

Minkowski functionals of large-scale structure as a probe of modified gravityAoxiang Jiang^{Ⓧ, †}, Wei Liu, and Wenjuan Fang^{*}*CAS Key Laboratory for Research in Galaxies and Cosmology, Department of Astronomy,
University of Science and Technology of China, Hefei, Anhui 230026, China
and School of Astronomy and Space Sciences, University of Science and
Technology of China, Hefei, Anhui 230026, China*

Baojiu Li

*Institute for Computational Cosmology, Department of Physics, Durham University, South Road,
Durham DH1 3LE, United Kingdom*Cristian Barrera-Hinojosa[Ⓧ]*Instituto de Física y Astronomía, Universidad de Valparaíso, Gran Bretaña 1111, Valparaíso, Chile;
Department of Computer Science, Durham University, South Road, Durham DH1 3LE, United Kingdom;
and Institute for Computational Cosmology, Department of Physics, Durham University,
South Road, Durham DH1 3LE, United Kingdom*

Yufei Zhang

College of Mathematics and Physics, Leshan Normal University, Leshan 614000, China

(Received 28 September 2023; accepted 18 March 2024; published 29 April 2024)

In this study, we explore the potential of utilizing the four Minkowski functionals, which can fully describe the morphological properties of the large-scale structures, as a robust tool for investigating the modified gravity, particularly on nonlinear and quasilinear scales. With the assistance of the N -body simulation, we employ the Minkowski functionals to probe the Hu-Sawicki $f(R)$ gravity model. The focus is on understanding the morphological properties extracted by the Minkowski functionals and their sensitivity to modified gravity. Our analysis involves a comprehensive examination of the cosmic variance arising from finite simulation volumes. By systematically varying smoothing scales and redshifts, we quantify the information encoded in the Minkowski functionals measured from the dark-matter density field. The goal is to assess the capacity of the Minkowski functionals to constrain the model and explore potential improvements through their combination. Additionally, we investigate the impact of using biased tracers such as dark matter halos and the halo occupation distribution galaxies on the modified gravity signatures within the Minkowski functionals of the large-scale structures. Furthermore, we evaluate the influence of the redshift space distortion on the observed results. In summary, our study suggests that the Minkowski functionals of the large-scale structures hold promise as a stringent tool for constraining modified gravity and offer valuable insights into the morphological features of the cosmic web.

DOI: [10.1103/PhysRevD.109.083537](https://doi.org/10.1103/PhysRevD.109.083537)**I. INTRODUCTION**

Einstein's theory of general relativity (GR) has achieved great success within the last hundred years due to its remarkable agreement with various experimental and observational tests, ranging from tests in the laboratory and the Solar System to the emission of gravitational waves from merging binary objects [1,2], among others. However, the aforementioned tests probe the law of gravity on small scales, and any attempt to extend GR's applicable scale

to the whole Universe could represent a challenge. The discovery of the late-time accelerated Hubble expansion in the 1990s [3,4], which since then has been supported by several observational evidences (e.g., [5–8] for recent works), raises the possibility that GR might be flawed on cosmological scales.

As alternatives to GR, theories of modified gravity (MG) challenge the applicability of GR on large scales (see, e.g., [9–11] for MG reviews), arguing that the need to introduce additional dark energy, which has negative pressure and does not have any nongravitational support, follows from the inappropriate application of GR on large scales. By modifying the way gravity works, these theories offer the

*Corresponding author: wjfang@ustc.edu.cn†jax9709@mail.ustc.edu.cn

possibility of cosmic self-acceleration without introducing an additional dark energy component.

One way to test MG theories is by using the large-scale structure (LSS) of the Universe. Since the structure of the cosmic density field is formed from tiny primordial density fluctuations that grow large and nonlinear due to gravitational collapse, in general the predictions of MG gravity do not match those from GR. For instance, in $f(R)$ gravity [12–14], which is a popular class of MG theories that contains a nonlinear function $f(R)$ of the Ricci scalar R in the Einstein-Hilbert action, the enhancement of gravity leads to more nonlinear growth. That indicates to us the great importance to focus on the small scales, where the traditional two-point statistics cannot provide us with full information of the structure and higher-order statistics should be considered.

Thanks to a wealth of galaxy redshift surveys such as SDSS [15], BOSS [16], and WiggleZ [17], one can extract the evolution information encoded in the 3D galaxy distribution by two-point statistics such as the galaxy power spectrum, or correlation function, and thus probe any departure from GR and put constraints on the parameter space of MG models. Furthermore, the ongoing and upcoming redshift surveys such as DESI [18], CSST [19], and Euclid [20], will also provide us with huge survey volume and a large number of galaxies to obtain more stringent constraints from the LSS.

The two-point statistics mentioned above could fully capture the information contained in the LSS if it were Gaussian. However, extracting the non-Gaussian information efficiently is a more complicated task, especially with the increasing number of galaxies in future surveys. $N(> 2)$ -point statistics can offer us additional information, but they require much more computational resources and are challenging to model accurately. As an alternative, the Minkowski functionals (MFs), which can fully describe the morphological properties of the LSS, have been proposed [21]. Compared with N -point statistics, MFs can principally capture information at all orders, and they are easier to measure. Since being introduced in the 1990s [21], they have been widely used to study primordial non-Gaussianity [22,23], weak lensing [24–26], neutrino mass [27–29], and redshift space distortions [30–32], among others.

MFs of the two-dimensional weak lensing maps have been shown to have the potential to detect the signatures of MG [33,34]. MFs of the three-dimensional LSS have also been proposed as a probe of gravity by [35]. Specifically, they measured the MFs of the matter distribution from N -body simulations and found strong signals in the MFs for discriminating the normal-branch Dvali-Gabadadze-Porrati (nDGP) [36] and $f(R)$ [14] MG models from GR. However, to apply this new probe in real galaxy surveys, one needs to take into account systematic effects such as redshift distortion and tracer bias. Moreover, since the modification to gravity is usually scale- and redshift-dependent, it is

worthwhile to study how the constraining power of the MFs varies with these two parameters.

In this work, motivated by the unprecedented precision achievable by upcoming LSS surveys, we study several aspects regarding the application of the LSS’s MFs as a probe of modified gravity in galaxy surveys. First, based on N -body simulations, we construct the Fisher matrix to quantize the constraining power of the MFs on modified gravity parameters and study how sample variance or survey volume affects the constraints by measuring the MFs in simulation boxes with different sizes. Then we analyze the effects of MG on the MFs measured from the matter density field and see how they change with the smoothing scale and redshift. We also investigate how the constraints on modified gravity parameters change with these two quantities, and we examine the improvement in the combined constraint of different smoothing scales and redshifts. Tracer bias is also taken into account in this work. By comparing the difference in the MFs between MG and GR measured from the distributions of halo and matter, respectively, we discover the significant impact of halo bias on the MG signature in the MFs. (For a detailed analysis of how redshift distortion affects the MFs of LSS, see our recent work of [32].) We then study how the constraints on modified gravity parameters change when using biased tracers of the LSS.

The investigation presented in this work contributes to deepening our understanding on MG, especially its dependence on scale, redshift, and tracer bias. They also set an example for the applications of the MFs of LSS in constraining cosmological parameters other than the MG parameters, such as the fractional matter density Ω_m , neutrino mass M_ν , and the amplitude of the density fluctuations σ_8 . The layout of this paper is as follows: We briefly describe the Hu-Sawicki $f(R)$ theory and the simulations used in this work in Sec. II. In Sec. III, we introduce the MFs and the Fisher matrix we construct to forecast the constraints on MG parameters. In Secs. IV and V, we show our results for the MFs measured from the distributions of dark matter and dark matter halos, respectively. In Sec. VI, we study how the redshift space distortion affect our results and study the results measured from the halo occupation distribution galaxies. Conclusions and discussions are presented in Sec. VII.

II. MODELS AND SIMULATIONS

An alternative way to solve the enigma of the late-time cosmic acceleration other than introducing dark energy is to modify the left-hand side of the Einstein equation and introduce extra degrees of freedom [9–11]. One of the simplest and most popular examples is $f(R)$ gravity, in which the Einstein-Hilbert action is extended to contain a generic function of the Ricci scalar R ,

$$S = \int d^4x \sqrt{-g} \left[\frac{R + f(R)}{16\pi G} \right] + \mathcal{L}_m, \quad (1)$$

where g , G , and \mathcal{L}_m are the determinant of the metric, Newton's constant, and the matter Lagrangian, respectively.

