mpc}^{-1}$ ).

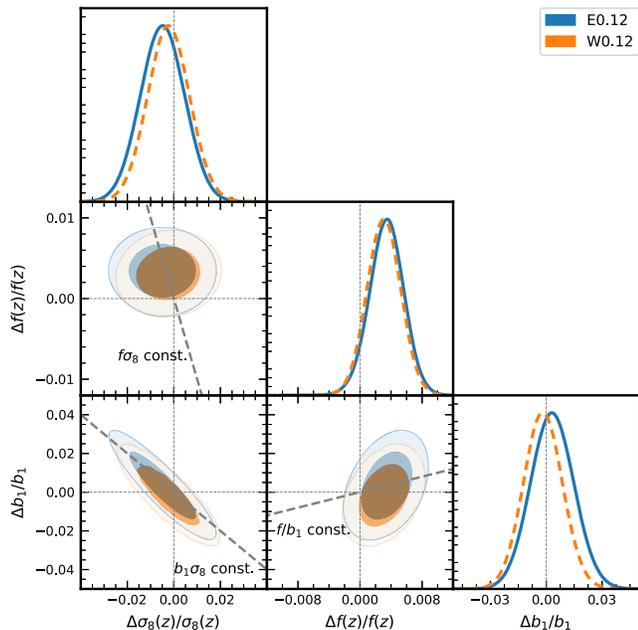


FIG. 10. Two-dimensional marginalized posterior distributions for amplitude-related parameters relevant for the RSD measurement from the analyses at  $k_{\text{max}} = 0.12 h \text{ Mpc}^{-1}$ . The expected degeneracy directions— $f\sigma_8$ ,  $b_1\sigma_8$ , or  $f/b_1$  is constant—expected from linear RSD measurements are shown by the dashed lines. Note that  $f(z)$  and  $\sigma_8(z)$  are derived parameters fully fixed once  $\Omega_m$  and  $A_s$  are given within the flat  $\Lambda$ CDM model.

from our MCMC chains, together with some other useful parameters, are displayed in Fig. 9.

Apart from the shape-related parameters, the determination of the amplitude parameter is of interest. We can see in Fig. 8 that the posterior of the amplitude parameter,  $A_s$ , is strongly correlated with the linear bias parameter  $b_1$ . To understand this more clearly, we show the constraints on the parameters relevant for the measurement of RSD (the one-dimensional and the two-dimensional marginalized posteriors in Figs. 9 and 10, respectively). In the two-dimensional contour plot, we can see that the amplitude parameter scaled to the redshift of the survey volume,  $\sigma_8(z) = [D_+(z)/D_+(z=0)]\sigma_8$ , is strongly degenerate with the linear bias parameter  $b_1$ , just as we have seen for  $A_s$  and  $b_1$ . In fact, they are expected to be fully degenerate in the absence of RSD information in linear theory. We can also see in Fig. 7 that  $b_1$  starts to depend weakly on  $k_{\text{max}}$  above  $\sim 0.14 h \text{ Mpc}^{-1}$  with statistical significance, and a similar departure from the ground truth value happens at the same place but to the opposite direction in  $\sigma_8(z)$ , as shown in Fig. 9. The other famous degeneracy directions,  $f\sigma_8$  or  $f/b$ , which are the direct observables from linear RSD, do not appear in our contours in Fig. 10. This is again due to the fact that the flat  $\Lambda$ CDM assumption makes  $f$  a dependent variable fully determined by  $\Omega_m$ . What we see here is that the constraint on  $\Omega_m$  through the shape and distance measurement discussed above, combined with the

measurement of  $f\sigma_8$  from RSD, allows us to constrain  $\sigma_8$  (and thus  $A_s$ ) directly.

## VI. CONCLUSION AND OUTLOOK

In this paper we have presented results of the blinded cosmology challenge initiated to test theoretical models for redshift-space galaxy clustering. The task was to assess whether the theoretical model, here EFTofLSS, can recover the blinded cosmological parameters in N-body simulation from the mock data of redshift-space power-spectrum multipoles for BOSS-like galaxies. The sufficiently large volume, dynamical range, and high resolution of the challenge simulation allow one to pin down any potential inaccuracy of theoretical modeling, compared to the statistical errors for the BOSS-like survey.

The simulations were run by a team (“Japan team”) that kept the true parameters a secret. The mock data were analyzed by two other independent teams (“East Coast team” and “West Coast team”) who volunteered to participate in the challenge. The rule of the challenge was that the true parameters can be unblinded only when the analyzing teams submitted their final results to the simulation team. All three teams agreed that the submitted results would be presented in this paper, without any change, after the unblinding.

Both analyzing teams used the same theoretical model based on the effective field theory of large-scale structure. However, there exist some nontrivial differences, whose impact on the final cosmological inference should be tested quantitatively with care. The corresponding pipelines were the ones applied to the real BOSS data in Refs. [28–30]. We have discussed in detail methodological and technical differences between these two pipelines. Despite these differences, both teams successfully recovered the true cosmological parameters within expected statistical error bars. This suggests that perturbation theory, once consistently implemented, can be used as a standard tool for unbiased estimation of cosmological parameters from galaxy surveys.

The enormously large total simulation volume used in the challenge helped to assess systematic error due to the incomplete theoretical modeling by suppressing the statistical error to a level much lower than the current surveys. The biased cosmological inference beyond the reported maximum wave number used for the challenge,  $k_{\text{max}} = 0.12 h \text{ Mpc}^{-1}$ , consistently determined by both teams, indicates the typical systematic error one can make from actual surveys with much smaller observed volume (see, e.g., Fig. 7 up to  $k_{\text{max}} = 0.2 h \text{ Mpc}^{-1}$ ). For instance, the analyses of SDSS BOSS galaxies by Refs. [28–30] adopt  $k_{\text{max}}$  around  $0.2 h \text{ Mpc}^{-1}$  (0.18 to 0.25 depending on the paper and the redshift bin of the galaxy sample). While the detailed choice of varied cosmological parameters, as well as the ways to combine them with CMB constraints, are different from what is presented here, one can make a

reasonable guess on the potential systematic biases on the inferred cosmological parameters of these papers out of our results.

