Extracting high-order cosmological information in galaxy surveys with power spectra

Check for updates

Yuting Wang [1,2], Gong-Bo Zhao [1,2,3] ✉, Kazuya Koyama [4] ✉, Will J. Percival [5,6,7] ✉,
Ryuichi Takahashi [8], Chiaki Hikage [9], Héctor Gil-Marín [10], ChangHoon Hahn [11], Ruiyang Zhao [1,3],
Weibing Zhang [1,3], Xiaoyong Mu [1], Yu Yu [12], Hong-Ming Zhu [13] & Fei Ge [14]

The reconstruction method was proposed more than a decade ago to boost the signal of baryonic acoustic oscillations measured in galaxy redshift surveys, which is one of key probes for dark energy. After moving the observed overdensities in galaxy surveys back to their initial position, the reconstructed density field is closer to a linear Gaussian field, with higher-order information moved back into the power spectrum. We find that by jointly analysing power spectra measured from the pre- and post-reconstructed galaxy samples, higher-order information beyond the 2-point power spectrum can be efficiently extracted, which generally yields an information gain upon the analysis using the pre- or post-reconstructed galaxy sample alone. This opens a window to easily use higher-order information when constraining cosmological models.

The science driver for massive galaxy spectroscopic surveys is to extract cosmological information from the clustering of galaxies in the past lightcone. Baryon acoustic oscillations (BAO)[1], formed in the early Universe due to interactions between photons and baryons under pressure and gravity, yield a special clustering pattern of galaxies around a characteristic comoving scale around 150 Mpc, which is one of key probes for dark energy[2,3]. The increasing size of galaxy redshift surveys over the decade 2000–2010 led ultimately to a 5σ detection of BAO by the Baryon Oscillation Spectroscopic Survey (BOSS)[4]. This enabled the BAO to be used as an accurate *standard ruler* to measure the geometry of the Universe and constrain the cosmic expansion history. A compilation of results from the Sloan Digital Sky Survey (SDSS) galaxy survey recently demonstrated the power of this technique[5].

The BAO feature is generally blurred by the nonlinear evolution of the Universe reducing its strength as a standard ruler, and various reconstruction methods have been developed to sharpen the BAO peak by undoing the nonlinear evolution of the density field. The commonly used Lagrangian

reconstruction, for example, linearises the density field by shifting the galaxies using the displacement field[6–8], while for the Eulerian reconstruction, manipulation is performed at the field level without moving the galaxies[9].

Although designed to boost the BAO signal originally, the reconstruction method can, in principle, also better extract the general cosmological information from the clustering. For example, the redshift space distortions (RSD)[10–12], which is caused by peculiar motions of galaxies under gravity, can also be better constrained using the reconstructed sample[13].

The standard method for BAO-reconstruction alters the over-density field so that it is more correlated with the initial linear field[14], and the level of mode-coupling can be highly reduced[15]. It achieves this by inferring the bulk-flows using the observed galaxy field and then removing these displacements from both the galaxy positions and the map of expected density setting the baseline from which the over-densities are found. The power spectrum of the post-reconstructed sample (P_{post}) provides additional information for cosmology compared with the pre-reconstructed sample

[1]National Astronomy Observatories, Chinese Academy of Sciences, Beijing 100101, PR China. [2]Institute for Frontiers in Astronomy and Astrophysics, Beijing Normal University, Beijing 102206, PR China. [3]School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, PR China. [4]Institute of Cosmology and Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth PO1 3FX, UK. [5]Waterloo Centre for Astrophysics, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada. [6]Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada. [7]Perimeter Institute for Theoretical Physics, 31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada. [8]Faculty of Science and Technology, Hirosaki University, 3 Bunkyo-cho, Hirosaki, Aomori 036-8561, Japan. [9]Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8583, Japan. [10]ICC, University of Barcelona, IEEC-UB, Martí i Franquès, 1, E-08028 Barcelona, Spain. [11]Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA. [12]Department of Astronomy, Shanghai Jiao Tong University, Shanghai 200240, PR China. [13]Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada. [14]Department of Physics and Astronomy, University of California, Davis, CA 95616, USA. ✉e-mail: gbzhao@nao.cas.cn; kazuya.koyama@port.ac.uk; will.percival@uwaterloo.ca
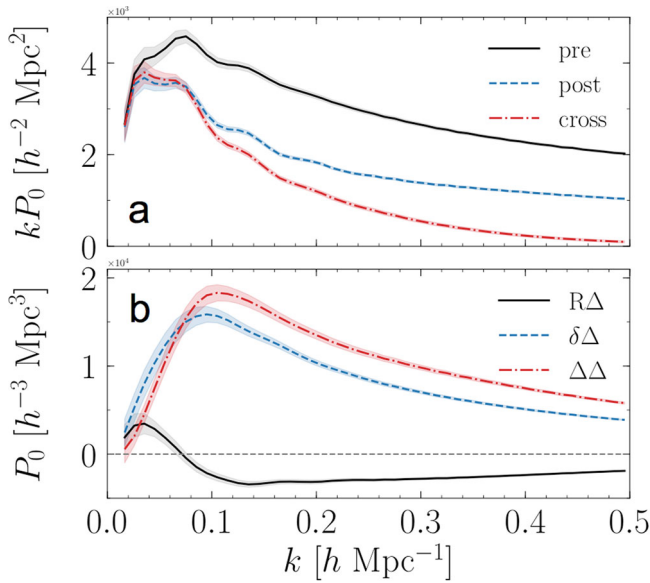
Fig. 1 | The measured power spectrum monopole. a The monopole (multiplied by k) of three types of power spectra indicated in the legend, measured from the MOLINO galaxy mocks; **b** The rotated power spectra monopole defined in Eq. (1). In both panels, the lines in the centre denote the mean of the mocks and the shades represent the 68% confidence level uncertainty.

($P_{\rm pre}$) because the reconstruction restores the linear signal reduced by the non-linear evolution.

The higher order statistics such as B, the bispectrum, induced by the non-linear evolution are, in turn, reduced. Thus we can extract more cosmological information encoded in the linear density field from $P_{\rm post}$ than $P_{\rm pre}$. On the other hand, the non-linear field contains information on small-scale clustering such as galaxy biases, which can provide better constraints on cosmological parameters by breaking the degeneracy between them. For the pre-reconstructed sample, this information can be extracted by combining the power spectrum with higher-order statistics. However, the power spectrum and higher order statistics such as the bispectrum are correlated, reducing our ability to estimate cosmological parameters. If we instead consider the post-reconstructed power spectrum, the covariance between the power spectrum and the bispectrum is reduced and we can extract the information more efficiently. In this work, we show that the same improvement can be achieved by a joint analysis of $P_{\rm pre}$, $P_{\rm post}$ and $P_{\rm cross}$ (the cross-power spectrum between the pre- and post-reconstructed density fields). Due to the restored linear signal in the reconstructed density field, $P_{\rm post}$ is decorrelated with $P_{\rm pre}$ on small scales, which are dominated by the non-linear information. On these scales, the combination of $P_{\rm pre}$, $P_{\rm post}$ and $P_{\rm cross}$ has a similar ability to extract cosmological information as the combination of $P_{\rm post}$ or $P_{\rm pre}$ with the bispectrum because we are able to use the linear information in $P_{\rm post}$ and higher-order information in $P_{\rm pre}$ separately.