There are various functional forms of $f(R)$ proposed in the literature. Here, as a concrete example to study the MG effects on the MFs, we choose the Hu-Sawicki (HS) model [14],

$$f(R) = -m^2 \frac{c_1(R/m^2)^n}{c_2(R/m^2)^n + 1}, \quad (2)$$

where $m = H_0 \Omega_M^{1/2}$ is a mass scale, and c_1 and c_2 are free parameters. To match the background expansion of Λ CDM, they should satisfy $c_1/c_2 = 6\Omega_\Lambda/\Omega_M$, where Ω_m and Ω_Λ are fractional matter and dark energy density parameters, respectively. n is the power index, and we take $n = 1$ in this work. The dynamical degree of freedom in this model is given by the derivative $f_R \equiv \frac{df(R)}{dR}$, which can be taken as a scalar field, often called the scalaron f_R . Its present value f_{R0} is conventionally used as the parameter of the model to determine the deviation from GR, with a smaller value of $|f_{R0}|$ representing less deviation from GR.

The enhancement of gravity in the HS model can be treated as a ‘‘fifth force’’ transmitted by the scalaron, which plays a role on all scales. However, the Solar System and laboratory experiments require that GR must be recovered in dense, high-curvature regions. In $f(R)$ gravity, the chameleon mechanism achieves this requirement because the scalaron has a position-dependent mass m_{eff} ,

$$m_{\text{eff}}^2 \simeq \frac{1}{3f_{RR}} \equiv \frac{1}{3} \left| \frac{d^2 f(R)}{dR^2} \right|^{-1}, \quad (3)$$

so that, with a Yukawa-type potential, the fifth force will be exponentially screened when m_{eff} is large. In other words, the effective mass m_{eff} characterizes the Compton wavelength $\lambda_C = m_{\text{eff}}^{-1}$, which determines how far the fifth force can propagate. Below the Compton wavelength, gravity will be enhanced with a factor of 4/3.

For this work, we use the ‘‘Extended LEnsING PHysics using ANalytic ray Tracing (ELEPHANT)’’ dark-matter-only N -body simulations, which were run using the ECOSMOG codes [37,38]. This code is designed for high-accuracy, high-resolution, and large-volume cosmological simulations for a wide class of modified gravity and dynamical dark energy theories based on the RAMSES code [39]. Adaptively refined meshes are created in high-density regions to obtain high resolutions when solving the Poisson and scalaron equations. Thanks to the efficient parallelization and the usage of the multigrid relaxation to solve the equations on both the regular domain grid and refinements, high efficiency can be achieved.

The cosmological parameters for GR were chosen as the best-fit values from WMAP-9 yr [40] results, $\{\Omega_b, \Omega_{\text{CDM}}, h, n_s, \sigma_8\} = \{0.046, 0.235, 0.697, 0.971, 0.82\}$. The amplitudes of the primordial power spectrum A_s are the same for

$f(R)$ and GR simulations, hence different gravity models give different σ_8 . Each simulation evolves $N_p = 1024^3$ dark matter particles with mass $m_p = 7.8 \times 10^{10} h^{-1} M_\odot$ in a cubic box with comoving size $L_{\text{box}} = 1024 h^{-1}$ Mpc. We use five independent realizations for each gravity model in this work, and for both GR and MG models, each realization has the same initial conditions (ICs), generated by the MPGRAFIC code [41] at initial redshift $z_{\text{ini}} = 49$.

There are three HS models with parameters $|f_{R0}| = 10^{-4}$ (F4), 10^{-5} (F5), and 10^{-6} (F6) simulated in the ELEPHANT simulations. Below, we mainly use F5 for our analysis. To estimate the covariance matrix of the MFs, we use 30 independent Λ CDM realizations [42] based on RAMSES code with similar conditions: 1024^3 dark matter particles in a cubic box with size $L_{\text{box}} = 1 h^{-1}$ Gpc and $z_{\text{ini}} = 49$. These simulations use the same level of adaptive mesh refinement as ELEPHANT. The cosmological parameters of the simulations are $\{\Omega_b, \Omega_{\text{CDM}}, h, A_s, n_s\} = \{0.048, 0.259, 0.68, 2.1 \times 10^{-9}, 0.96\}$, from which the matter density variance parameter can be derived as $\sigma_8 = 0.824$.

III. METHODS

In this section, we first introduce the meaning of the four Minkowski functionals and how we measure them. Then we describe the construction of the Fisher matrix, which we use to obtain the forecast constraints on the parameter of $f(R)$ gravity.

A. Minkowski functionals and their measurement

According to Hadwiger's theorem [43,44], the morphological properties of patterns in a d -dimension space can be fully described by $d + 1$ quantities which satisfy motion invariance, additivity, etc.—namely, the Minkowski functionals. In this work, the patterns are chosen to be the excursion sets characterized by the density contrast δ —i.e., regions with density contrast $u(\mathbf{x})(\equiv \rho(\mathbf{x})/\bar{\rho} - 1)$ exceeding a specific δ , where ρ and $\bar{\rho}$ are the density and the mean density of the field, respectively. Geometrically, the four MFs V_i with $i = 0, 1, 2$ and 3 in three-dimensional space represent the excursion sets' volume, their surface area, integrated mean curvature, and Euler characteristic χ , respectively.

When studying the MFs, one often chooses to use their spatial density instead of the MFs themselves to make comparisons between samples with different volumes. Therefore, we divide the measured MFs by the volume of the simulation box. Thus, in this work, V_0 is the volume fraction occupied by the patterns, while V_1 , V_2 , and V_3 refer to the surface area, the integrated mean curvature, and the Euler characteristic per unit volume, respectively.

We derive the MG-induced features in the MFs, ΔV_i , from the differences between the averaged values in MG and GR models, which are measured from five independent

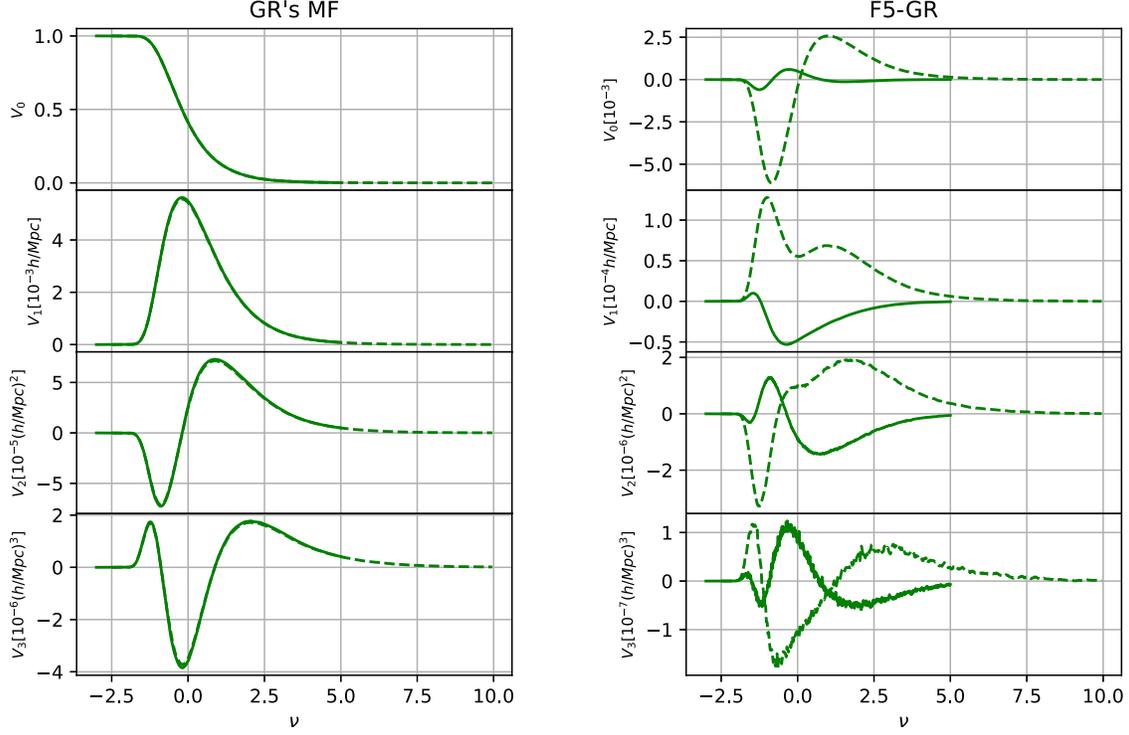


FIG. 1. The MFs (left) and the differences in MFs (right) between F5 ($|f_{R0}| = 10^{-5}$) and GR as a function of the density contrast $\delta = \rho/\bar{\rho} - 1$ (dashed lines) and the normalized density $\nu = (\rho - \bar{\rho})/\sigma$ (solid lines) measured with $R_G = 10h^{-1}$ Mpc at $z = 0$, where ρ , $\bar{\rho}$, and σ are the density, mean density, and the variance of the density field. The dashed lines are rescaled with a factor of $1/\sigma(\text{GR})$ to keep the same x scale with the solid lines.

realizations with $L_{\text{box}} = 1024h^{-1}$ Mpc for each model. We estimate the error bars using the same number of realizations. The procedures to measure the MFs are as follows: First, we construct the density field from the spatial distribution of the objects (dark matter particles or halos) using the cloud-in-cell mass-assignment scheme. Next, we smooth the field with a Gaussian window function with a smoothing scale R_G . Then, we characterize the excursion sets with the density threshold, and the MFs are measured as functions of the density contrast $\delta (\equiv (\rho - \bar{\rho})/\bar{\rho})$. When measuring the MFs, we use the integral method developed by [45]. We have tested the two standard methods they developed—the Koenderink invariant based on differential geometry, and Crofton’s formula based on integral geometry—and find that they give consistent results.