Out of the three cosmological parameters that we considered here, the scalar amplitude parameter  $A_s$ , is most severely biased beyond  $k_{\max} = 0.12 h \text{Mpc}^{-1}$ , reaching  $\sim 4\%$ <sup>11</sup> at  $k_{\max} = 0.2 h \text{Mpc}^{-1}$ , while the other two parameters,  $\Omega_m$  and  $H_0$ , are fairly unbiased, even when the EFT template starts to fail. This indicates that the latter two are mostly constrained through the shape of the spectrum (mostly the distinctive BAO feature). The situation should be the same in actual observational data analyses such as the one listed above. Although the precise value of the detected parameter bias on  $A_s$  can depend on the detail in the halo-galaxy connection mainly through the uncertainty in the strength of the redshift-space distortions, it is assuring to observe that our worst-case value of 4% is still below the statistical errors from Refs. [28–30], which are 12% to 19% (68% C.L.). Future experiments with even larger survey volume and higher galaxy number density will allow us to lower these uncertainties, and in that case, we have to be more careful on the parameter bias due to the model inaccuracy, either by lowering  $k_{\max}$  or by improving the model itself. We investigate the parameter constraints for a hypothetical survey with the volume of the Dark Energy Spectroscopic Instrument (DESI, Ref. [4]) in Appendix C.

We are currently exploring a number of various post-blinded research directions. The first one includes a thorough investigation of the information content of redshift galaxy surveys. Second, it would be interesting to see how much the  $k_{\max}$  value where one-loop perturbation theory breaks down depends on the properties of the galaxy population, i.e., assembly bias or satellite fraction. Third, it would also be interesting to see how well perturbation theory performs for other observables, e.g., the galaxy-galaxy weak lensing or the redshift-space bispectrum. These research avenues are left for future work.

We have presented the results obtained by analyzing teams in such a way that the true parameters are still blinded to the readers. This is done in case some other researchers would like to test their theory models on the challenge spectra. All challenge data are available online at [46]. We encourage all groups working on galaxy clustering analysis to participate in the challenge.

## ACKNOWLEDGMENTS

T. N., L. S., M. T., and M. Z. acknowledge the warm hospitality of the BCCP-IAS workshop “The Nonlinear Universe 2018” held at Smartno, Slovenia, where this work was initiated. This work is supported in part by World

Premier International Research Center Initiative (WPI Initiative), MEXT, Japan, and by MEXT/JSPS KAKENHI Grants No. JP17K14273 (T. N.), No. JP15H05887 (M. T.), No. JP15H05893 (M. T.), No. JP15K21733 (M. T.), and No. JP19H00677 (T. N. and M. T.). T. N. also acknowledges financial support from Japan Science and Technology Agency (JST) CREST Grant No. JPMJCR1414 and by JST AIP Acceleration Research Grant No. JP20317829, Japan. Numerical computations were carried out on Cray XC50 at the Center for Computational Astrophysics, National Astronomical Observatory of Japan. G. D. A. is partially supported by Simons Foundation Origins of the Universe program (Modern Inflationary Cosmology Collaboration). L. S. is partially supported by Simons Foundation Origins of the Universe program (Modern Inflationary Cosmology Collaboration) and by NSF Award No. 1720397. M. Z. is supported by NSF Grants No. AST1409709 and No. PHY-1820775, the Canadian Institute for Advanced Research (CIFAR) program on Gravity and the Extreme Universe, and the Simons Foundation Modern Inflationary Cosmology initiative. M. I. is partially supported by the Simons Foundations Origins of the Universe program and by the RFBR Grant No. 20-02-00982 A.

## APPENDIX A: GALAXY KERNELS

The explicit expressions for the galaxy kernels appearing in the one-loop power spectrum are given here (for a derivation, see [72]):

$$\begin{aligned}
 Z_1(\mathbf{q}_1) &= K_1(\mathbf{q}_1) + f\mu_1^2 G_1(\mathbf{q}_1) = b_1 + f\mu_1^2, \\
 Z_2(\mathbf{q}_1, \mathbf{q}_2, \mu) &= K_2(\mathbf{q}_1, \mathbf{q}_2) + f\mu_2^2 G_2(\mathbf{q}_1, \mathbf{q}_2) \\
 &\quad + \frac{1}{2} f\mu q \left( \frac{\mu_2}{q_2} G_1(\mathbf{q}_2) Z_1(\mathbf{q}_1) + \text{perm.} \right), \\
 Z_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mu) &= K_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) + f\mu_{123}^2 G_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \\
 &\quad + \frac{1}{3} f\mu q \left( \frac{\mu_3}{q_3} G_1(\mathbf{q}_3) Z_2(\mathbf{q}_1, \mathbf{q}_2, \mu_{123}) \right. \\
 &\quad \left. + \frac{\mu_{23}}{q_{23}} G_2(\mathbf{q}_2, \mathbf{q}_3) Z_1(\mathbf{q}_1) + \text{cyc.} \right), \quad (\text{A1})
 \end{aligned}$$

where  $\mu = \mathbf{q} \cdot \hat{\mathbf{z}}/q$ ,  $\mathbf{q} = \mathbf{q}_1 + \dots + \mathbf{q}_n$ , and  $\mu_{i_1 \dots i_n} = \mathbf{q}_{i_1 \dots i_n} \cdot \hat{\mathbf{z}}/q_{i_1 \dots i_n}$ ,  $\mathbf{q}_{i_1 \dots i_m} = \mathbf{q}_{i_1} + \dots + \mathbf{q}_{i_m}$ , with  $\hat{\mathbf{z}}$  being the unit vector in the direction of the line of sight, and  $n$  is the order of the kernel  $Z_n$ . Note that  $K_i$  and  $G_i$  are the galaxy density and velocity kernels, respectively. We choose to work in the basis of descendants (this is the first complete set of bias coefficients for LSS, established in [19,100] and with some typos corrected in [101]; see [19,100] for connection to former bases of bias coefficients, as in for example [14]). Notice that while the new terms introduced in [19] happen to be degenerate with the standard bias terms at one-loop order, this will not be

<sup>11</sup>We estimate this theory-systematic error as the distance from the truth of the  $1\sigma$  region of the posterior, as done in [28].

the case once one goes to higher orders. For the one-loop power spectrum, all kernels can be described with four bias parameters  $b_i$ .