Let us rewrite the non-linear over-density field as $\delta = R + \Delta$, where R is the over-density field after reconstruction, which is closer to the linear field. It is then straightforward to express $P_{\rm pre}$, $P_{\rm cross}$ in terms of $P_{RR}(=P_{\rm post})$, $P_{\Delta\Delta}$ (the power spectrum of Δ) and $P_{\Delta R}$ (the cross-power spectrum between Δ and R). Using perturbation theory[16], we can show that, at the leading order, $P_{\Delta\Delta}$ contains the integrated contribution from the bispectrum of squeezed-limit triangles while $P_{R\Delta}$ contains the integrated contribution from the trispectrum (T) of folded/squeeze-limit quadrilaterals (See Supplementary Note 1 for an explanation). In this fashion when combining $P_{\rm pre}$ and $P_{\rm cross}$ with $P_{\rm post}$, we are essentially adding in higher-order signal, thus naturally gaining information. Note that in order to match the information obtained by adding these two extra statistics, it is not enough to consider the bispectrum signal of the pre-reconstructed field, but both bispectrum and trispectrum signals. For this reason the information content of $P_{\rm pre} + B$ is different from that contained in $P_{\rm pre} + P_{\rm cross} + P_{\rm post}$. However, it is important to note that higher-order information that reconstruction brings is only a part from the total contained in the full bispectrum and trispectrum data-vectors. This is why a full analysis using $P + B + T$ will always provide more information. However, such an analysis is not very practical because of the size of the full data-vector and the computational time typically required to measure B and especially T directly. In this paper we show that $P_{\rm pre} + P_{\rm cross} + P_{\rm post}$ is an efficient alternative for extracting the relevant information from higher-order statistics for cosmological analyses.

To demonstrate the power of jointly using density fields before and after the reconstruction, we perform an anisotropic Lagrangian reconstruction (see Methods for details) on each realisation of the MOLINO galaxy mocks[17], which is a large suite of realistic galaxy mocks produced from the Quijote simulations[18] at $z = 0$. We then use these mocks to calculate the data covariance matrix and derivatives numerically for a Fisher matrix analysis[19] using the measured multipoles (up to $\ell = 4$) of $P_{\rm pre}$, $P_{\rm post}$ and $P_{\rm cross}$ on the parameter set $\Theta \equiv \{\Omega_{\rm m}, \Omega_{\rm b}, h, n_s, \sigma_8, M_\nu, {\bf H}\}$ where $\bf H$ denotes the Halo Occupation Distribution (HOD) parameters, i.e., ${\bf H} \equiv \{\log M_{\rm min}, \sigma_{\log M}, \log M_0, \alpha, \log M_1\}$[20] (see Methods for details).

Results and discussion

Panel **a** in Fig. 1 shows the measured power spectra monopole (the quadrupole and hexadecapole are shown in Supplementary Fig. 1), and we see that $P_{\rm cross}$ decreases dramatically with scale compared to $P_{\rm pre}$ and $P_{\rm post}$. This indicates a decorrelation between $P_{\rm pre}$ and $P_{\rm post}$ below quasi-nonlinear scales ($k \gtrsim 0.1\,{\rm hMpc}^{-1}$), which is largely due to the difference in infra-red effects contained in density fluctuations before and after the BAO reconstruction[21].

From the original data vector $\{P_{\rm pre}, P_{\rm post}, P_{\rm cross}\}$, we can construct their linear combinations, $P_{R\Delta}$, $P_{\Delta\Delta}$, $P_{\delta\Delta}$ defined as

$$P_{R\Delta} = P_{\rm cross} - P_{\rm post}\,; P_{\Delta\Delta} = P_{\rm pre} + P_{\rm post} - 2P_{\rm cross}\,; P_{\delta\Delta} = P_{\rm pre} - P_{\rm cross}.$$
(1)

Figure 1b shows these power spectra. As discussed above, these power spectra involving Δ contain the information of part of the high-order statistics such as bispectrum and trispectrum.

The derivatives of $\{P_{\rm pre}, P_{\rm post}, P_{\rm cross}\}$ with respect to cosmological parameters and HOD parameters are presented in Supplementary Figs. 2–4. We have checked and confirmed the convergence of our Fisher matrix result given the number of mocks available, demonstrating the robustness of our result (see Supplementary Note 2 and Supplementary Fig. 5 for details).

The correlation matrix for the monopole of power spectrum and bispectrum (only the correlation with the squeezed-limit of B_0 is visualised for brevity) is shown in Fig. 2. It is seen that $P_0^{\rm pre}$ highly correlates with B_0, confirming that the bispectrum is induced by nonlinearities. In contrast, $P_0^{\rm post}$ weakly correlates with B_0, or with $P_0^{\rm pre}$ and $P_0^{\rm cross}$ on nonlinear scales (e.g., at $k \gtrsim 0.2\,{\rm hMpc}^{-1}$). This, however, does not mean that $P_{\rm post}$ is irrelevant to the bispectrum—it actually is a mixture of $P_{\rm pre}$ and certain integrated forms of the bispectrum and trispectrum information[9,16]. Therefore by combining $P_{\rm post}$ with $P_{\rm pre}$ and $P_{\rm cross}$, one can in principle decouple the leading contribution in the power spectrum, bispectrum and trispectrum. The integrated form of the bispectrum information dominates $P_{\Delta\Delta}$, which strongly correlates with B_0, as shown in Supplementary Fig. 6 (see Supplementary Note 3). The fact that $P_{\rm post}$ barely correlates with B_0 implies that the information content in $P_{\rm post}$ combined with B_0 may be similar to that in $P_{\rm post}$ combined with $P_{\rm pre}$ and $P_{\rm cross}$, which is confirmed to be the case by the results from the Fisher analysis presented below.

The cumulative signal-to-noise ratio (SNR) for power spectrum multipoles (up to $\ell = 4$), bispectrum monopole and various data combinations are shown in Supplementary Fig. 7, in which we can see that the joint 2-point statistics, $P_{\rm all}$, is measured with a greater SNR than that of $P_{\rm post}$ or $P + B_0$, which may mean that $P_{\rm all}$ can be more informative than $P_{\rm post}$ or $P + B_0$ for constraining cosmological parameters.

To confirm the constraining power of $P_{\rm all}$, we then project the information content in the observables onto cosmological parameters using a

Fisher matrix approach. Contour plots for $(\log M_0, \sigma_8)$ derived from different datasets with two choices of k_{max} (the maximal k for the observables used in the analysis) are shown in Fig. 3 (more complete contour plots are shown in Supplementary Figs. 8–12). The smoothing scale is set to be $S = 10\,h^{-1}\mathrm{Mpc}$ when performing the reconstruction. The degeneracies between parameters using P_{pre}, P_{post} and P_{cross} are generally different, because P_{pre}, P_{post} and P_{cross} differ to a large extent in terms of nonlinearity on small scales. This easier to see in Supplementary Fig. 8, in which contours for the same parameters are shown for observables used in several k intervals. The contours derived from P_{pre} and P_{post} generally rotate as k
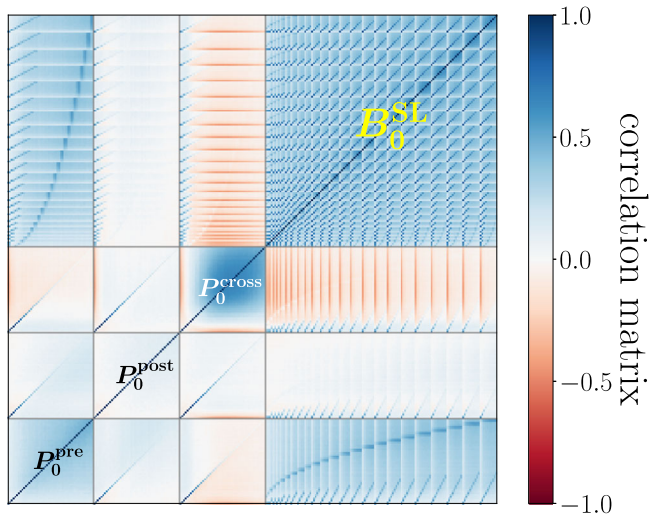