The MFs are more frequently measured as a function of normalized density threshold ν or a rescaled volume filling factor ν_A in the literature [30,46–51]. The former by definition is $\nu = \delta/\sigma$, where σ is the root mean square of the density fluctuations. The latter is defined according to the volume Minkowski functional V_0 and makes it match with a corresponding Gaussian field: $V_0 = f_A = \frac{1}{2\pi} \int_{\nu_A}^{\infty} e^{-t^2/2} dt$, where f_A is the fractional volume of the field above ν_A . One should not be puzzled by these choices, as they play different roles. When representing the MFs as a function of ν , we disregard the variance of the field but focus more on

higher-order properties. If the rescaled factor ν_A is chosen, one would further lose all Gaussian information due to rescaling, meanwhile obtaining other advantages: The MFs are invariant under a local monotonic transformation of the density field, and a properly parametrized bias scheme would not introduce non-Gaussian correction to the MFs [30,51].

When the excursion sets are characterized by the density contrast δ , all physical and unphysical factors that change the variance of the field would rescale the δ , and hence change the MFs. That is, more information will be captured by the MFs as a function of δ ; meanwhile, systematics affect the MFs more. Normalized density threshold ν and volume-filling factor ν_A are better choices if one deals with a weakly non-Gaussian case and wants to compare their measurements with the theoretical predictions. However, that is not our point. The purpose of this work is using the MFs to distinguish the modified gravity and general relativity. Thus, we follow the works [27–29,32,52] to directly choose the density contrast to define the excursion sets, in which case we would capture all information including both Gaussian and non-Gaussian properties of the field. In the two panels of Fig. 1, we show the measured MG signatures when the MFs are chosen as functions of δ and ν , respectively. One can find from the larger amplitudes of ΔV_i for the dashed lines that the difference in the variance of the field between two gravity models leads to enhancement of the MG signatures.

B. Fisher matrix

To forecast the constraints on $f(R)$ gravity from the Minkowski functionals, we use the Fisher matrix technique [53] to predict error on the modified gravity parameter. Following [54–56], we forecast the error on $\log_{10}(|f_{R0}|)$ to avoid nonzero likelihood for unphysical values of f_{R0} . Specifically, a Gaussian likelihood is assumed in the $\log_{10}(|f_{R0}|)$ space around the fiducial value $\log_{10}(|f_{R0}|) = -5$. Ignoring the dependence of the covariance matrix on the model parameters, we construct the Fisher matrix as

$$F_{\alpha\beta} = \frac{\partial \mathbf{v}^T}{\partial p_\alpha} \mathbf{C}^{-1} \frac{\partial \mathbf{v}}{\partial p_\beta}, \quad (4)$$

where p_α and p_β represent model parameters, while \mathbf{v} refers to the vector of observables, which are the MFs for different thresholds and orders, and in the most general case are measured with different smoothing scales and at different redshifts. \mathbf{C} is the covariance matrix. Note that, since we have only one parameter here—i.e., $\log_{10}(|f_{R0}|)$ —our Fisher matrix is a simple 1×1 matrix. Theoretically, the error on $\log_{10}(|f_{R0}|)$ can be estimated by $\sigma_{\log_{10}(|f_{R0}|)} = \sqrt{F^{-1}}$.

To estimate the derivative, we use the following approximation:

$$\frac{\partial \mathbf{v}}{\partial p} = \frac{\bar{\mathbf{v}}(p) - \bar{\mathbf{v}}(p^-)}{p - p^-}, \quad (5)$$

where $\bar{\mathbf{v}}$ represents the average of the MFs over all realizations of the corresponding model. Given that until very recently, there have been very few situations that explore the parameter space of f_{R0} in great detail, most existing simulations use a fixed set of $\log_{10}(|f_{R0}|)$ values—e.g., $\log_{10}(|f_{R0}|) = -5, -6$. As a result, we follow the common practice, as found in but not limited to Refs. [34,57–59], and choose p and $p^- = -5$ and -6 , respectively, for our derivative estimation. It is essential to note a potential caveat: a 20% difference in $\log_{10}(|f_{R0}|)$ in our derivative estimation might not meet the desired level of precision for convergence. A high-order estimator such as the one used in [29] may also provide us with more precise results. However, compared to GR simulations, MG simulations are much more expensive to run, which restricts the possibility of simulating many models with different values of f_{R0} to achieve a more accurate estimation of the derivatives. Due to these reasons, our results can be viewed as intuitive values to illustrate the distinguishability between the two models. For an accurate estimate of the constraint precision, one generally needs to use the Markov chain–Monte Carlo method with likelihoods estimated using full covariance matrices based on proper mock galaxy catalogues, which is also future work [60] we plan to do.

Due to the limited number of realizations, we have to use the subsample method to get enough samples for the covariance estimation. Our covariance matrix is estimated by

$$C_{\mu\nu} = \frac{1}{n-1} \sum_{i=1}^n (v_\mu^{\text{obs},i} - \bar{v}_\mu)(v_\nu^{\text{obs},i} - \bar{v}_\nu), \quad (6)$$

where n is the total number of sub-boxes, $v^{\text{obs},i}$ represents the MF measured from the i th sub-box, and the subscripts μ and ν refer to different thresholds, different orders, different smoothing scales, and different redshifts. For an unbiased estimation, we follow the correction suggested by Hartlap *et al.* in [61] and correct the inverse of the covariance matrix with a factor $\frac{n-p-2}{n-1}$, where p is the degree of freedom of observables. For each order of the MFs, we uniformly choose 21 density thresholds in the density range $[\delta_{\min}, \delta_{\max}]$ for our Fisher analysis, where δ_{\min} (δ_{\max}) is related to $V_0 = 0.99$ (0.01). This simple linear spacing bin choice may not be optimal. To further improve the constraints, one can try using nonlinear spacing bins or selecting different threshold regions for different orders of MF, as done in [29]. We have tried selecting the threshold regions according to [29] and found no improvements in our results, likely because our number of bins is limited to be much smaller than that used by [29]. Finding the most optimal bin choice is an important topic and is a comprehensive study left as future work [60].

For our covariance estimation, we use the 30 RAMSES Λ CDM simulations and divide them into 1920 sub-boxes with $L_{\text{box}} = 250h^{-1}$ Mpc when we study the effects of the smoothing scales, redshifts, and halo bias. Moreover, when studying the effects of cosmic variance, we further divide each sub-box into 8 or 64 smaller sub-boxes with $L_{\text{box}} = 125h^{-1}$ Mpc or $62.5h^{-1}$ Mpc. Then we estimate the covariance matrices from these subsamples for the three sub-box volumes. We test the convergence of our covariance matrix in Appendix A.

Since we choose F5 as the fiducial model and estimate the covariance matrix with RAMSES runs which have both different cosmological parameters and different gravity, the concern is also raised about how the covariance matrices are dependent on these factors. To test the issue, we take a comparison between the results obtained from RAMSES subsamples and those from ELEPHANT subsamples. The results are shown in Appendix B.

IV. CONSTRAINTS FROM THE MFS OF DARK MATTER

In this section, we quantify the MG information content encoded in the MFs of the dark matter density field by the Fisher matrix technique mentioned above. It is valuable to evaluate how survey volume affects the constraint on the MG parameters. We first discuss how cosmic variance will affect our constraint in Sec. IV A. Next, considering the scale dependence of the MG effects, we study the MG

signal encoded in the MFs measured with different smoothing scales R_G in Sec. IV B. Then, in Sec. IV C, we investigate the MG effects at different redshifts. Finally, in Sec. IV D, we try to combine the observables measured with different smoothing scales and at different redshifts to tighten the constraints.