The first- and second-order galaxy density kernels are

$$K_1 = b_1,$$

$$K_2(\mathbf{q}_1, \mathbf{q}_2) = b_1 \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1^2} + b_2 \left( F_2(\mathbf{q}_1, \mathbf{q}_2) - \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1^2} \right) + b_4 + \text{perm.} \quad (\text{A2})$$

The galaxy velocity kernels  $G_n$  are simply the standard perturbation theory ones since the galaxy velocity field follows the dark matter velocity field, up to higher-derivative terms which are degenerate with other counterterms that appear in the renormalization of the redshift-space expression (see e.g., [69] for the expressions of  $F_n$  and  $G_n$ ).

The third-order galaxy density kernel has a much more involved expression. However, for the one-loop calculation, degeneracies appear in the one-loop diagram obtained from  $\langle \delta^{(3)} \delta^{(1)} \rangle$ , when UV divergences are removed and the integral over the angular coordinates is performed, leading to the following simple expression:

$$K_3(k, q) = \frac{b_1}{504k^3q^3} \left( -38k^5q + 48k^3q^3 - 18kq^5 \right. \\ \left. + 9(k^2 - q^2)^3 \log \left[ \frac{k-q}{k+q} \right] \right) \\ + \frac{b_3}{756k^3q^5} \left( 2kq(k^2 + q^2)(3k^4 - 14k^2q^2 + 3q^4) \right. \\ \left. + 3(k^2 - q^2)^4 \log \left[ \frac{k-q}{k+q} \right] \right). \quad (\text{A3})$$

## APPENDIX B: POST-UNBLINDING ANALYSES

While both analysis teams worked on a specific bias parametrization in the main text, it is worth exploring different options for a better understanding and to make a better connection to some of the recent works on observational data. This Appendix presents two such possibilities.

### 1. Residual shot noise

It is known that dark matter halos or associated galaxies are not a Poisson sample of the underlying hypothetical continuous distribution (e.g., [114,115]). As explained in Sec. III C, the standard shot-noise contribution is already

subtracted in the power-spectra data files provided by the Japan team. The subtracted shot-noise contribution, strictly speaking, is not really an estimate of the additional fluctuations associated with the connection between the underlying smooth field and the discrete point distribution but simply the “zero-lag” correlator inherent in a point process. Therefore, the assumption of the zero shot-noise-like term adopted in the blinded analyses presented in the main text is not guaranteed to be valid. Here we study the impact of adding a nuisance parameter to model the residual shot term, which is relevant for the monopole moment.

The green contours in Fig. 11 show the result from the East Coast team at four different  $k_{\text{max}}$  as indicated in the figure legend. Figure 12 shows the same analysis done by the West Coast team. Both results are compared with the open dotted contours from the blinded analysis. Introducing one more free parameter indeed results in slightly looser constraints with tilted degeneracy directions. An interesting observation is that the biases that the East Coast team displays on  $H_0$  and  $A_s$  at  $k_{\text{max}} = 0.16$  and  $0.18 \text{ h Mpc}^{-1}$  are reduced. Similarly, the West Coast team has a bias just on  $A_s$ , which is also reduced. Although the introduction of the non-Poissonian shot-noise term ceases to mitigate the parameter bias at  $k_{\text{max}} = 0.2 \text{ h Mpc}^{-1}$ , probably due to the absence of terms higher than one loop in the theoretical template, this parameter would allow for a more robust analysis in an actual analysis of observational data. This is exactly what was done by both teams in their analyses of the BOSS data.

### 2. Floating $b_{\Gamma_3}$

The East Coast team did not vary  $b_{\Gamma_3}$  in their analysis because they found that it does not affect the parameter constraints. This is explicitly illustrated in this Appendix. Note that  $b_{\Gamma_3}$  has very little impact on the power-spectrum constraints because this bias parameter is strongly degenerate with  $b_{\mathcal{G}_2}$ . To break this degeneracy, the East Coast team used the Gaussian prior centered at the prediction of the coevolution model,

$$b_{\Gamma_3} \sim \mathcal{N} \left( \frac{23}{42} (b_1 - 1), 1 \right). \quad (\text{B1})$$

The results of the analysis with this prior are shown in Fig. 11. One clearly sees that varying  $b_{\Gamma_3}$  or fixing it to a constant value has no noticeable impact on the cosmological parameters or the linear bias  $b_1$ .