Fig. 2 | Part of the correlation matrix between the power spectra and bispectrum. The correlation matrix for the monopole of three types of power spectra $(P_0^{pre}, P_0^{post}, P_0^{cross})$, and of the bispectrum in the squeezed limit (B_0^{SL}), i.e. $k_1 = k_2 \gg k_3$, derived from the MOLINO galaxy mocks. The horizontal and vertical lines separate each block for visualisation. For all blocks, the associated k or k_1 increases from 0.01 to 0.5 hMpc^{-1}, from left to right and from bottom to top.
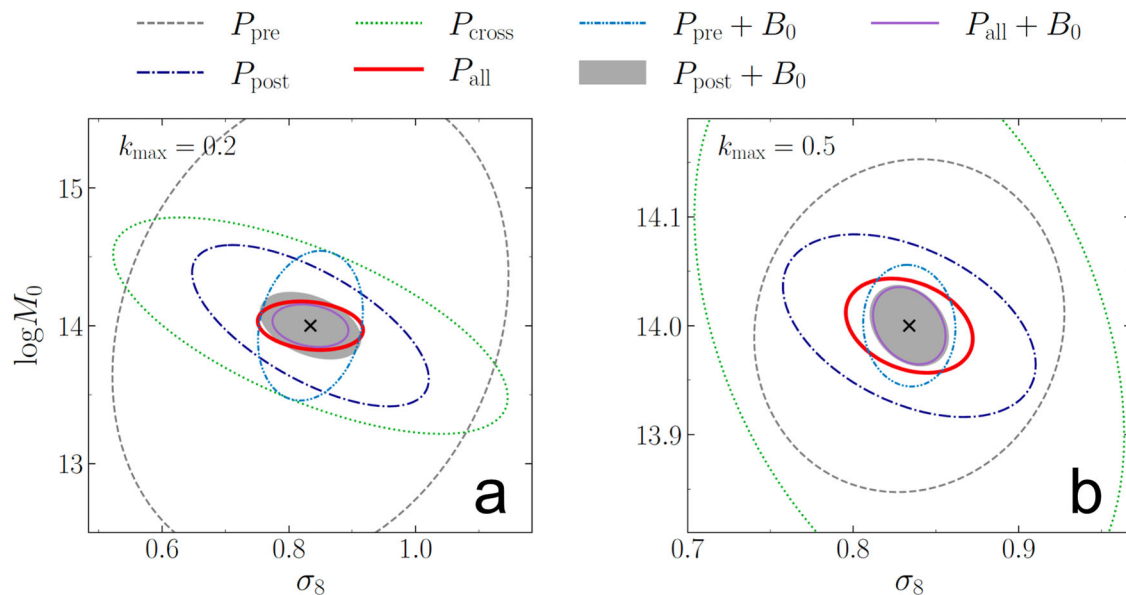
increases because of the kick-in of nonlinear effects, which affects P_{pre} and P_{post} at different levels on the same scale. This significantly improves the constraint when these power spectra are combined, labelled as P_{all}, which is tighter than that from the traditional joint power spectrum-bispectrum analysis ($P_{pre} + B_0$). It is found that P_{all} can even win against $P_{post} + B_0$ in some cases, demonstrating the robustness of this method. The contour plots with 1D posterior distributions for all parameters with $S = 10$ and $20\,h^{-1}$Mpc and $k_{max} = 0.2$ and 0.5 hMpc^{-1} are shown in Supplementary Figs. 9–12, respectively. In all cases, P_{all} offers competitive constraints on all parameters, even compared to the joint $P_{post} + B_0$ analysis.

To further quantify our results, in Fig. 4 we compare the square root of the Fisher matrix element for each parameter, with and without marginalising over others, derived from P_{all} and $P_{pre} + B_0$, respectively, with two choices of k_{max}.

For $k_{max} = 0.2\,h\,\mathrm{Mpc}^{-1}$, we see that the Fisher information for each parameter (panel a: without marginalising over others) derived from P_{all} is identical or even greater than that in $P_{pre} + B_0$. In other words, combining all power spectra we can efficiently extract the information in $P_{pre} + B_0$. After marginalising over other parameters, panel b shows that the uncertainty on each parameter gets redistributed due to the degeneracy. The ratios for the HOD parameters are all greater than unity especially for $\log M_0$ and $\sigma_{\log M}$, demonstrating the power of our method on constraining HOD parameters. The information content for cosmological parameters in $P_{pre} + B_0$ is well recovered by using P_{all}, although the recovery for M_ν is relatively worse. The overall trend for the case of $k_{max} = 0.5\,h\,\mathrm{Mpc}^{-1}$ is similar, although the advantage of using P_{all} over $P_{pre} + B_0$ gets degraded to some extent. However, P_{all} is still competitive: it almost fully recovers the information for the HOD parameters in $P_{pre} + B_0$ with or without marginalisation, and largely wins against $P_{pre} + B_0$ after marginalisation. Regarding the cosmological parameters, P_{all} recovers all information in $P_{pre} + B_0$ before the marginalisation, although the recovery is slightly worse for M_ν. After marginalisation when the uncertainties are redistributed, the constraint from P_{all} is generally worse than $P_{pre} + B_0$, especially for M_ν.

The 68% confidence level constraints on each parameter fitting to various datasets are shown in Table 1. To quantify the information gain, we evaluate the Figure-of-Merit (FoM) defined as $[\mathbf{det}(F)]^{1/(2N_p)}$, where F

Fig. 3 | The 68% confidence level contour plots on σ_8 and $\log M_0$ derived from various data combinations. a, b Show the constraint with $k_{max} = 0.2$ and 0.5 hMpc^{-1}, respectively. In each panel, the constraints are from the pre-reconstructed power spectrum (P_{pre}) alone (grey dashed line), post-reconstructed power spectrum (P_{post}) alone (dark blue dash-dotted line), cross power spectrum between the pre- and post-reconstructed density fields (P_{cross}) alone (green dotted

line), the combination of pre-, post-reconstructed and cross-power spectra (P_{all}) (red solid line), the combination of pre-reconstructed power spectrum and bispectrum ($P_{pre} + B_0$) (light blue dash-dot-dotted line), the combination of post-reconstructed power spectrum and bispectrum ($P_{post} + B_0$) (grey filled region) and the combination of P_{all} and bispectrum ($P_{all} + B_0$) (purple solid line).

denotes the Fisher matrix and N_P is the total number of free parameters. For the ease of comparison, for cases with different k_{max}, we normalise all the quantities using the corresponding one for P_{pre}. As shown, for $k_{max} = 0.2\,h\,Mpc^{-1}$, $(FoM)_{P_{all}}$ is greater than all others, namely, it is larger than $(FoM)_{P_{pre}}$ and $(FoM)_{P_{post}}$ by a factor of 2.7 and 1.7, respectively and it is even greater than $(FoM)_{P_{post}+B_0}$ by ~13%. For $k_{max} = 0.5\,h\,Mpc^{-1}$, $(FoM)_{P_{all}}$ is also more informative than $(FoM)_{P_{pre}}$ and $(FoM)_{P_{post}}$ by a factor of 2.1 and 1.5, respectively and is the same as $(FoM)_{P_{pre}+B_0}$, but is less than $(FoM)_{P_{post}+B_0}$ by ~10% in this case.

To highlight the constraining power on cosmological parameters, we also list FoM_{cos}, which is the FoM with all HOD parameters fixed. It shows a similar trend as $FoM_\odot$: P_{all} is the most informative data combination for $k_{max} = 0.2\,h\,Mpc^{-1}$, but it is outnumbered by $P_{pre} + B_0$ and $P_{post} + B_0$ by 13% and 30%, respectively, for the case of $k_{max} = 0.5\,h\,Mpc^{-1}$.

Conclusions

As demonstrated in this analysis, a joint analysis using P_{pre}, P_{post} and P_{cross} is an efficient way to extract high-order information from galaxy catalogues, and in some cases, P_{all} is more informative even than $P_{post} + B_0$, which is computationally much more expensive.

In this example, the k-binning for P and B are different, namely, $\Delta k(B) = 3k_f \sim 0.019\,hMpc^{-1} \sim 1.9\Delta k(P)$ where k_f denotes the fundamental k mode given the box size of the simulation. We have checked that using a finer k-binning for B only improves the constraints marginally[22], namely, the FoM can only be raised by ~10% when $\Delta k(B)$ is reduced from $3k_f$ to k_f, which is largely due to the strong mode-coupling in B as shown in Fig. 2. Such a fine binning is not practical anyway as, for example, using $\Delta k(B) = k_f$ up to $k = 0.5\,hMpc^{-1}$, we end up with more than 50,000 data points to measure for B_0.