A. Cosmic variance

The cosmic variance is related to the finite volume of a galaxy survey. It is important for both cosmological parameter estimation and future survey design, and it has been extensively studied for various LSS statistics [62–67]. Because our main results in this work are based on the analysis of sub-boxes with $L_{\text{box}} = 250h^{-1}$ Mpc, while the sizes of redshift surveys nowadays are often much larger than this value, it is necessary to quantify how the cosmic variance will affect the constraints from the MFs.

To do this, we separately estimate the covariance matrix from the subsample sets with different box volumes. These subsamples are optimistically assumed to only receive the influence of the box volume, and the MFs are measured from each box with $R_G = 5h^{-1}$ Mpc at $z = 0$. We show in Fig. 2 the constraints as a function of box volume V_{box} . The three data points from left to right represent constraints from sub-boxes with $V_{\text{box}} = 62.5^3$, 125^3 , and $250^3 h^{-3}$ Mpc³, respectively. We find that the value of $\sigma_{\log_{10}(f_{R0})}$ decreases by a factor of 2.0 when the box volume increases 8 times from $62.5^3 h^{-3}$ Mpc³ to $125^3 h^{-3}$ Mpc³, and by another factor of

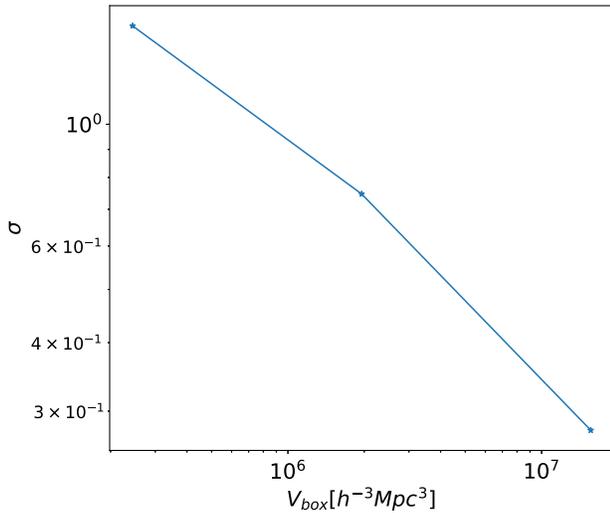


FIG. 2. Constraints on $\log_{10}(f_{R0})$ from the MFs estimated by the Fisher matrix technique, as a function of the forecasted box volume V_{box} . The MFs are measured from the dark matter density field smoothed with a Gaussian window of size $R_G = 5h^{-1}$ Mpc at $z = 0$. The three data points from left to right are obtained with the covariance matrices estimated from subsamples with $V_{\text{box}} = 62.5^3 h^{-3}$ Mpc³, $125^3 h^{-3}$ Mpc³, and $250^3 h^{-3}$ Mpc³, respectively.

2.7 when the volume increases from $125^3 h^{-3}$ Mpc³ to $250^3 h^{-3}$ Mpc³.

To understand this scaling, one can assume that the boxes in a sub-sample set are statistically independent and identically distributed. When we divide N large boxes into nN sub-boxes, the statistics S with additivity satisfy $S_{\text{large}} = nS_{\text{sub}}/n = S_{\text{sub}}$, and their standard deviations satisfy $\sigma_{\text{large}} = \sqrt{n}\sigma_{\text{sub}}/n = \sigma_{\text{sub}}/\sqrt{n}$. Hence, one can expect that $\sigma_{f_{R0}}$ should improve by a factor of $\sqrt{8}$ if the volume is enlarged by 8 times. However, the basic assumption in the subsample method, that all subsamples are fully independent of each other, is not really satisfied, which leads to systematical bias when estimating the covariance [68]. And the bias might be more significant for subsamples with smaller volumes because of the stronger correlations between small-scale structures. This may be the reason why the constraint is not strictly scaled as $1/\sqrt{V}$. There are methods to study the effects of the sample variance more accurately, such as by running more independent simulations with different volumes, or by Monte Carlo sampling for a large survey. We will consider them in future work.

This scaling with V can in principle be extended to other cosmological parameters, indicating that with the huge volume of future surveys, we can obtain much more accurate constraints on both MG and other cosmological parameters from the MFs. For example, the DESI spectroscopic survey [18] covers a sky area of ~ 14000 deg², and it is expected to detect objects in the redshift range $z \lesssim 1.9$, corresponding to a comoving volume $\sim 60h^{-3}$ Gpc³. Both of them are much larger than the volume $V_{\text{box}} = 1.56 \times 10^{-2} h^{-3}$ Gpc³ that we use to obtain our main predictions.

B. Smoothing scales

Before the measurement of the MFs, the density field is often smoothed by a Gaussian window function with a smoothing scale R_G to reduce the shot noise (e.g., [27,35,45,69]). For example, in Ref. [27], it was demonstrated that choosing an optimal value of R_G can improve the signal-to-noise (S/N) ratio of neutrino signatures when employing MFs to study the impact of massive neutrinos on LSS. Furthermore, Ref. [28] quantified the cosmological constraint obtained from the MFs of the simulated dark matter (and neutrino) field with $R_G = 5h^{-1}$ Mpc and $10h^{-1}$ Mpc, and found their combination can improve the results. Similarly to neutrinos, MG's effects on LSS are also scale dependent. Hence, in this work, we also try different R_G 's to study the scale-dependent effects on the MFs of LSS from MG.

We show in Fig. 3 the MF V_i measured with $R_G = 5h^{-1}$ Mpc, $10h^{-1}$ Mpc, and $20h^{-1}$ Mpc at $z = 0$ from the GR and F5 models and the difference ΔV_i between the MFs of the two models. The values of V_i are the means measured from the five independent simulations with