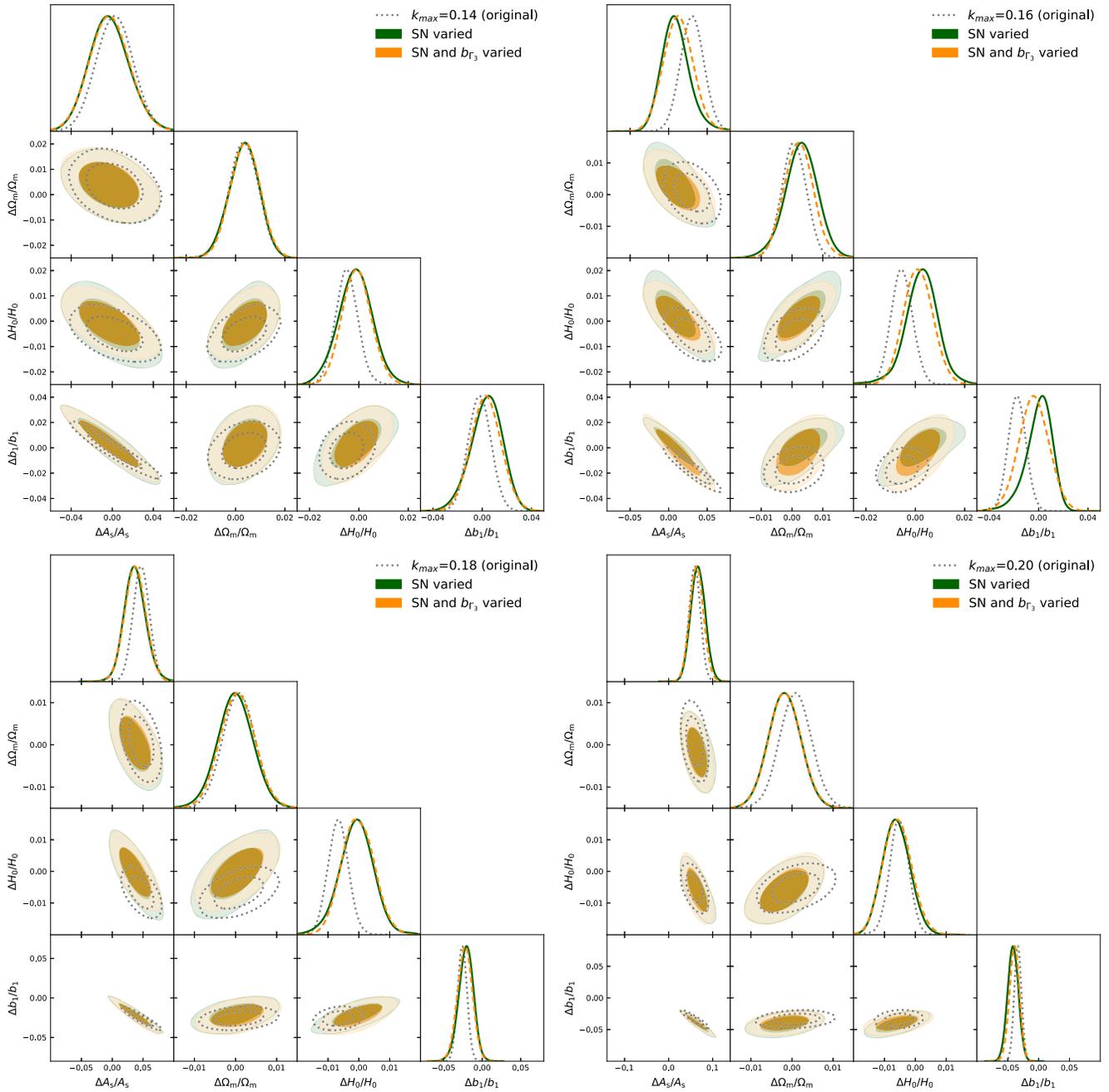


FIG. 11. Posterior distributions from the post-unblinding analyses where one or two additional bias parameters are floated. The results are from the pipeline by the East Coast team.

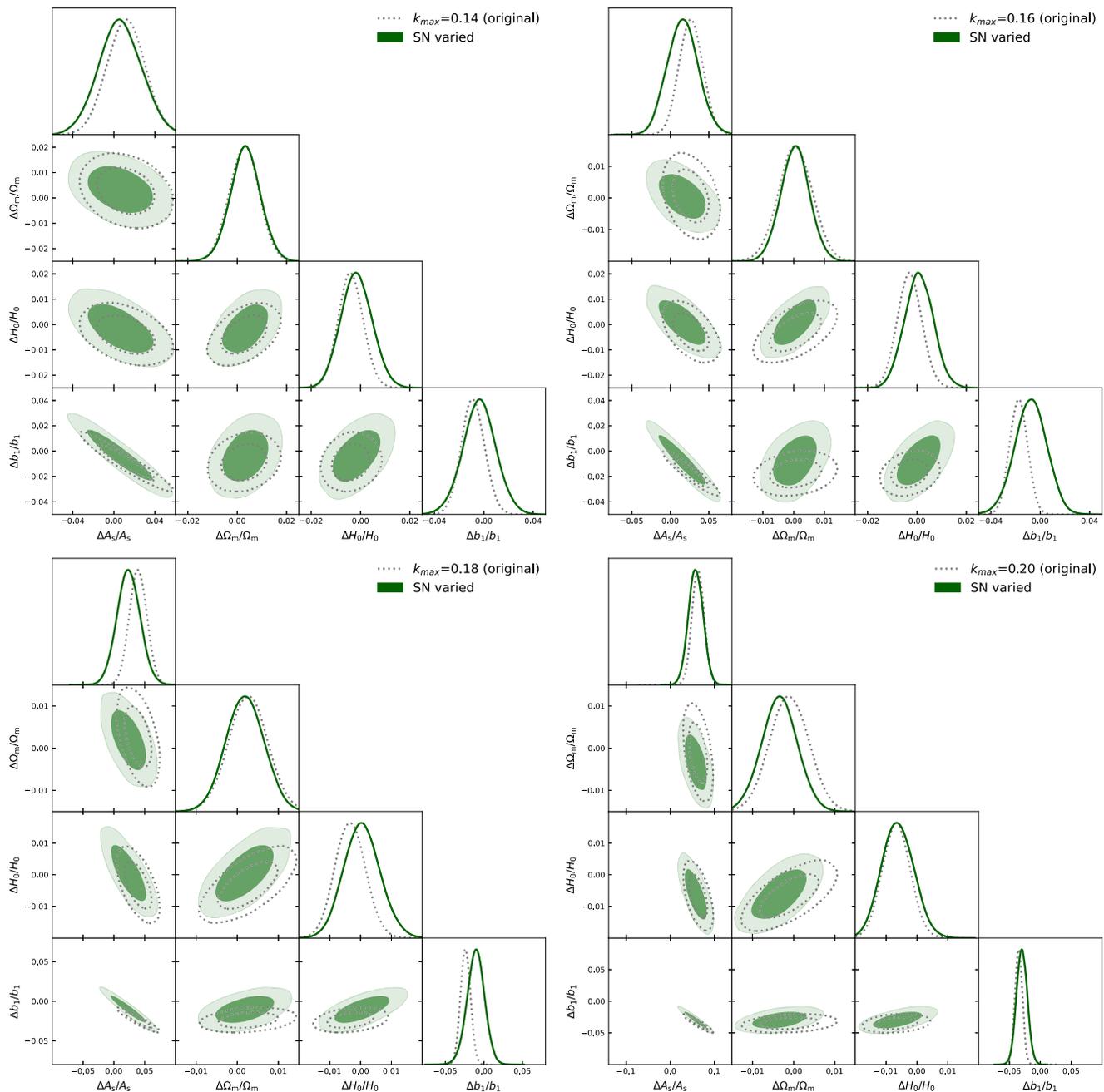


FIG. 12. Similar to Fig. 11, but from the pipeline by the West Coast team for the case of a varied residual shot-noise term.