Note that the MOLINO mock is produced at $z = 0$, where the nonlinear effects are the strongest. At higher redshifts, the density fields are more linear and Gaussian, thus we may expect less gain from our method. This can be
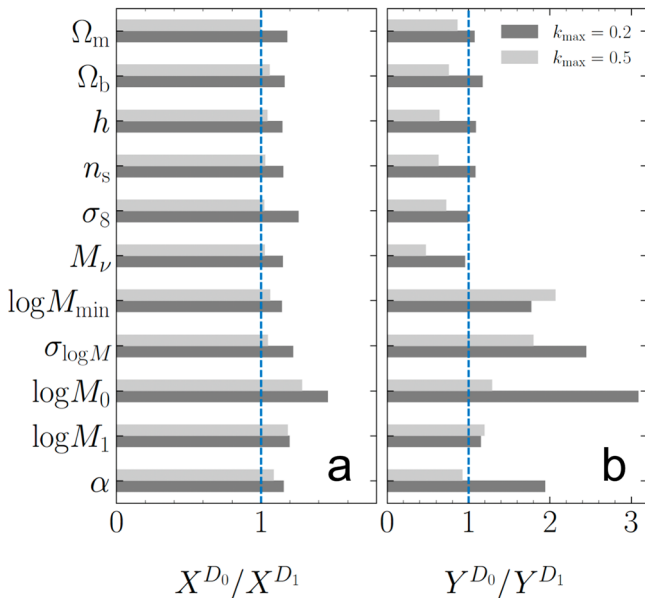

Fig. 4 | The Figure of Merit (FoM) of each individual parameter derived from all power spectrum combined, rescaled by those derived from $P_{pre} + B_0$. The quantities X in (**a**) and Y in (**b**) are defined as the FoM of each individual parameter with or without all other parameters fixed. Specifically, $X \equiv \sqrt{F_{ii}}$ and $Y \equiv 1/\sqrt{C_{ii}}$ where F is the Fisher matrix and $C \equiv F^{-1}$. D_0 and D_1 denote P_{all} and $P_{pre} + B_0$ respectively. The dark and light grey bars in each panel show the cases with $k = 0.2$ and $0.5\,hMpc^{-1}$, respectively. The vertical dashed lines show a full recovery of information from dataset $P_{pre} + B_0$. The smoothing scale is set to be $10\,h^{-1}Mpc$ when performing the reconstruction.

Table 1 | A quantification of the information content extracted from various observables measured from the MOLINO mocks

	$k_{max} = 0.2\,h\,Mpc^{-1}$							$k_{max} = 0.5\,h\,Mpc^{-1}$						
	P_{pre}	P_{post}	P_{all}	P_{all}^{SN}	$P_{pre} + B_0$	$P_{post} + B_0$	$P_{all} + B_0$	P_{pre}	P_{post}	P_{all}	P_{all}^{SN}	$P_{pre} + B_0$	$P_{post} + B_0$	$P_{all} + B_0$
$\sigma(\Omega_m)$	1	0.66	0.41	0.44	0.44	0.39	0.32	1	0.86	0.51	0.52	0.44	0.37	0.33
$\sigma(\Omega_b)$	1	0.69	0.44	0.47	0.52	0.44	0.37	1	0.76	0.50	0.51	0.38	0.33	0.31
$\sigma(h)$	1	0.75	0.43	0.47	0.47	0.43	0.34	1	0.86	0.48	0.49	0.31	0.28	0.26
$\sigma(n_s)$	1	0.79	0.41	0.44	0.44	0.42	0.33	1	0.92	0.45	0.45	0.28	0.27	0.25
$\sigma(\sigma_8)$	1	0.60	0.27	0.28	0.26	0.25	0.19	1	0.81	0.41	0.41	0.30	0.25	0.23
$\sigma(M_\nu[eV])$	1	0.88	0.36	0.48	0.34	0.33	0.25	1	0.84	0.46	0.50	0.22	0.22	0.20
FoM_{cos}	1	1.6	2.5	2.6	2.0	2.4	2.9	1	1.6	2.3	2.4	2.6	3.0	3.2
$\sigma(\log M_{min})$	1	0.66	0.21	0.10	0.38	0.39	0.14	1	0.54	0.19	0.14	0.39	0.31	0.13
$\sigma(\sigma_{\log M})$	1	0.59	0.18	0.16	0.43	0.38	0.14	1	0.50	0.16	0.17	0.30	0.21	0.12
$\sigma(\log M_0)$	1	0.33	0.10	0.10	0.31	0.14	0.09	1	0.55	0.28	0.28	0.37	0.24	0.23
$\sigma(\log M_1)$	1	0.89	0.27	0.15	0.31	0.36	0.22	1	0.96	0.43	0.29	0.51	0.48	0.35
$\sigma(\alpha)$	1	0.54	0.14	0.12	0.27	0.20	0.12	1	0.94	0.41	0.28	0.38	0.34	0.32
$FoM_\odot$	1	1.6	2.7	3.4	1.9	2.4	3.2	1	1.4	2.1	2.4	2.0	2.3	2.8

The 68% confidence level uncertainty of parameters derived from various datasets, with two different choices of the maximal wavenumber k_{max}. The column P_{all}^{SN} shows the result derived from P_{all} with the shot-noise component of P_{cross} kept. The last row, $FoM_\odot$ shows the Figure-of-Merit (FoM), defined as $|\det(F)|^{1/(2N_p)}$, where F denotes the Fisher matrix and N_p is the total number of free parameters. The row of FoM_{cos} shows the same FoM as defined above, but with all HOD parameters fixed. All quantities are normalised by that of the P_{pre} result.

seen from panel a of Supplementary Fig. 13, in which the correlation between P_{pre} and P_{post} at various redshifts is shown. As expected, P_{pre} and P_{post} are more correlated at higher redshifts, e.g., the correlation approaches 0.95 at $z = 5$ around $k \sim 0.3 \, h\mathrm{Mpc}^{-1}$, which implies that almost no information gain can be obtained at such high redshifts. As argued previously, the decorrelation at lower z is due to the fact that P_{pre} and P_{post} contain different levels of nonlinearity, as illustrated in panel **b**, thus are complementary. Also, the Alcock-Paczyński (AP) effect[23], which is a geometric distortion due to the discrepancy between the true cosmology and the fiducial one used to convert redshifts to distances, is irrelevant at $z = 0$[24]. As studied[25], the AP effect can make the small-scale bispectrum more informative for constraining the standard ruler than the power spectrum (P_{pre}), thus it is worth revisiting the case in which P_{post} and P_{cross} are added to the analysis.

To further demonstrate the efficacy of our method, we perform another analysis at a higher redshift using P and B_0 from an independent set of mocks: a suite of 4000 high-resolution N-body mocks (512^3 particles in a box with $500 \, h^{-1}\mathrm{Mpc}$ a side) produced at $z = 1.02$. This allows us to include the AP effect when performing the BAO and RSD analysis. This test confirms that P_{pre}, P_{post} and P_{cross} are complementary for constraining cosmological parameters, and that P_{all} contains almost all the information in P combined with B_0, which is consistent with our findings from the MOLINO analysis (see Methods and Supplementary Figs. 16–19 for more details).

Stage-IV redshift surveys including the Dark Energy Spectroscopic Instrument (DESI)[26], Euclid[27] and the Prime Focus Spectrograph[28] will release galaxy maps over a wide range of redshifts with an exquisite precision. As long as the distribution of a tracer in a given redshift range is not too sparse, namely, the number density is not lower than $10^{-4} \, h^3\mathrm{Mpc}^{-3}$ so that a reconstruction can be efficiently performed[29], the method presented in this work can be directly applied to extract high order statistics for constraining cosmological parameters from 2-point measurements, which is computationally much more efficient to perform. Since the reconstruction will be performed anyway for most ongoing and forthcoming galaxy surveys to improve the BAO signal, our proposed analysis can be performed at almost no additional computational cost.