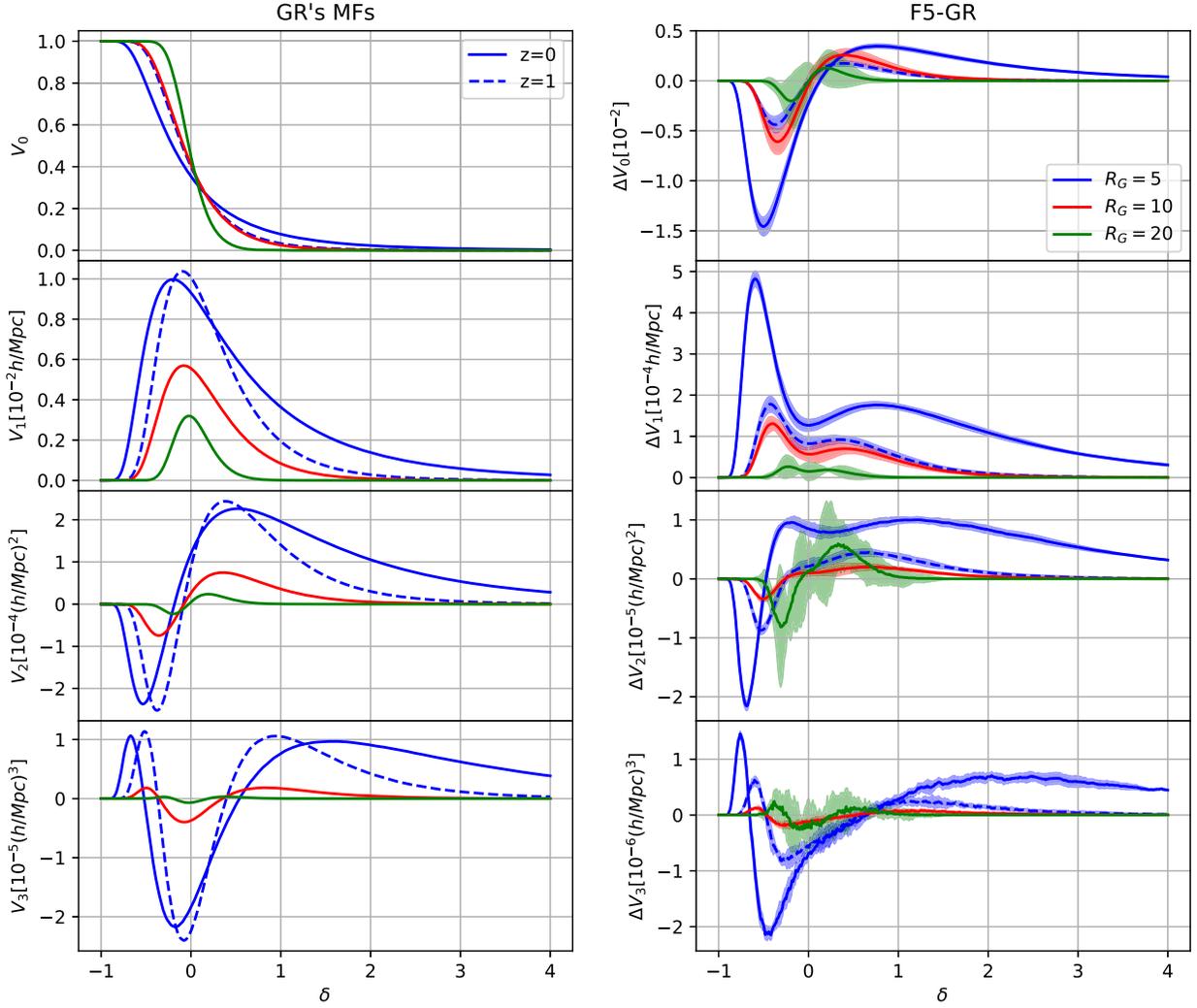


FIG. 3. The MFs (left) and the differences in MFs (right) between F5 ($|f_{R0}| = 10^{-5}$) and GR as a function of the density contrast $\delta = \rho/\bar{\rho} - 1$. The MFs are measured from the dark matter density field at $z = 0$ (solid lines) and $z = 1$ (dashed lines). We smooth the field by a Gaussian window function with scales $R_G = 5, 10, 20$ (unit: h^{-1} Mpc), respectively. (For $z = 1$, we only show the results with $R_G = 5h^{-1}$ Mpc, because the results at the two redshifts change similarly with changing R_G .) The shaded regions represent the error bands. These errors are estimated from five independent realizations with volume $V_{\text{box}} \simeq 1.07h^{-3}$ Gpc³. ΔV_2 , ΔV_3 , and their associated errors are artificially enlarged by a factor of 20 when $R_G = 20h^{-1}$ Mpc in the figure for ease of visualization.

$L_{\text{box}} = 1024h^{-1}$ Mpc, while the error bars represent their standard deviations. These are shown as a function of dark matter density contrast δ for $\delta \in [-1, 4]$. The three R_G choices from the smallest to the largest values roughly represent a nonlinear, a quasilinear, and a linear scale, respectively. We will briefly interpret the curves of V_i measured with $R_G = 15h^{-1}$ Mpc from the halo field and compare them to results for the matter density field in Sec. V, as the results with $R_G = 5h^{-1}$ Mpc have been discussed in detail in [35]. In this section, we only focus on the difference between the curves with different R_G .

We find from the curves of V_0 that when $\delta \gtrsim 0$, the excursion sets occupy a smaller fractional volume for larger R_G 's. And the fractional volume occupied by the excursion sets is larger when $\delta \lesssim 0$, which indicates a smaller

fractional volume occupied by their complements—i.e., the underdense regions. These results are consistent with an overall smaller density fluctuation for a larger R_G , because the process of smoothing erases structures with scales smaller than R_G . Other orders of V_i have similar trends with lower amplitudes when we increase R_G due to the same reason. It is worth mentioning that as R_G increases, all V_i curves tend to approach the Gaussian case.

In the right panels of Fig. 3, we find that the curves of ΔV_i also have similar shapes for different R_G 's, but lower amplitudes for larger R_G 's. The signal-to-noise ratio also decreases when we increase the smoothing scale. The Compton wavelength for F5 is $\sim 7.5h^{-1}$ Mpc at $z = 0$. One can expect more MG signals below the scale. Note that the process of Gaussian smoothing does not erase all

TABLE I. Values of $\sigma_{\log_{10}(f_{R0})}$ forecasted using the Fisher matrix of MFs based on N -body simulations. The results are obtained from the dark matter field and from three halo populations with their properties listed in Table II at two redshifts, $z = 0$ and 1 [for $z = 0$, the additional halo population \mathcal{H}_{2m} , which keeps the same lower mass limit for GR and $f(R)$, is also considered], and in real (\mathcal{R}) and redshift (\mathcal{Z}) space. C_0 and C_1 represent the results obtained by combining the MFs measured with $R_G = 5, 10,$ and $20h^{-1}$ Mpc at $z = 0$ and 1, respectively. $C_0 + C_1$ represents the result of combining C_0 and C_1 . For each order of MFs, 21 threshold bins are chosen. We estimate the covariance matrix from the subsample of RAMSES realizations with their volume $V_s = 1.56 \times 10^{-2} h^{-3} \text{ Mpc}^3$.

$R_G(h^{-1} \text{ Mpc})$		5	10	15	20
Matter, \mathcal{R}	$z = 0$	0.28	0.73	1.69	2.96
	$z = 1$	0.57	2.41	6.44	14.20
	C_0			0.26	
	C_1			0.47	
	$C_0 + C_1$			0.23	
Halo, \mathcal{R}	$z = 0, \mathcal{H}_1$	1.10	1.39
	\mathcal{H}_2	1.12	1.41
	\mathcal{H}_{2m}	1.11	1.45
	\mathcal{H}_3	1.25	1.39
Halo, \mathcal{Z}	$z = 0, \mathcal{H}_1$	1.24	1.67
	\mathcal{H}_2	1.24	1.53
	\mathcal{H}_3	1.36	1.73
Halo, \mathcal{R}	$z = 1, \mathcal{H}_1$	1.09	1.33
	\mathcal{H}_2	1.23	1.56
	\mathcal{H}_3	0.95	1.09
Halo, \mathcal{Z}	$z = 1, \mathcal{H}_1$	1.35	1.45
	\mathcal{H}_2	1.49	1.70
	\mathcal{H}_3	1.01	1.29

information on scales smaller than R_G , but rather gives an exponentially suppressed weight to structures on these scales. This is the reason why we can still find MG signals for R_G larger than the Compton wavelength.

To quantify the influence of smoothing scales, we estimate the parameter error $\sigma_{\log_{10}(f_{R0})}$ by the Fisher matrix, which is constructed from the MFs of the dark matter field with four R_G choices— $R_G = 5, 10, 15,$ and $20h^{-1}$ Mpc—at $z = 0$. The covariance matrix is estimated from the 1920 subsamples with size $L_{\text{box}} = 250h^{-1}$ Mpc. We show the results in Table I. The constraint becomes monotonically worse as we increase R_G , consistent with the findings in Fig. 3. The best constraint, $\sigma_{\log_{10}(f_{R0})} = 0.28$, is obtained with $R_G = 5h^{-1}$ Mpc.

We propose $R_G = 5h^{-1}$ Mpc as a realistic choice that is achievable for next-generation galaxy surveys [18]. However, as a theoretical analysis, we can push R_G onto smaller scales to check if there is an optimal scale where we could obtain the best constraint. In Fig. 4, we show the ratio of the constraint obtained from the MFs measured with Gaussian smoothing scale R_G to that with $R_G = 5h^{-1}$ Mpc.

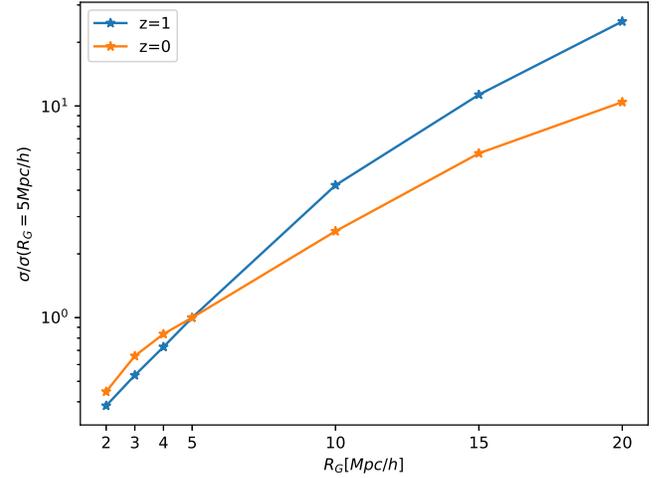


FIG. 4. Ratio of the constraint on $\log_{10}(f_{R0})$, $\sigma_{\log_{10}(f_{R0})}$, obtained from the MFs measured with Gaussian smoothing scale R_G to that with $R_G = 5h^{-1}$ Mpc. The blue and orange lines represent the ratios at $z = 1$ and $z = 0$, respectively.

Note that due to the high resolution of the N -body simulation, even if we smooth the field with R_G as small as $R_G = 2h^{-1}$ Mpc, the root mean square of the smoothed density field is still much larger than shot noise. The result shows a monotonical trend, as the constraint becomes better when R_G is decreased.