### APPENDIX C: SCALING TO REALISTIC SURVEYS

We have focused on the statistical inference from mock spectra measured from an unrealistically large total volume in the main text. To gain further insights into more realistic observations, we redo the analysis after unblinding with a much smaller volume. We consider a hypothetical survey with the volume of  $25.5(h^{-1} \text{ Gpc})^3$  and scale the error bars from the mock simulations according to the volume ratio. This is close to the expected survey volume of the DESI survey. The results of this Appendix should be taken with

care since the effective redshift of the DESI survey is higher than the one we use here, which is similar to the one of the BOSS survey instead. As a consequence, the effect of the nonlinear corrections, and so the systematic error measured here, is larger than what we expect for the actual DESI survey.

We show the results using the same analysis pipeline as the one used by the West Coast team in the main text. Due to the larger error bars in the mock spectra, we can push to smaller scales without bias in the inferred cosmological parameters. Figure 13 shows the  $1\sigma$  and  $2\sigma$  credible regions

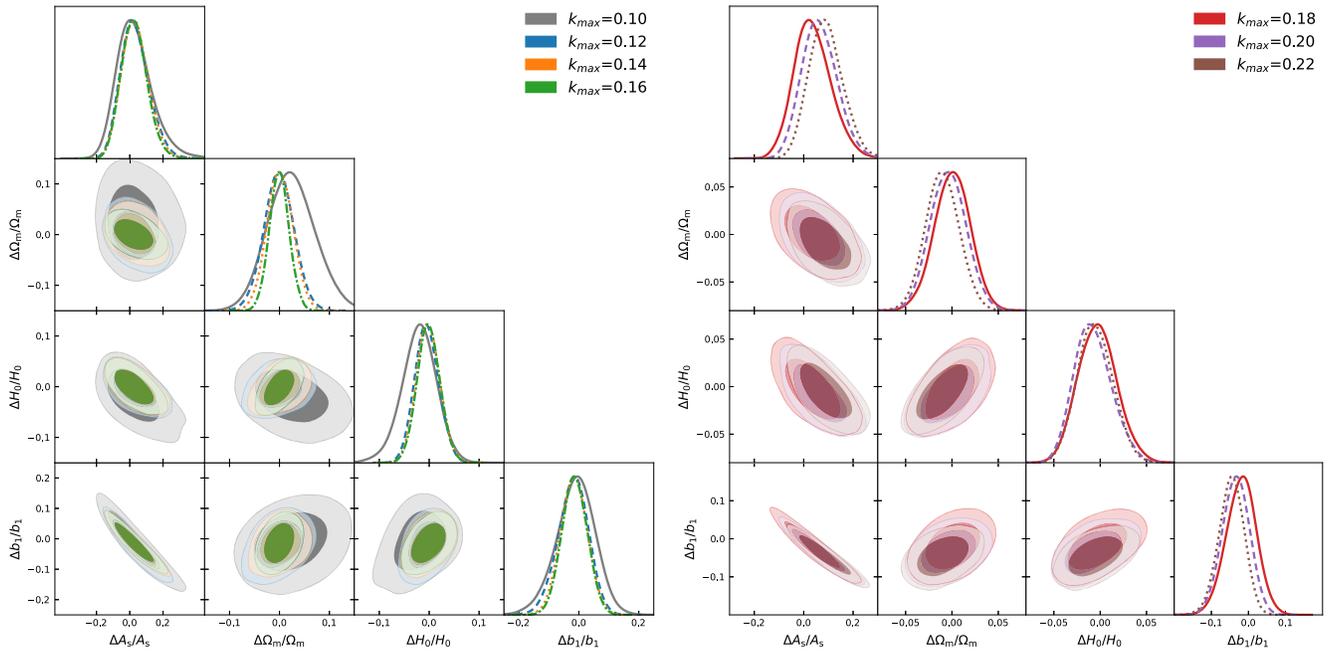


FIG. 13. Expected constraints from a DESI-like survey at different maximum wave numbers (left: 0.1, 0.12, 0.14, and 0.16  $h \text{ Mpc}^{-1}$ ; right: 0.18, 0.2, and 0.22  $h \text{ Mpc}^{-1}$ ). We use the pipeline by the West Coast team in the main text for this plot.

for the three varied cosmological parameters as well as the linear bias parameter from this analysis. We show the results at different values of  $k_{\text{max}}$  up to  $0.22 h \text{ Mpc}^{-1}$  as indicated by the figure legend in two panels for ease of visibility (lower  $k_{\text{max}}$  in the left and higher  $k_{\text{max}}$  in the right panel).

Overall, we can see that the inferred parameters are unbiased compared to the statistical error level expected from a DESI-like survey for all the  $k_{\text{max}}$  values considered here. With respect to the results of the BOSS analysis

performed in [28], the error bars at  $k_{\text{max}} = 0.2 h^{-1} \text{ Mpc}$  shrink from 3.2% to 2% on  $\Omega_m$ , from 3.2% to 2.1% on  $H_0$ , and from 13% to 6.7% on  $A_s$ . However, from the runs described in the main text, there is a systematic error on  $A_s$  of  $\sim 4\%$  at  $k_{\text{max}} = 0.2 h^{-1} \text{ Mpc}$ , which corresponds to about 2/3 of the statistical error here. These results are particularly encouraging in view of the fact that DESI will survey a higher redshift than the one of the simulations, where nonlinear corrections will be less important.