Additional work is required to build a link between cosmological parameters to the full shape of power spectra for a likelihood analysis, and this is challenging using perturbation-theory-based models on (quasi-) nonlinear scales, especially for the reconstructed power spectrum and the cross power spectrum. However, model-free approaches including the simulation-based emulation[30–32], can be used for performing the P_{all} analysis down to nonlinear scales, in order to extract the cosmological information from the power spectra to the greatest extent. The emulator-based P_{all} analysis was recently performed and validated[33], which well demonstrates the idea proposed in this work.

Methods
The mock catalogues—MOLINO galaxy mocks at $z = 0$
The MOLINO catalogues[17] are a suite of publicly available galaxy mock catalogues that were constructed to quantify the total cosmological information content of different galaxy clustering observables using Fisher matrix forecasting. They are constructed from the QUIJOTE suite of N-body simulations[18] using the halo occupation distribution (HOD) framework. HOD provides a statistical prescription for populating dark matter halos with central and satellite galaxies and has been successful in reproducing a wide range of observed galaxy clustering statistics. In particular, the MOLINO catalogues use the standard HOD model[20], which has five free parameters: $\{ \log M_{min}, \sigma_{\log M}, \log M_0, \alpha, \log M_1 \}$. MOLINO includes 15,000 galaxy catalogues that are constructed at a fiducial set of cosmological parameters ($\Omega_m = 0.3175, \Omega_b = 0.049, h = 0.6711, n_s = 0.9624, \sigma_8 = 0.834, M_\nu = 0$) and HOD parameters ($\log M_{min} = 13.65, \sigma_{\log M} = 0.2, \log M_0 = 14.0, \alpha = 1.1, \log M_1 = 14.0$), which are based on the best-fit HOD parameters for the SDSS $M_r < -21.5$ and -22 samples[20].

The 15,000 MOLINO mocks for the fiducial cosmology are designed for accurately estimating the covariance matrices of the galaxy clustering observables, including the power spectra and bispectra. In addition, a

separate set of the MOLINO mocks are produced for estimating the derivatives with respect to cosmological parameters (including the HOD ones) using the finite difference method (see Supplementary Note 2 for details). For this purpose, 60,000 galaxy mocks are constructed at 24 cosmologies that are slightly different from the fiducial one[17].

Since the data covariance matrices and the derivatives are all evaluated numerically using mocks, it is important to ensure that the result derived from the Fisher matrix approach is robust against numerical issues, as argued in[17,34–36]. We, therefore, perform numerical tests to check the dependence of our Fisher matrix calculation on the number of mocks and find that the marginalised uncertainties of all the concerning parameters are well converged given the number of mocks available. The details are presented in the Supplementary Note 2.

The mock catalogues—4000 high-resolution N-body mocks at $z = 1.02$
To confirm our findings from the MOLINO mocks, we perform an independent mock test on a suite of 4000 high-resolution N-body simulations with 512^3 dark matter particles in a $L = 500 \, h^{-1}\mathrm{Mpc}$ box at $z = 1.02$[13]. The fiducial cosmology used for this set of mocks is consistent with the Planck 2015[37] observations.

The mock catalogues—COLA mocks at multiple redshifts
To investigate how the decorrelation between P_{pre} and P_{post} varies with redshifts, we perform another set of N-body simulations using the COmoving Lagrangian Acceleration (COLA)[38] method with the MG-PICOLA code[39]. The mocks are performed using 256^3 dark matter particles in a $L = 256 \, h^{-1}\mathrm{Mpc}$ box, and snapshots at $z = 0, 1, 2, 3, 5, 10, 15$ are analysed, to cover a sufficiently wide range of redshifts. Although the COLA mocks are approximate, the accuracy and reliability has been well demonstrated in the literature[38–40].

The reconstruction process
An anisotropic reconstruction[41] is performed on each realisation of the MOLINO galaxy mocks with two choices of the smoothing scale, $S = 10$ and $20 \, h^{-1}\mathrm{Mpc}$ (All results presented in the main text are for the $10 \, h^{-1}\mathrm{Mpc}$ case, while results for $20 \, h^{-1}\mathrm{Mpc}$ are shown in the Supplementary information). Specifically, a smoothing is performed by convolving density field with the kernel $K(k) = \exp\left[-(kS)^2/2\right]$ in Fourier space. Note that in this procedure the information on scales below the smoothing scales gets erased, and there are studies on choosing the proper smoothing scale[15]. In principle, the smoothing scale can be made sufficiently small to restore more information, for example, no smoothing is needed at all in the nonlinear reconstruction methods[14] and we will apply our pipeline to those reconstruction schemes for further investigation. After the smoothing, the displacement vector is solved using the Zeldovich approximation, i.e., $\tilde{\mathbf{s}}(\mathbf{k}) = -\frac{i\mathbf{k}}{k^2} \frac{\delta(\mathbf{k})}{b_{in} + f_{in}\mu^2} K(k)$, where δ denotes the nonlinear redshift-space overdensity, b_{in} and f_{in} are the *input* linear bias and the logarithmic growth rate for the density field, respectively. Note that $\{b_{in}, f_{in}\}$ does not have to be identical to the true underlying $\{b, f\}$ of the density field, thus they are *not* free parameters to be determined. The post-reconstructed power spectrum for a given $\{b_{in}, f_{in}\}$ can be modelled using either the perturbation theory[42], or an emulation approach, as developed in ref. 33 It is true that an inappropriate choice of $\{b_{in}, f_{in}\}$, e.g., a set of $\{b_{in}, f_{in}\}$ that is significantly different from the truth, may affect the efficiency of the BAO reconstruction, but the impact from using $\{b_{in}, f_{in}\}$ can be well modelled and corrected for, so this process is not expected to generate bias or uncertainties.

To demonstrate that the result would not get biased by an inappropriate set of $\{b_{in}, f_{in}\}$, ref. 33 uses a significantly wrong set of $\{b_{in}, f_{in}\}$ for the reconstruction, namely, $\{b_{in} = 0.9b, f_{in} = 0.7f\}$, where $\{b, f\}$ are the true b and f of the density field. This level of deviation from the true value is greater than 3σ level, given the uncertainty of b and f constrained by the BOSS (DR12) survey[43]. The impact of using such a wrong set of $\{b_{in}, f_{in}\}$ is corrected for by the properly trained emulator, and as demonstrated in Fig. 6 of

ref. 33, using this set of $\{b_{in}, f_{in}\}$ does not bias, or dilute the final parameter constraint. In summary, it is expected that the choice of $\{b_{in}, f_{in}\}$ used in this work does not bias the result and a more in-depth assessment on the potential influence of the choice on $\{b_{in}, f_{in}\}$ is left for a future study on a joint P_{all} analysis using the actual observational data.

An inverse Fourier transformation on $\tilde{s}$ returns the configuration-space displacement field $s(x)$, which is used to move both the galaxies and randoms. We also perform the anisotropic Lagrangian reconstruction[15] on each realisation of the N-body mocks, but only with a smoothing scale $S = 10\,h^{-1}$Mpc.

Note that the information content in the reconstructed power spectrum is the same no matter whether the RSD is kept or not during the reconstruction process, and we have numerically confirmed this by performing the analysis with the isotropic reconstruction[15], in which the RSD is removed using the fiducial f and b used for producing the mocks.

Also note that the BAO reconstruction procedure is not always required for extracting geometric information in the galaxy clustering. For example, when using the information in the linear point[44–47], no reconstruction is required. Also, the estimated α from the traditional BAO methods and from the linear point approach may conceptually differ and a comparison is beyond the scope of this work.

Measurement of the power spectrum multipoles
The multipoles (up to $\ell = 4$) of both the pre- and post-reconstructed density fields are measured using an FFT-based estimator[48] implemented in N-body kit[49]. The shot-noise, which reflects the discreteness of the density field, is removed as a constant for the monopole of the auto-power. The k-binning is $\Delta k = 0.01\,$hMpc^{-1} for both the MOLINO and N-body mocks.

Care needs to be taken when measuring the cross-power spectrum between the pre-and post-reconstructed density fields since the raw measurement using the FFT-based estimator is contaminated by a scale-dependent shot-noise: on large scales, the post-reconstructed field resembles the unreconstructed one, making the cross-power spectrum essentially an auto-power, thus it is subject to a shot-noise component. On small scales, however, the shot-noise largely drops because the two fields effectively decorrelate.