C. Redshifts

The motivation to investigate the MFs of LSS at different redshifts is that the effects of the modification to gravity are generally redshift dependent and accumulate over time. In $f(R)$, the Compton wavelength λ_C that determines how far the scalaron can propagate is a function of z . The large-scale distribution of tracers at different redshifts can provide us with complementary information.

We show the MFs V_i measured from the dark matter density field with $R_G = 5h^{-1}$ Mpc at $z = 1$ and the corresponding differences ΔV_i in the MFs between two models as dashed lines in Fig. 3.

The morphological properties for a Gaussian random field are statistically equivalent for overdense and underdense regions. Thus, in the Gaussian case, the curves of V_0 and V_2 are central-symmetric about the origin, while the curves of V_1 and V_3 are symmetric about $\delta = 0$. Because of nonlinear evolution on small scales, these curves deviate from the Gaussian case, and there are greater deviations at $z = 0$ compared to $z = 1$, due to more nonlinear evolution at low redshift. From the curve of V_0 , we find that the excursion sets occupy a smaller fractional volume at $z = 1$ for $\delta \gtrsim 0$. Also, they occupy a larger fraction of total volume when $\delta \lesssim 0$; hence, their complements, the underdense regions, occupy a smaller volume fraction. These findings reflect a relatively smaller density fluctuation at $z = 1$. To confirm the result, we measure the root mean square of the smoothed density field by $\text{rms} = \sqrt{\langle(\rho - \bar{\rho})^2\rangle}$, where ρ and $\bar{\rho}$ are the density and

mean density of the field, respectively, and the bracket represents averaging over total grids. We find that the rms is $\sim 40\%$ smaller at $z = 1$ than at $z = 0$, which is consistent with our findings.

From the right panels of Fig. 3, we find that all curves of ΔV_i have similar shapes for the two redshifts, but the amplitudes are lower at the higher redshift. Considering the curves of the MFs in the left panels, one can easily find that the relative difference of $\Delta V_i/V_i$ between the two models is more significant at low redshift. That is expected, because the Compton wavelength at $z = 1$ is $\sim 3.6h^{-1}$ Mpc, which is smaller than its value at $z = 0$, and the modification to gravity is stronger at low redshift.

We also perform the Fisher matrix analysis to quantify the constraining power of the MFs at $z = 1$, and the results are presented in Table I. We find that, compared to the results at $z = 0$, the constraints are worse at higher redshift for each R_G . This is consistent with the findings in Fig. 3. Additionally, it keeps the trend that the constraint is better for smaller R_G , and the best constraint is $\sigma_{\log_{10}(lf_{R0})} = 0.57$, with $R_G = 5h^{-1}$ Mpc. We can also find from Fig. 4 that the trend when tracing to smaller scales is the same as the result at $z = 0$, while the curve at $z = 1$ is a little steeper.

D. Combined constraints

In previous subsections, we have evaluated the constraining power of the MFs measured with specific R_G and at specific redshift from the simulated dark matter field with a typical subsample volume $V_{\text{box}} = 1.56 \times 10^{-2} h^{-3} \text{ Gpc}^3$. The volume roughly corresponds to a redshift survey with $z_{\text{max}} = 0.1$ and sky coverage $\sim 6000 \text{ deg}^2$. We note that the next-generation redshift surveys can be much larger than this choice for both the redshift and the sky coverage; hence, we can expect more precise results.

In addition, we can consider the combination of the observables measured with different smoothing scales and at different redshifts to improve our constraints. Since the smoothing process suppresses the structures with scales smaller than R_G , the combination of observables with different R_G 's is equivalent to collecting multiscale information. We assume no correlation between the observables measured at different redshifts because of the large separation between the two redshifts we consider.

C_0 , C_1 , and $C_0 + C_1$ in Table I show the constraints obtained with the combined observables. C_0 and C_1 represent the results obtained by combining the MFs measured with $R_G = 5, 10, \text{ and } 20 h^{-1}$ Mpc at $z = 0$ and 1, respectively. $C_0 + C_1$ represents the result of combining C_0 and C_1 . We find that the improvement is not very significant for all these combinations. There are two reasons why the improvements after combining different smoothing scales are not significant: First, the correlations between the MFs with different R_G 's are strong. Second, the MG effects are not significant on large scales—i.e., R_G is $10h^{-1}$ or

$20h^{-1}$ Mpc—but rather on small scales, as one can find in Fig. 3. These findings indicate we should try our best to use small scales and low redshifts in future surveys if we want to obtain a good result.

V. MFS OF DARK MATTER HALOS

In this section, we study the MG signals encoded in the MFs of halo distribution and compare them with the results for dark matter. We briefly describe how we construct the halo catalogue in Sec. VA. Due to the low number density of the halo catalogue, we also discuss how the shot noise affects our results in Sec. VB. Then, in Sec. VC, we present our analysis of how the halo bias affects the MG signals contained in the MFs. Finally, we quantify the constraining power on the MG parameter from the MFs of the halo distribution in Sec. VD.

A. Construction of the halo catalogue

Dark matter halos serve as the foundational units of large-scale structures, and they are the hosts of galaxies. Hence, studying their statistical properties, such as clustering and abundance, is important in understanding the nature of gravity. The halo catalogues are constructed using the public ROCKSTAR [70] algorithm, which is a friend-of-friend halo finder in 6D phase space. We define the mass of a halo as m_{200c} , the mass within a sphere of radius r_{200c} , which is the radius in which the mean overdensity is 200 times the critical density of the Universe.

Figure 5 shows the halo mass function, which is the dark matter halo number density as a function of their mass, for two modified gravity models and GR at $z = 0$ and 1. Due to the enhancement of the gravity in $f(R)$ scenarios, both of the two $f(R)$ models predict more halos at almost all masses compared with the GR case [71–74], except the high-mass end for F6. This is because the chameleon mechanism works efficiently for these halos, and the enhancement of the gravity is suppressed.

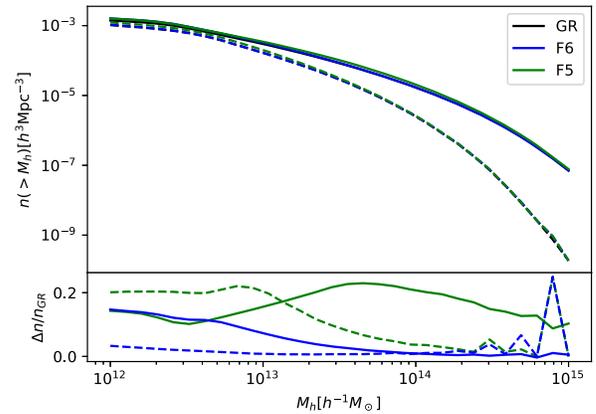


FIG. 5. Top: Halo mass functions of three models at $z = 0$ (solid lines) and $z = 1$ (dashed lines). Bottom: Relative difference in the halo mass function between modified gravity and GR.

TABLE II. Properties of the halo populations used in this work. We list the minimum halo mass M_{\min} [$10^{12}h^{-1}M_{\odot}$] used to obtain a required number density \bar{n} [$10^{-3}h^3 \text{ Mpc}^{-3}$] for a given redshift and gravity model. We also measure a scale-independent bias for these halo populations and show them in the table as a reference to help understand our MF results.

		\bar{n}	M_{\min}			Bias		
			GR	F6	F5	GR	F6	F5
$z = 0$	\mathcal{H}_1	1.0	2.57	3.04	2.96	1.26	1.23	1.23
	\mathcal{H}_2	0.7	3.98	4.45	4.45	1.36	1.33	1.34
	\mathcal{H}_3	0.4	7.33	7.96	8.50	1.51	1.47	1.48
$z = 1$	\mathcal{H}_1	1.0	1.09	1.17	1.87	2.12	2.10	2.03
	\mathcal{H}_2	0.7	2.65	2.81	3.43	2.21	2.20	2.15
	\mathcal{H}_3	0.4	5.07	5.15	5.85	2.47	2.47	2.37

To understand the MG signatures contained in the MFs of the halos, we set the same lower halo mass cut, $M_{\min} = 3.98 \times 10^{12}h^{-1}M_{\odot}$, for $f(R)$ and GR scenarios (hereafter, \mathcal{H}_{2m}). However, considering a realistic case where the number density of observables in a galaxy survey is always fixed, we also set a fixed number density (hereafter, \mathcal{H}_2), $\bar{n} = 7 \times 10^{-4}h^3 \text{ Mpc}^{-3}$, which is the number density of the GR model in \mathcal{H}_{2m} for each of the halo samples. According to Fig. 5, the halo mass cut should be higher in the $f(R)$ case. We also consider the halo populations with a fixed but higher/lower halo number density than \mathcal{H}_2 to study how different halo populations affect the MG signals we detect. The number densities of the two populations are $\bar{n} = 1.0 \times 10^{-3}h^3 \text{ Mpc}^{-3}$ (hereafter, \mathcal{H}_1) and $4.0 \times 10^{-4}h^3 \text{ Mpc}^{-3}$ (hereafter, \mathcal{H}_3). Because we only need H_{2m} to understand the full influence of the MG effects, we measure the MFs from it at $z = 0$. But for the other three halo populations, we consider both redshifts $z = 0$ and 1 to study how the MG effects evolve with time.