- 
- [1] M. Takada, R. S. Ellis, M. Chiba, J. E. Greene, H. Aihara, N. Arimoto, K. Bundy, J. Cohen, O. Doré, G. Graves *et al.*, *Publ. Astron. Soc. Jpn.* **66**, R1 (2014).
- [2] R. Laureijs, J. Amiaux, S. Arduini, J. Auguères, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet *et al.*, [arXiv:1110.3193](https://arxiv.org/abs/1110.3193).
- [3] P. A. Abell, J. Allison, S. F. Anderson, J. R. Andrew, J. R. P. Angel, L. Armus, D. Arnett, S. J. Asztalos, T. S. Axelrod *et al.* (LSST Science Collaboration), [arXiv:0912.0201](https://arxiv.org/abs/0912.0201).
- [4] M. Levi, L. E. Allen, A. Raichoor, C. Baltay, S. BenZvi, F. Beutler, A. Bolton, F. J. Castander, C.-H. Chuang, A. Cooper *et al.*, *Bull. Am. Astron. Soc.* **51**, 57 (2019), <https://ui.adsabs.harvard.edu/abs/2019BAAS...51g..57L/abstract>.
- [5] N. Kaiser, *Astrophys. J. Lett.* **284**, L9 (1984).
- [6] M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, S. Bird, D. Nelson, and L. Hernquist, *Nature (London)* **509**, 177 (2014).
- [7] S. Genel, M. Vogelsberger, V. Springel, D. Sijacki, D. Nelson, G. Snyder, V. Rodriguez-Gomez, P. Torrey, and L. Hernquist, *Mon. Not. R. Astron. Soc.* **445**, 175 (2014).
- [8] M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, D. Nelson, and L. Hernquist, *Mon. Not. R. Astron. Soc.* **444**, 1518 (2014).
- [9] R. A. Crain, J. Schaye, R. G. Bower, M. Furlong, M. Schaller, T. Theuns, C. Dalla Vecchia, C. S. Frenk, I. G. McCarthy, J. C. Helly *et al.*, *Mon. Not. R. Astron. Soc.* **450**, 1937 (2015).
- [10] J. Schaye, R. A. Crain, R. G. Bower, M. Furlong, M. Schaller, T. Theuns, C. Dalla Vecchia, C. S. Frenk, I. G. McCarthy, J. C. Helly *et al.*, *Mon. Not. R. Astron. Soc.* **446**, 521 (2015).
- [11] Y. Dubois, C. Pichon, C. Welker, D. Le Borgne, J. Devriendt, C. Laigle, S. Codis, D. Pogosyan, S. Arnouts,