To obtain a measured cross-power spectrum whose mean value reflects the true power spectrum in the data such that no subtraction of the noise component is required, we adopt the half-sum and half-difference (HS-HD) approach[50]. We start by randomly dividing the catalogue into two halves, dubbed δ_1 and δ_2 and the corresponding reconstructed density fields are R_1 and R_2, respectively.

Let

$$\text{HS} \equiv \frac{\delta_1 + \delta_2}{2}, \quad \text{HD} \equiv \frac{\delta_1 - \delta_2}{2}, \tag{2}$$

and

$$\text{HS}^R \equiv \frac{R_1 + R_2}{2}, \quad \text{HD}^R \equiv \frac{R_1 - R_2}{2}. \tag{3}$$

Then HS$^{(R)}$ contains both the signal and noise, but HD$^{(R)}$ only contains the noise. Hence the cross-power spectrum estimator is,

$$\hat{P}_{cross} = \langle \text{HS}, \text{HS}^R \rangle - \langle \text{HD}, \text{HD}^R \rangle = \frac{\langle \delta_1, R_2 \rangle + \langle \delta_2, R_1 \rangle}{2}. \tag{4}$$

The scatter of $\hat{P}_{cross}$ around the mean value allows for an estimation of the covariance matrix, which is a 4-point function[51], shown in Fig. 2. By comparing $\hat{P}_{cross}$ with that measured without splitting the samples, we can obtain the noise power spectrum, as shown in Supplementary Figs. 14, 15 (for cases with $S = 10$ and $20\,h^{-1}$Mpc, respectively), which is apparently scale-dependent. The noise is anisotropic, and thus it affects even for multipoles with $\ell \neq 0$.

Since a change in HOD parameters can result in a change in the number density of the galaxy sample and thus affect the shot-noise, the shot-noise can in principle be used to constrain the HOD parameters. We, therefore, perform an additional Fisher projection with the shot-noise kept in the spectra, and find that the constraints on HOD parameters can be improved in general, but the constraint on cosmological parameters is largely unchanged (see the P_{all}^{SN} column in Table 1).

Measurement of the bispectrum monopole
We measure the galaxy bispectrum monopole, B_0, for all of the mock catalogues using the publicly available pySpectrum package[17,34]. Galaxy positions are first interpolated onto a grid using a fourth-order interpolation scheme and then Fourier transformed to obtain $\delta(k)$. Afterwards B_0 is estimated using

$$B_0(k_1, k_2, k_3) = \frac{1}{V_B} \int_{k_1} d^3 q_1 \int_{k_2} d^3 q_2 \int_{k_3} d^3 q_3 \, \delta_D(q_{123}) \, \delta(q_1) \, \delta(q_2) \, \delta(q_3)$$
$$- B_0^{SN} \tag{5}$$

where δ_D is the Dirac delta function, V_B is the normalisation factor proportional to the number of triplets that can be found in the k_1, k_2, k_3 triangle bin and B_0^{SN} is the Poisson shot noise correction term. Triangle configurations are defined by k_1, k_2, k_3 and for the MOLINO mocks, the width of the bins is $\Delta k = 3k_f$, where $k_f = 2\pi/(1000\,h^{-1}$Mpc$)$ and for the N-body mocks, $\Delta k = 0.02\,$hMpc^{-1}.

An AP test performed on the MOLINO mocks
Although the AP effect plays no role for the MOLINO mock since it is produced at $z = 0$[24], we perform a test by isotropically stretching the scales and angles using pairs of AP parameters calculated at a non-zero redshift. This gives us an idea about whether this artificial and exaggerated AP effect can change the main conclusion of this work that the cosmological information content in P_{all} is almost the same as or more than that in $P_{pre} + B_0$. In practice, we use the $(\alpha_{\parallel}$ and $\alpha_{\perp})$ pairs computed at $z_{eff} = 0.5$ and 1.0 respectively to stretch the wave numbers along and across the line of sight directions and repeat the analysis. As shown in the Supplementary Table (see Supplementary Note 3), this added 'artificial' AP effects can generally tighten the constraint, but the relative constraints from P_{all} and $P_{pre} + B_0$ are largely unchanged, meaning that the main conclusion of this paper remains the same if the AP effect is taken into account.

An AP test on the N-body mocks
We perform an additional Fisher matrix analysis[19] on the AP parameters using 4000 realisations of N-body particle mocks produced at $z = 1.02$ in redshift space. Part of the observables (the power spectrum monopole) are shown in Supplementary Fig. 16. From panel a we see that the amplitude of P_{cross} decreases dramatically with scales, indicating a decorrelation between P_{pre} and P_{post} below quasi-nonlinear scales, which is confirmed by the correlation coefficient (the normalised covariance) plotted in panel b. This decorrelation, which is not caused by the shot noise given the negligible noise level in the mocks, is a clear evidence of the complementarity among the power spectra.

The cumulative signal-to-noise ratio (SNR) is shown in panel a of Supplementary Fig. 17, in which we see that P_{all} is more informative than P_{pre} and that $P + B_0$ has slightly higher SNR on small scales.

We first perform an AP test on the isotropic dilation parameter α_{iso}, which is defined as the ratio of the true spherically-averaged scale of the standard ruler to the fiducial one. This dilation parameter depends on cosmological parameters, and can be constrained using the monopole of the power spectrum and bispectrum. The wavenumber k gets dilated by α_{iso} due to the AP effect, thus the

observables are,

$$P_0^T(k') = A_0 \frac{1}{\alpha_{\rm iso}^3} P_0^T(k/\alpha_{\rm iso}) \tag{6}$$

$$B_0(k_1', k_2', k_3') = A_B \frac{1}{\alpha_{\rm iso}^6} B_0(k_1/\alpha_{\rm iso}, k_2/\alpha_{\rm iso}, k_3/\alpha_{\rm iso}), \tag{7}$$

where T denotes the type of P_0, namely, $T = \{\text{pre, post, cross}\}$, and the parameters A_0 and A_B are used to parameterise the overall amplitudes of power spectrum monopole and bispectrum monopole, respectively. Since the purpose of this test is to study the impact of AP parameters, the relevant parameters are $\{\alpha_{\rm iso}, \ln A_0, \ln A_B\}$, and these are free parameters in this calculation. Other parameters are held fixed to avoid confusion. The derivative with respect to $\alpha_{\rm iso}$ is evaluated semi-analytically as

$$\frac{\partial P_0^T}{\partial \alpha} = -3P_0^T - \frac{dP_0^T}{d\ln k}, \tag{8}$$

$$\frac{\partial B_0}{\partial \alpha_{\rm iso}} = -6B_0 - \left(\frac{\partial B_0}{\partial \ln k_1} + \frac{\partial B_0}{\partial \ln k_2} + \frac{\partial B_0}{\partial \ln k_3}\right). \tag{9}$$

Then the constraint on $\alpha_{\rm iso}$ is derived after marginalising over the amplitudes A_0 and A_B, and it is shown in panel b of Supplementary Fig. 17. The FoM of $\alpha_{\rm iso}$ shows up step-like features due to the BAO feature, as previously discovered[25] and $P_{\rm all}$ offers the greatest FoM, until overtaken by $P + B_0$ at $k_{\rm max} \gtrsim 0.37\, h{\rm Mpc}^{-1}$.