The lower number densities have been chosen to be representative of the samples observed in the last-stage galaxy surveys such as SDSS [15] and BOSS [16], and a high number density is expected in the current—e.g., DESI—survey [18]. The minimum halo mass for each halo sample at a given redshift is listed in Table II.

We construct the halo number density field using the cloud-in-cell (CIC) mass assignment scheme with the grid unit $4h^{-1} \text{ Mpc}$. We have tested mass assignment schemes with different accuracy levels—NGP, CIC, TSC, and PCS—using the public PYLIANS code, and we find that they provide us with consistent results if R_G is larger than the grid unit.

B. Choice of smoothing scale

The estimation accuracy of the MFs is expected to increase in a fixed volume with the choice of a smaller R_G because of the increasing number of structures with a

typical scale R_G . However, R_G should be limited by a minimum value to satisfy some criteria [48,75]. First, R_G should be larger than twice the grid cell, which is related to Nyquist resolution frequency and can also avoid the systematics caused by the mass assignment scheme. Second, R_G should be larger than the mean separation of the objects, to avoid treating a single object as a structure. That limits our minimum value of R_G to $15h^{-1} \text{ Mpc}$.

Third, the shot noise should be properly handled after the density field is smoothed with R_G . We follow the discussions in [29] to show the validity of our R_G choices. Under the Gaussian limit, the four MFs of a biased (halo/galaxy) field with a linear bias b should be determined by

$$\left(\frac{\hat{\sigma}_0}{\hat{\sigma}_1}\right)^2 = \frac{b^2\sigma_0^2 + ((4\pi)^{3/2}R_G^3\bar{n})^{-1}}{b^2\sigma_1^2 + 3((16\pi)^{3/2}R_G^5\bar{n})^{-1}}, \quad (7)$$

where σ_i is the i th moment of the field, which, with and without hat, represents the biased and matter field, respectively. \bar{n} is the mean number density of the biased tracers.

Therefore, the effect of shot noise for a biased field with number density \bar{n}_b can be equivalent to that for a matter field with number density $\bar{n}_m \sim \bar{n}_b \times b^2$. We approximately measure the bias for our three halo populations by

$$b^2 = \langle P_h(k)/P_m(k) \rangle, \quad (8)$$

where the bracket represents averaging over all k with $k < 0.5$. The values of halo bias are also listed in Table II. The values of the galaxy bias can be found in Table 1 in Ref. [76]. We thus choose the smallest, $\bar{n}_m \sim 9 \times 10^{-4}$, which is related to \mathcal{H}_3 at $z = 0$, as a representative to study how the shot noise affects our results. To do this, we use the public MG-PICOLA code [77] to fast-run realizations with the same cosmology and simulation parameters but two different resolutions with the number density of the dark matter particles $n = 1.7 \times 10^{-2}$ (high) and $n = 9 \times 10^{-4}$ (low) $h^3 \text{ Mpc}^{-3}$. The results are shown in Fig. 6 as solid and dashed lines for high and low resolutions, respectively.

One can find that all V_i 's show good consistency between the high- and low-resolution results for the two larger R_G 's, but there are large differences between the solid and dashed lines when $R_G = 10h^{-1} \text{ Mpc}$. However, we can make the corrections by multiplying a factor $((\sigma_0/\sigma_1)/(\hat{\sigma}_0/\hat{\sigma}_1))^i$ on i th-order MF V_i , and rescaling the density threshold from δ to $(\sigma_0/\hat{\sigma}_0)\delta$ to correct the shot noise effects on the measured MFs. We show the results after correcting as dotted lines, and we find that they are well consistent with the high-resolution results.

Similarly to the statements in [29], where they have tried to correct the shot noise effects on the MFs of the mock galaxy catalogue and find that the correction does not significantly affect the Fisher forecast, one should also not worry about how the shot noise will affect our Fisher results for the same reason: Shot noise effects mainly depend on

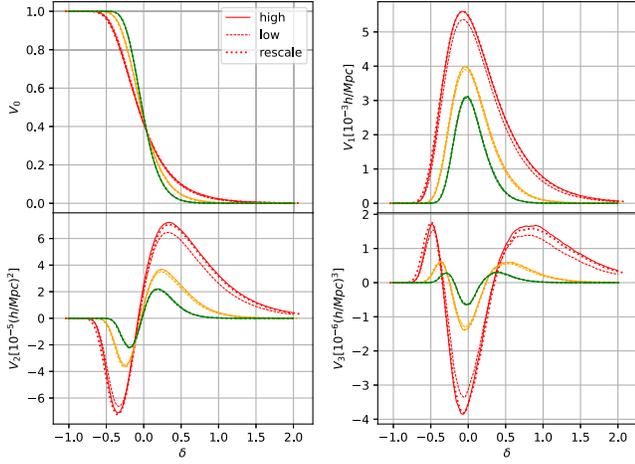


FIG. 6. The MFs measured from the COLA realizations with the number density of dark matter particles $n = 1.7 \times 10^{-2}$ (high, solid lines) and $9 \times 10^{-4} h^3 \text{ Mpc}^{-3}$ (low, dashed lines). The colors red, orange, and green represent the MFs measured with $R_G = 10, 15,$ and $20 h^{-1} \text{ Mpc}$, respectively. The results after correcting the shot noise effects are shown as dotted lines.

the number density of the field and should play a similar role in different cosmology or gravity models; thus, they almost cancel out during the process of evaluating the Fisher matrix. The density threshold δ is rescaled similarly for all the MFs present in both the derivative and covariance terms of the Fisher matrix, and the amplitude change in the derivative term should be normalized by the inverse of

the covariance matrix. Therefore, below, we only show the results without the corrections.

C. MG signatures in the MFs of halos

We show V_i and $\Delta V_i = V_i(\text{F5}) - V_i(\text{GR})$ in the left and middle panels of Fig. 7, respectively, which are measured from the distributions of the matter and the four halo populations at $z = 0$. The density field is smoothed with $R_G = 15 h^{-1} \text{ Mpc}$. We do not show the error bars here or below for which one can find the error level from the Fisher matrix results. Additionally, we show the difference in the MFs between F5 and F6 in the right panel of Fig. 7, for an intuitive sense of the derivative term in the Fisher matrix. Below, we mainly focus on the left and middle panels to understand the MG signatures.

We first discuss the overall influences of the halo bias—that is, the differences between the purple and other lines in the left panel of Fig. 7. One can directly find from Fig. 13 in Ref. [52] that the Gaussian assumption does not work for the biased field, where they show the MFs of the mock galaxies and the corresponding Gaussian field. We refer the readers to [30,51] for detailed discussions of the bias effects on the MFs under a weakly non-Gaussian assumption and to Sec. IV.3 in Ref. [29] for a brief summarization of these discussions. More intuitively, the curves of the MFs of the matter field will expand a lot to both the low and high ends of δ when transforming to the MFs of the halos due to a larger-than-one halo bias, and the larger the bias is, the wider they expand. Then, the non-Gaussian properties—for