- K. Benabed *et al.*, *Mon. Not. R. Astron. Soc.* **444**, 1453 (2014).
- [12] V. Springel, R. Pakmor, A. Pillepich, R. Weinberger, D. Nelson, L. Hernquist, M. Vogelsberger, S. Genel, P. Torrey, F. Marinacci *et al.*, *Mon. Not. R. Astron. Soc.* **475**, 676 (2018).
- [13] D. Nelson, V. Springel, A. Pillepich, V. Rodriguez-Gomez, P. Torrey, S. Genel, M. Vogelsberger, R. Pakmor, F. Marinacci, R. Weinberger *et al.*, *Comput. Astrophys. Cosmol.* **6**, 2 (2019).
- [14] P. McDonald and A. Roy, *J. Cosmol. Astropart. Phys.* **08** (2009) 020.
- [15] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga, *J. Cosmol. Astropart. Phys.* **07** (2012) 051.
- [16] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore, *J. High Energy Phys.* **09** (2012) 082.
- [17] V. Assassi, D. Baumann, D. Green, and M. Zaldarriaga, *J. Cosmol. Astropart. Phys.* **08** (2014) 056.
- [18] L. Senatore and M. Zaldarriaga, *J. Cosmol. Astropart. Phys.* **02** (2015) 013.
- [19] L. Senatore, *J. Cosmol. Astropart. Phys.* **11** (2015) 007.
- [20] L. Senatore and M. Zaldarriaga, [arXiv:1409.1225](https://arxiv.org/abs/1409.1225).
- [21] M. Lewandowski, A. Perko, and L. Senatore, *J. Cosmol. Astropart. Phys.* **05** (2015) 019.
- [22] M. Lewandowski, L. Senatore, F. Prada, C. Zhao, and C.-H. Chuang, *Phys. Rev. D* **97**, 063526 (2018).
- [23] V. Desjacques, D. Jeong, and F. Schmidt, *Phys. Rep.* **733**, 1 (2018).
- [24] N. Hand, U. Seljak, F. Beutler, and Z. Vlah, *J. Cosmol. Astropart. Phys.* **10** (2017) 009.
- [25] Y. Kobayashi, T. Nishimichi, M. Takada, and R. Takahashi, *Phys. Rev. D* **101**, 023510 (2020).
- [26] K. Osato, T. Nishimichi, F. Bernardeau, and A. Taruya, *Phys. Rev. D* **99**, 063530 (2019).
- [27] A. G. Riess, *Nat. Rev. Phys.* **2**, 10 (2020).
- [28] G. D'Amico, J. Gleyzes, N. Kokron, D. Markovic, L. Senatore, P. Zhang, F. Beutler, and H. Gil-Marn, [arXiv:1909.05271](https://arxiv.org/abs/1909.05271).
- [29] M. M. Ivanov, M. Simonović, and M. Zaldarriaga, *J. Cosmol. Astropart. Phys.* **05** (2020) 042.
- [30] T. Colas, G. D'Amico, L. Senatore, P. Zhang, and F. Beutler, [arXiv:1909.07951](https://arxiv.org/abs/1909.07951).
- [31] T. Tröster, A. G. Sánchez, M. Asgari, C. Blake, M. Crocce, C. Heymans, H. Hildebrandt, B. Joachimi, S. Joudaki, A. Kannawadi *et al.*, *Astron. Astrophys.* **633**, L10 (2020).
- [32] H. Hildebrandt, M. Viola, C. Heymans, S. Joudaki, K. Kuijken, C. Blake, T. Erben, B. Joachimi, D. Klaes, and L. Miller, *Mon. Not. R. Astron. Soc.* **465**, 1454 (2017).
- [33] T. M. C. Abbott, F. B. Abdalla, A. Alarcon, J. Aleksić, S. Allam, S. Allen, A. Amara, J. Annis, J. Asorey, S. Avila *et al.*, *Phys. Rev. D* **98**, 043526 (2018).
- [34] C. Hikage, M. Oguri, T. Hamana, S. More, R. Mandelbaum, M. Takada, F. Köhlinger, H. Miyatake, A. J. Nishizawa, and H. Aihara, *Publ. Astron. Soc. Jpn.* **71**, 43 (2019).
- [35] P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett *et al.* (Planck Collaboration), *Astron. Astrophys.* **594**, A13 (2016).
- [36] A. Lewis, A. Challinor, and A. Lasenby, *Astrophys. J.* **538**, 473 (2000).
- [37] P. J. E. Peebles and J. T. Yu, *Astrophys. J.* **162**, 815 (1970).
- [38] R. A. Sunyaev and Y. B. Zeldovich, *Astrophys. Space Sci.* **7**, 3 (1970), <https://link.springer.com/article/10.1007/BF00653471>.
- [39] J. R. Bond and G. Efstathiou, *Astrophys. J. Lett.* **285**, L45 (1984).
- [40] J. R. Bond and G. Efstathiou, *Mon. Not. R. Astron. Soc.* **226**, 655 (1987).
- [41] J. A. Holtzman, *Astrophys. J. Suppl. Ser.* **71**, 1 (1989).
- [42] J. C. Jackson, *Mon. Not. R. Astron. Soc.* **156**, 1P (1972).
- [43] N. Kaiser, *Mon. Not. R. Astron. Soc.* **227**, 1 (1987).
- [44] C. Alcock and B. Paczynski, *Nature (London)* **281**, 358 (1979).
- [45] S. Alam, F. D. Albareti, C. Allende Prieto, F. Anders, S. F. Anderson, T. Anderton, B. H. Andrews, E. Armengaud, É. Aubourg, S. Bailey *et al.*, *Astrophys. J. Suppl. Ser.* **219**, 12 (2015).
- [46] <http://www2.yukawa.kyoto-u.ac.jp/~takahiro.nishimichi/data/PTchallenge/>.
- [47] K. S. Dawson, D. J. Schlegel, C. P. Ahn, S. F. Anderson, É. Aubourg, S. Bailey, R. H. Barkhouser, J. E. Bautista, A. r. Beifiori, A. A. Berlind *et al.*, *Astron. J.* **145**, 10 (2013).
- [48] T. Nishimichi, A. Shirata, A. Taruya, K. Yahata, S. Saito, Y. Suto, R. Takahashi, N. Yoshida, T. Matsubara, N. Sugiyama *et al.*, *Publ. Astron. Soc. Jpn.* **61**, 321 (2009).
- [49] P. Valageas and T. Nishimichi, *Astron. Astrophys.* **527**, A87 (2011).
- [50] R. Scoccimarro, *Mon. Not. R. Astron. Soc.* **299**, 1097 (1998).
- [51] M. Crocce, S. Pueblas, and R. Scoccimarro, *Mon. Not. R. Astron. Soc.* **373**, 369 (2006).
- [52] T. Nishimichi, M. Takada, R. Takahashi, K. Osato, M. Shirasaki, T. Oogi, H. Miyatake, M. Oguri, R. Murata, Y. Kobayashi *et al.*, *Astrophys. J.* **884**, 29 (2019).
- [53] B. Marcos, T. Baertschiger, M. Joyce, A. Gabrielli, and F. Sylos Labini, *Phys. Rev. D* **73**, 103507 (2006).
- [54] M. Joyce and B. Marcos, *Phys. Rev. D* **76**, 103505 (2007).
- [55] L. H. Garrison, D. J. Eisenstein, D. Ferrer, M. V. Metchnik, and P. A. Pinto, *Mon. Not. R. Astron. Soc.* **461**, 4125 (2016).
- [56] V. Springel, *Mon. Not. R. Astron. Soc.* **364**, 1105 (2005).
- [57] P. S. Behroozi, R. H. Wechsler, and H.-Y. Wu, *Astrophys. J.* **762**, 109 (2013).
- [58] M. White, M. Blanton, A. Bolton, D. Schlegel, J. Tinker, A. Berlind, L. da Costa, E. Kazin, Y. T. Lin, M. Maia *et al.*, *Astrophys. J.* **728**, 126 (2011).
- [59] S. More, H. Miyatake, R. Mandelbaum, M. Takada, D. N. Spergel, J. R. Brownstein, and D. P. Schneider, *Astrophys. J.* **806**, 2 (2015).
- [60] E. Sefusatti, M. Crocce, R. Scoccimarro, and H. M. P. Couchman, *Mon. Not. R. Astron. Soc.* **460**, 3624 (2016).
- [61] Y. P. Jing, *Astrophys. J.* **620**, 559 (2005).
- [62] H. A. Feldman, N. Kaiser, and J. A. Peacock, *Astrophys. J.* **426**, 23 (1994).
- [63] F. Beutler, H.-J. Seo, S. Saito, C.-H. Chuang, A. J. Cuesta, D. J. Eisenstein, H. Gil-Marín, J. N. Grieb, N. Hand, F.-S. Kitaura *et al.*, *Mon. Not. R. Astron. Soc.* **466**, 2242 (2017).
- [64] T. Baldauf, M. Mirbabayi, M. Simonović, and M. Zaldarriaga, *Phys. Rev. D* **92**, 043514 (2015).