We use the first three even multipole moments to assemble the two-dimensional power spectrum, i.e.,

$$P(k, \mu) \simeq \sum_{\ell=0,2,4} P_\ell(k) \mathcal{L}_\ell(\mu), \tag{10}$$

The bispectrum is similarly assembled using the first three even multipoles with $m = 0$[52], which are the most informative ones[53], i.e.,

$$B(k_1, k_2, k_3, \mu_1, \mu_2) \simeq \sum_{\ell=0,2,4} B_{\ell,m=0}(k_1, k_2, k_3) Y_{\ell,m=0}(\theta, \phi). \tag{11}$$

The wave-number k_i and the cosine of the angle to the line-of-sight μ_i are stretched by two dilation parameters $\alpha_\perp$ and $\alpha_\parallel$ due to the AP effect[54,55],

$$q_i = \frac{k_i}{\alpha_\perp}\left[1 + \mu_i^2\left(\frac{1}{F^2} - 1\right)\right]^{1/2}, \quad \nu_i = \frac{\mu_i}{F}\left[1 + \mu_i^2\left(\frac{1}{F^2} - 1\right)\right]^{-1/2}, \quad F = \alpha_\parallel/\alpha_\perp. \tag{12}$$

The power spectrum multipoles (the index for the type is omitted for brevity) and bispectrum monopole including the AP effect are respectively given as,

$$P_\ell(k) = \frac{(2\ell + 1)}{2\alpha_\perp^2\, \alpha_\parallel} \int_{-1}^{1} d\mu P(q, \nu) \mathcal{L}_\ell(\mu), \tag{13}$$

$$B_0(k_1, k_2, k_3) = \frac{1}{4\pi\alpha_\perp^4 \alpha_\parallel^2} \int_{-1}^{1} d\mu_1 \int_{0}^{2\pi} d\phi B(q_1, q_2, q_3, \nu_1, \nu_2). \tag{14}$$

The free parameters are $\{\alpha_\perp, \alpha_\parallel, \ln A_\ell, \ln A_B\}$, where $A_\ell (\ell = 0, 2, 4)$ denotes the overall amplitudes of the power spectrum multipoles, and A_B is the amplitude of the bispectrum monopole. The derivatives with respect to the parameters $\alpha_\perp$ and $\alpha_\parallel$ are evaluated numerically by

$$\frac{\partial \boldsymbol{O}}{\partial \alpha_i} = \frac{\boldsymbol{O}(\alpha_i^+) - \boldsymbol{O}(\alpha_i^-)}{2\Delta\alpha_i}, \tag{15}$$

where $\boldsymbol{O} \in \{P_\ell, B_0\}$ denotes the observables, and the step size $\Delta\alpha_i = 0.01$. Then the constraints on $\alpha_\perp$ and $\alpha_\parallel$ are derived after marginalising over the amplitudes A_ℓ and A_B.

The FoM for $\alpha_\perp, \alpha_\parallel$ is shown in panel c of Supplementary Fig. 17, and it shows a similar trend as FoM($\alpha_{\rm iso}$). The contour plot for $\alpha_\perp, \alpha_\parallel$ with $k_{\rm max} = 0.4\, h{\rm Mpc}^{-1}$ is shown in Supplementary Fig. 18, further highlighting the strong constraining power of $P_{\rm all}$ in comparison to that of $P + B_0$.

A joint BAO and RSD analysis on the *N*-body mocks

In addition to $\alpha_\perp, \alpha_\parallel$, we add one more parameter to the analysis, which is $\Delta\nu$, the parameter describing the change of velocities along the line of sight. This parameter mimics the change of the linear growth rate on large scales, but it also changes the velocity of particles coherently on small scales. We compute the derivatives with respect to $\Delta\nu$ numerically.

The projection onto the parameters, shown in Supplementary Fig. 19, demonstrates the advantage of performing a joint analysis using $P_{\rm pre}, P_{\rm post}$ and $P_{\rm cross}$. On large scales, $P_{\rm pre}$ and $P_{\rm post}$ are both determined by the linear density field, making the power spectra highly correlated. As shown in panels c_1 and c_5, the contours derived from $P_{\rm pre}$ and $P_{\rm post}$ have similar orientations and we do not gain by combining them. For $k > 0.15\, h{\rm Mpc}^{-1}$, the correlation between $P_{\rm pre}$ and $P_{\rm post}$ decreases as the pre-reconstructed density field is dominated by the non-linear field while $P_{\rm post}$ still retains the correlation with the linear density field. The contours shown in lines in panel c_7, which are derived from power spectra in the k range of $[0.2, 0.25]\, h{\rm Mpc}^{-1}$, are almost orthogonal to each other, making the constraint from the combined spectra, as illustrated in the shaded region, significantly tightened. On smaller scales, the post-reconstructed density field is also dominated by the non-linear field and the orientations of the contours are again aligned and the complementarity on smaller scales weakens. This shows that the level of nonlinearity in the power spectrum determines the degeneracies between parameters. Since $P_{\rm pre}$ and $P_{\rm post}$ are affected by different levels of nonlinearities on a given scale, which gives rise to different degeneracies, a joint analysis using both $P_{\rm pre}$ and $P_{\rm post}$ (and $P_{\rm cross}$) can yield a better constraint by breaking the degeneracies.

Data availability

The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

References

1. Eisenstein, D. J. & Hu, W. Baryonic features in the matter transfer function. *Astrophys. J.* **496**, 605–614 (1998).
2. Riess, A. G. et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.* **116**, 1009–1038 (1998).
3. Perlmutter, S. et al. Measurements of Ω and Λ from 42 high-redshift supernovae. *Astrophys. J.* **517**, 565–586 (1999).
4. Dawson, K. S. et al. The Baryon Oscillation Spectroscopic Survey of SDSS-III. *Astron. J.* **145**, 10 (2013).
5. Alam, S. et al. Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: cosmological implications from two decades of spectroscopic surveys at the Apache Point Observatory. *Phys. Rev. D* **103**, 083533 (2021).
6. Eisenstein, D. J., Seo, H.-j, Sirko, E. & Spergel, D. Improving cosmological distance measurements by reconstruction of the baryon acoustic peak. *Astrophys. J.* **664**, 675–679 (2007).
7. Padmanabhan, N. et al. A 2 per cent distance to z = 0.35 by reconstructing baryon acoustic oscillations - I. Methods and application to the Sloan Digital Sky Survey. *Mon. Not. Roy. Astron. Soc.* **427**, 2132–2145 (2012).
8. Burden, A. et al. Efficient reconstruction of linear baryon acoustic oscillations in galaxy surveys. *Mon. Not. Roy. Astron. Soc.* **445**, 3152–3168 (2014).

9. Schmittfull, M., Feng, Y., Beutler, F., Sherwin, B. & Chu, M. Y. Eulerian BAO reconstructions and N -point statistics. *Phys. Rev. D* **92**, 123522 (2015).

10. Kaiser, N. Clustering in real space and in redshift space. *Mon. Not. Roy. Astron. Soc.* **227**, 1–27 (1987).

11. Lue, A., Scoccimarro, R. & Starkman, G. Differentiating between modified gravity and dark energy. *Phys. Rev. D* **69**, 044005 (2004).

12. Guzzo, L. et al. A test of the nature of cosmic acceleration using galaxy redshift distortions. *Nature* **451**, 541–545 (2008).

13. Hikage, C., Takahashi, R. & Koyama, K. Covariance of the redshift-space matter power spectrum after reconstruction. *Phys. Rev. D* **102**, 083514 (2020).

14. Zhu, H.-M., Yu, Y. & Pen, U.-L. Nonlinear reconstruction of redshift space distortions. *Phys. Rev. D* **97**, 043502 (2018).

15. Seo, H.-J., Beutler, F., Ross, A. J. & Saito, S. Modeling the reconstructed BAO in Fourier space. *Mon. Not. Roy. Astron. Soc.* **460**, 2453–2471 (2016).

16. Hikage, C., Koyama, K. & Takahashi, R. Perturbation theory for the redshift-space matter power spectra after reconstruction. *Phys. Rev. D* **101**, 043510 (2020).

17. Hahn, C. & Villaescusa-Navarro, F. Constraining M_ν with the bispectrum. Part II. The information content of the galaxy bispectrum monopole. *J.Cosmol. Astropart. Phys.* **2021**, 029 (2021).

18. Villaescusa-Navarro, F. et al. The Quijote Simulations. *Astrophys. J., Suppl. Ser.* **250**, 2 (2020).

19. Tegmark, M., Taylor, A. & Heavens, A. Karhunen-Loeve eigenvalue problems in cosmology: how should we tackle large data sets? *Astrophys. J.* **480**, 22 (1997).

20. Zheng, Z., Coil, A. L. & Zehavi, I. Galaxy evolution from halo occupation distribution modeling of DEEP2 and SDSS galaxy clustering. *Astrophys. J.* **667**, 760–779 (2007).

21. Sugiyama, N. Developing a Theoretical Model for the Resummation of Infrared Effects in the Post-Reconstruction Power Spectrum 2402.06142 (2024).

22. Yankelevich, V. & Porciani, C. Cosmological information in the redshift-space bispectrum. *Mon. Not. Roy. Astron. Soc.* **483**, 2078–2099 (2019).

23. Alcock, C. & Paczynski, B. An evolution free test for non-zero cosmological constant. *Nature* **281**, 358–359 (1979).

24. d'Amico, G. et al. The cosmological analysis of the SDSS/BOSS data from the effective field theory of large-scale structure. *J.Cosmol. Astropart. Phys* **2020**, 005 (2020).

25. Samushia, L., Slepian, Z. & Villaescusa-Navarro, F. Information content of higher order galaxy correlation functions. *Mon. Not. Roy. Astron. Soc.* **505**, 628–641 (2021).

26. DESI Collaboration. The DESI Experiment Part I: Science,Targeting, and Survey Design. *ArXiv e-prints* 1611.00036 (2016).

27. Laureijs, R. et al. Euclid Definition Study Report 1110.3193 (2011).

28. Ellis, R. et al. Extragalactic science, cosmology, and Galactic archaeology with the Subaru Prime Focus Spectrograph. *Publ. Astron. Soc. Jap.* **66**, R1 (2014).

29. White, M. Shot noise and reconstruction of the acoustic peak 1004.0250 (2010).

30. Lawrence, E. et al. The coyote universe. III. Simulation suite and precision emulator for the nonlinear matter power spectrum. *Astrophys. J.* **713**, 1322–1331 (2010).

31. Kobayashi, Y., Nishimichi, T., Takada, M., Takahashi, R. & Osato, K. Accurate emulator for the redshift-space power spectrum of dark matter halos and its application to galaxy power spectrum. *Phys. Rev. D* **102**, 063504 (2020).

32. Neveux, R. et al. Combined full shape analysis of BOSS galaxies and eBOSS quasars using an iterative emulator. *Mon. Not. R. Astron. Soc.* **516**, 1910–1922 (2022).

33. Wang, Y. et al. Emulating power spectra for pre- and post-reconstructed galaxy samples 2311.05848 (2023).

34. Hahn, C., Villaescusa-Navarro, F., Castorina, E. & Scoccimarro, R. Constraining M_ν with the bispectrum. Part I. Breaking parameter degeneracies. *J. Cosmol. Astropart. Phys.* **03**, 040 (2020).

35. Coulton, W. R. et al. Quijote PNG: the information content of the halo power spectrum and bispectrum. *Astrophys. J.* **943**,178 (2023).

36. Paillas, E. et al. Constraining $\nu\Lambda$CDM with density-split clustering. *Mon. Not. R. Astron. Soc.* **522**, 606–625. (2022).

37. Ade, P. A. R. et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.* **594**, A13 (2016).

38. Tassev, S., Zaldarriaga, M. & Eisenstein, D. Solving large scale structure in ten easy steps with COLA. *J. Cosmol. Astropart. Phys.* **06**, 036 (2013).

39. Winther, H. A., Koyama, K., Manera, M., Wright, B. S. & Zhao, G.-B. COLA with scale-dependent growth: applications to screened modified gravity models. *J. Cosmol. Astropart. Phys.* **08**, 006 (2017).

40. Howlett, C., Manera, M. & Percival, W. J. L-PICOLA: a parallel code for fast dark matter simulation. *Astron. Comput.* **12**, 109–126 (2015).

41. Chen, S.-F., Vlah, Z. & White, M. The reconstructed power spectrum in the Zeldovich approximation. *J. Cosmol. Astropart. Phys.* **09**, 017 (2019).

42. Sherwin, B. D. & White, M. The impact of wrong assumptions in BAO reconstruction. *J. Cosmol. Astropart. Phys.* **02**, 027 (2019).

43. Beutler, F. et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: anisotropic galaxy clustering in Fourier space. *Mon. Not. Roy. Astron. Soc.* **466**, 2242–2260 (2017).

44. Anselmi, S., Starkman, G. D., Corasaniti, P.-S., Sheth, R. K. & Zehavi, I. Galaxy correlation functions provide a more robust cosmological standard ruler. *Phys. Rev. Lett.* **121**, 021302 (2018).

45. Anselmi, S. et al. Cosmic distance inference from purely geometric BAO methods: Linear Point standard ruler and Correlation Function Model Fitting. *Phys. Rev. D* **99**, 123515 (2019).

46. O'Dwyer, M. et al. Linear Point and Sound Horizon as Purely Geometric standard rulers. *Phys. Rev. D* **101**, 083517 (2020).

47. Anselmi, S., Starkman, G. D. & Renzi, A. Cosmological forecasts for future galaxy surveys with the linear point standard ruler: toward consistent bao analyses far from a fiducial cosmology. *Phys. Rev. D* **107**, 123506 (2023).

48. Hand, N., Li, Y., Slepian, Z. & Seljak, U. An optimal FFT-based anisotropic power spectrum estimator. *J. Cosmol. Astropart. Phys.* **07**, 002 (2017).

49. Hand, N. et al. nbodykit: an open-source, massively parallel toolkit for large-scale structure. *Astron. J.* **156**, 160 (2018).

50. Ando, S., Benoit-Lévy, A. & Komatsu, E. Angular power spectrum of galaxies in the 2MASS Redshift Survey. *Mon. Not. Roy. Astron. Soc.* **473**, 4318–4325 (2018).

51. Sugiyama, N. S., Saito, S., Beutler, F. & Seo, H.-J. Perturbation theory approach to predict the covariance matrices of the galaxy power spectrum and bispectrum in redshift space. *Mon. Not. Roy. Astron. Soc.* **497**, 1684–1711 (2020).

52. Scoccimarro, R., Couchman, H. M. P. & Frieman, J. A. The bispectrum as a signature of gravitational instability in redshift-space. *Astrophys. J.* **517**, 531–540 (1999).

53. Gagrani, P. & Samushia, L. Information content of the angular multipoles of redshift-space galaxy bispectrum. *Mon. Not. Roy. Astron. Soc.* **467**, 928–935 (2017).

54. Ballinger, W. E., Peacock, J. A. & Heavens, A. F. Measuring the cosmological constant with redshift surveys. *Mon. Not. Roy. Astron. Soc.* **282**, 877–888 (1996).

55. Gil-Marín, H. et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: RSD measurement from the power spectrum and bispectrum of the DR12 BOSS galaxies. *Mon. Not. Roy. Astron. Soc.* **465**, 1757–1788 (2017).

Author contributions
Y.W. contributed to the idea and the development of the pipeline, performed the analysis on the MOLINO mocks, produced the results and contributed to the draft. G.B.Z. proposed the idea, developed the pipeline, performed the analysis on the N-body mocks and wrote the draft. K.K. contributed to the idea and pipeline, led the effort on the theoretical interpretation of the result and co-wrote the draft. W.J.P. contributed to the idea, theoretical interpretation and observational implications of this work, and co-wrote the draft. R.T. and C.H. provided the N-body mocks, performed the reconstruction and measured the power spectra and bispectrum. H.G.M. and C.H.H. contributed to the bispectrum analysis and to the draft. R.Z. built a theoretical model for the k-dependent noise for the cross-power spectrum, to confirm the measurement using the HSHD approach. W.Z., R.Z. and X.M. contributed to building an EFT-based theoretical model for the cross-power, which helped with interpretation of the result. Y.Y. and H.M.Z. provided another set of N-body mocks for a cross validation, and helped with interpreting the de-correlation effect. F.G. performed a test on the eBOSS mocks to confirm the de-correlation effect between P_{pre} and P_{post}.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42005-024-01624-7.

Correspondence and requests for materials should be addressed to Gong-Bo Zhao, Kazuya Koyama or Will J. Percival.

Peer review information *Communications Physics* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.