- [65] L. Senatore and G. Trevisan, *J. Cosmol. Astropart. Phys.* **05** (2018) 019.
- [66] M. Lewandowski and L. Senatore, [arXiv:1810.11855](https://arxiv.org/abs/1810.11855).
- [67] D. Blas, M. Garny, M. M. Ivanov, and S. Sibiryakov, *J. Cosmol. Astropart. Phys.* **07** (2016) 028.
- [68] R. A. Porto, L. Senatore, and M. Zaldarriaga, *J. Cosmol. Astropart. Phys.* **05** (2014) 022.
- [69] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro, *Phys. Rep.* **367**, 1 (2002).
- [70] D. Blas, M. Garny, M. M. Ivanov, and S. Sibiryakov, *J. Cosmol. Astropart. Phys.* **07** (2016) 052.
- [71] M. Mirbabayi, F. Schmidt, and M. Zaldarriaga, *J. Cosmol. Astropart. Phys.* **07** (2015) 030.
- [72] A. Perko, L. Senatore, E. Jennings, and R. H. Wechsler, [arXiv:1610.09321](https://arxiv.org/abs/1610.09321).
- [73] A. Chudaykin and M. M. Ivanov, *J. Cosmol. Astropart. Phys.* **11** (2019) 034.
- [74] M. M. Ivanov, M. Simonović, and M. Zaldarriaga, *J. Cosmol. Astropart. Phys.* **05** (2020) 042.
- [75] M. M. Ivanov and S. Sibiryakov, *J. Cosmol. Astropart. Phys.* **07** (2018) 053.
- [76] M. Simonović, T. Baldauf, M. Zaldarriaga, J. J. Carrasco, and J. A. Kollmeier, *J. Cosmol. Astropart. Phys.* **04** (2018) 030.
- [77] D. Blas, J. Lesgourgues, and T. Tram, *J. Cosmol. Astropart. Phys.* **07** (2011) 034.
- [78] A. Chudaykin, M. M. Ivanov, O. H. Philcox, and M. Simonović, *Phys. Rev. D* **102**, 063533 (2020).
- [79] B. Audren, J. Lesgourgues, K. Benabed, and S. Prunet, *J. Cosmol. Astropart. Phys.* **02** (2013) 001.
- [80] T. Brinckmann and J. Lesgourgues, [arXiv:1804.07261](https://arxiv.org/abs/1804.07261).
- [81] A. Lewis and S. Bridle, *Phys. Rev. D* **66**, 103511 (2002).
- [82] J. J. M. Carrasco, S. Foreman, D. Green, and L. Senatore, *J. Cosmol. Astropart. Phys.* **07** (2014) 056.
- [83] J. J. M. Carrasco, S. Foreman, D. Green, and L. Senatore, *J. Cosmol. Astropart. Phys.* **07** (2014) 057.
- [84] S. M. Carroll, S. Leichenauer, and J. Pollack, *Phys. Rev. D* **90**, 023518 (2014).
- [85] T. Baldauf, E. Schaan, and M. Zaldarriaga, *J. Cosmol. Astropart. Phys.* **03** (2016) 007.
- [86] S. Foreman, H. Perrier, and L. Senatore, *J. Cosmol. Astropart. Phys.* **05** (2016) 027.
- [87] T. Baldauf, L. Mercolli, and M. Zaldarriaga, *Phys. Rev. D* **92**, 123007 (2015).
- [88] M. Cataneo, S. Foreman, and L. Senatore, *J. Cosmol. Astropart. Phys.* **04** (2017) 026.
- [89] M. Lewandowski and L. Senatore, *J. Cosmol. Astropart. Phys.* **08** (2017) 037.
- [90] T. Konstandin, R. A. Porto, and H. Rubira, [arXiv:1906.00997](https://arxiv.org/abs/1906.00997).
- [91] E. Pajer and M. Zaldarriaga, *J. Cosmol. Astropart. Phys.* **08** (2013) 037.
- [92] A. A. Abolhasani, M. Mirbabayi, and E. Pajer, *J. Cosmol. Astropart. Phys.* **05** (2016) 063.
- [93] L. Mercolli and E. Pajer, *J. Cosmol. Astropart. Phys.* **03** (2014) 006.
- [94] M. McQuinn and M. White, *J. Cosmol. Astropart. Phys.* **01** (2016) 043.
- [95] R. E. Angulo, S. Foreman, M. Schmittfull, and L. Senatore, *J. Cosmol. Astropart. Phys.* **10** (2015) 039.
- [96] T. Baldauf, L. Mercolli, M. Mirbabayi, and E. Pajer, *J. Cosmol. Astropart. Phys.* **05** (2015) 007.
- [97] D. Bertolini, K. Schutz, M. P. Solon, and K. M. Zurek, [arXiv:1604.01770](https://arxiv.org/abs/1604.01770).
- [98] T. Baldauf, E. Schaan, and M. Zaldarriaga, *J. Cosmol. Astropart. Phys.* **03** (2016) 017.
- [99] S. Foreman and L. Senatore, *J. Cosmol. Astropart. Phys.* **04** (2016) 033.
- [100] R. Angulo, M. Fasiello, L. Senatore, and Z. Vlah, *J. Cosmol. Astropart. Phys.* **09** (2015) 029.
- [101] T. Fujita, V. Mauerhofer, L. Senatore, Z. Vlah, and R. Angulo, [arXiv:1609.00717](https://arxiv.org/abs/1609.00717).
- [102] E. O. Nadler, A. Perko, and L. Senatore, *J. Cosmol. Astropart. Phys.* **02** (2018) 058.
- [103] M. Lewandowski, A. Maleknejad, and L. Senatore, *J. Cosmol. Astropart. Phys.* **05** (2017) 038.
- [104] G. Cusin, M. Lewandowski, and F. Vernizzi, *J. Cosmol. Astropart. Phys.* **04** (2018) 005.
- [105] B. Bose, K. Koyama, M. Lewandowski, F. Vernizzi, and H. A. Winther, *J. Cosmol. Astropart. Phys.* **04** (2018) 063.
- [106] V. Assassi, D. Baumann, E. Pajer, Y. Welling, and D. van der Woude, *J. Cosmol. Astropart. Phys.* **11** (2015) 024.
- [107] V. Assassi, D. Baumann, and F. Schmidt, *J. Cosmol. Astropart. Phys.* **12** (2015) 043.
- [108] D. Bertolini, K. Schutz, M. P. Solon, J. R. Walsh, and K. M. Zurek, *Phys. Rev. D* **93**, 123505 (2016).
- [109] D. Bertolini and M. P. Solon, [arXiv:1608.01310](https://arxiv.org/abs/1608.01310).
- [110] L. Senatore and M. Zaldarriaga, [arXiv:1707.04698](https://arxiv.org/abs/1707.04698).
- [111] R. de Belsunce and L. Senatore, *J. Cosmol. Astropart. Phys.* **02** (2019) 038.
- [112] D. J. Eisenstein, H.-j. Seo, and M. J. White, *Astrophys. J.* **664**, 660 (2007).
- [113] T. Baldauf, M. Mirbabayi, M. Simonović, and M. Zaldarriaga, [arXiv:1602.00674](https://arxiv.org/abs/1602.00674).
- [114] R. E. Smith, R. Scoccimarro, and R. K. Sheth, *Phys. Rev. D* **75**, 063512 (2007).
- [115] U. Seljak, N. Hamaus, and V. Desjacques, *Phys. Rev. Lett.* **103**, 091303 (2009).