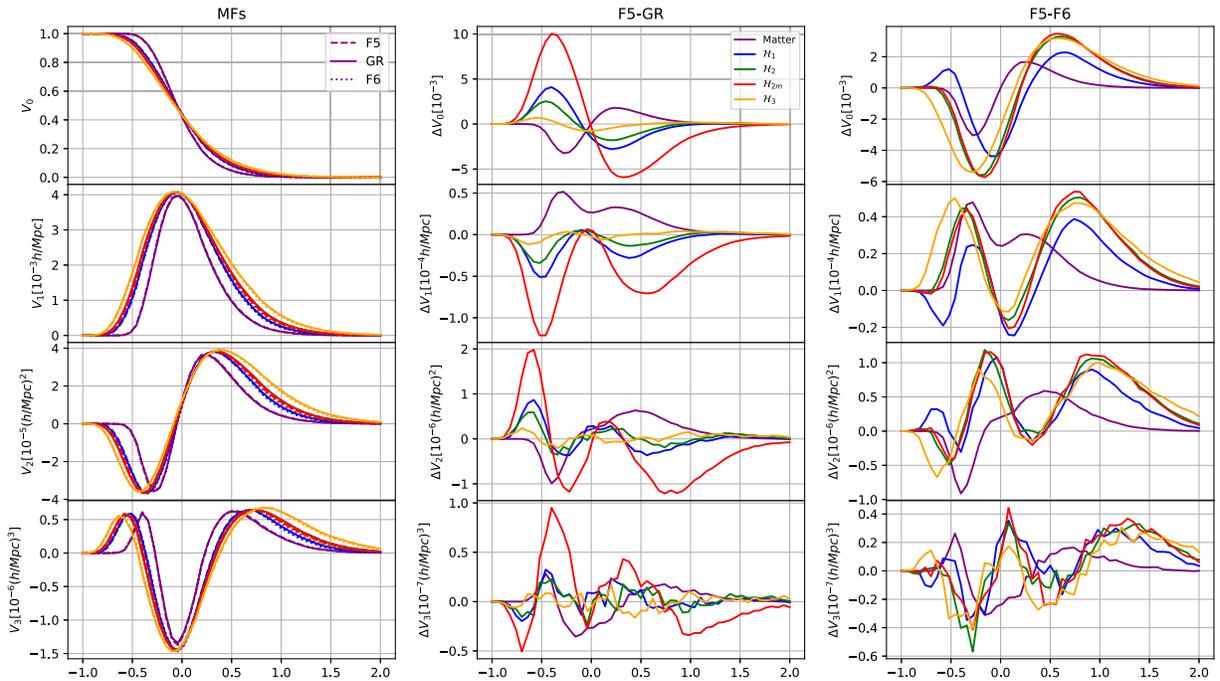


FIG. 7. Left: The MFs measured from the matter's and four halo populations' density fields for GR, F6, and F5, with $R_G = 15 h^{-1} \text{ Mpc}$ and at $z = 0$. Middle: Differences in the MFs between F5 and GR models. Right: Same as the middle panel, but between F5 and F6 models.

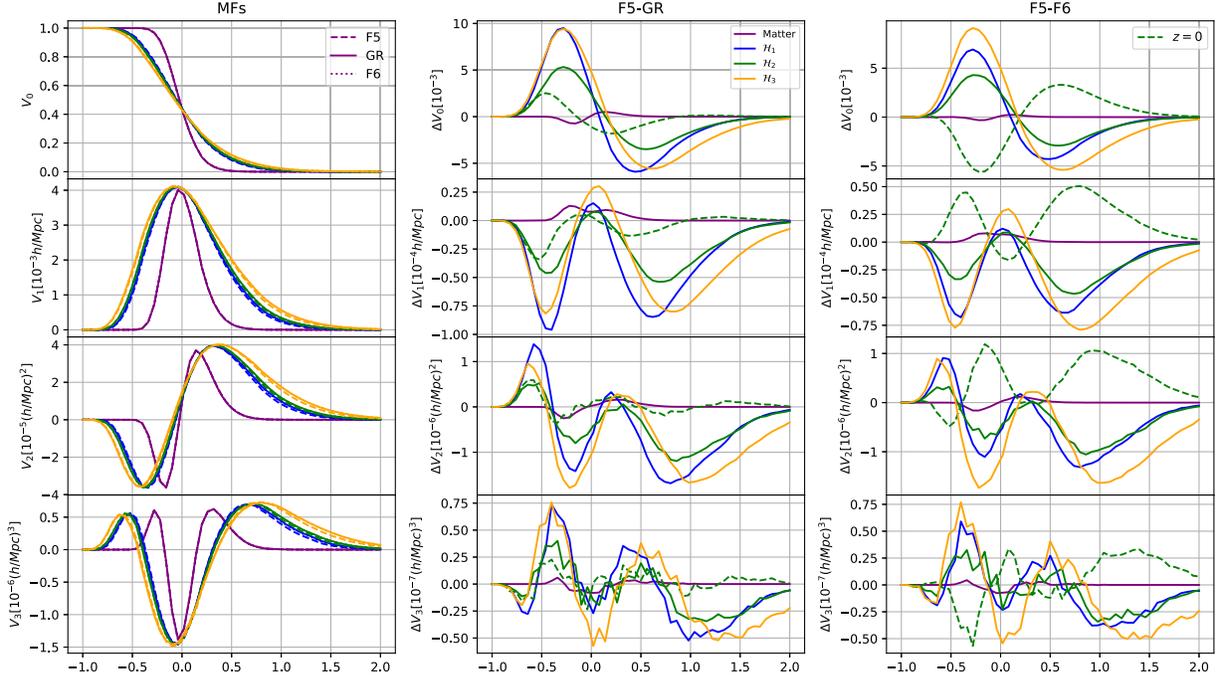


FIG. 8. Same as Fig. 7, but at $z = 1$. The green dashed lines in the middle and right panels are the results for \mathcal{H}_2 at $z = 0$, which we plot as a comparison.

example, the asymmetry that two peaks in V_3 have different widths and amplitudes—become more significant. The different values of the zero-order spectral momentum σ_0 for these fields will lead to the differences in the amplitudes of the last three MFs.

Next, we discuss the MG signatures encoded in the four MFs. We can find from the middle panel that the four halo populations have similar trends of ΔV_i , but the trends are different from the matter's. We also find that the difference between F5 and GR becomes smaller as the halo number density decreases from \mathcal{H}_1 to \mathcal{H}_3 . Moreover, the population with the same mass limit \mathcal{H}_{2m} of F5 shows much larger differences from the GR case. To understand all these findings, we can look at a fixed mass limit of the halos; due to the enhancement of gravity, there will be more halos originating from smaller-density peaks in $f(R)$ gravity when the initial conditions are the same for $f(R)$ and GR. Therefore, the overall model difference in halo clustering behaves differently from that for the matter. When the halo number density is fixed, theoretically, we are comparing the clustering properties of the same sets of density peaks, and thus the trend caused by the modified gravity will behave closely to the matter, as one can find from the right panel, where the behaviors are much more similar for the halo and the matter field. In the next subsection, we will explain the MG-induced signatures implied in the four MFs one by one based on the ΔV_i (red) curves of \mathcal{H}_{2m} .

We also show the results at $z = 1$ in Fig. 8 to see how the MG signatures encoded in the MFs of halos change with redshift. That is, the differences between F5 and GR

detected from the halo field at $z = 1$ are more significant than those at $z = 0$. Both higher- and lower-number-density halo populations show more significant MG signatures than \mathcal{H}_2 . As we have discussed above, these curves of ΔV_i receive great influence from the halo bias—e.g., the curves of ΔV_0 actually indicate a smaller halo bias in $f(R)$ scenarios, which we will discuss below. One can understand the results, as the difference in halo bias between GR and $f(R)$ is mass and redshift dependent (see Table II). In addition to the fact that the different halo populations may host different types of galaxies, our results hint that properly choosing the types of galaxies in the future survey might provide us with better constraints on the $f(R)$ gravity.

Below, we focus on the results of \mathcal{H}_{2m} in Fig. 7 and compare them with those of the matter field, and we clarify the morphological changes induced by the modified gravity.

V_0 is the volume fraction occupied by the excursion sets. From the curve of V_0 , we find that, compared to the matter field, for halos, regions with their number density above a threshold specified by a given δ occupy a larger fraction of total volume when $\delta \gtrsim 0$. But they occupy a smaller volume fraction when $\delta \lesssim 0$, which indicates a larger volume fraction occupied by the underdense regions. The results indicate an overall larger fluctuation for the halo number density field, which is consistent with the larger-than-one halo bias. From the curve of ΔV_0 , we find $\Delta V_0 < 0$ when $\delta \gtrsim 0$, and $\Delta V_0 > 0$ when $\delta \lesssim 0$ for halos. That is, the fractional volume occupied by regions with density above

an overdense threshold becomes smaller in the MG scenario, while for regions above an underdense threshold, it becomes larger. The latter indicates a smaller volume fraction occupied by regions with density below the underdense threshold. These results imply an overall smaller fluctuation for the halo field in $f(R)$ gravity. To confirm the result, we measure the root mean square of the smoothed halo number density field $R_G = 15h^{-1}$ Mpc, and we find that it is $\sim 2.6\%$ smaller for the F5 model when compared with that for GR.

However, the situation for the matter field is almost the opposite. As was discussed in [35], the curve of ΔV_0 for matter indicates that halos and voids are larger and/or more abundant in the F5 model because of the enhancement of gravity. For halos, our results indicate relatively weak clustering properties, which is consistent with a relatively smaller halo bias in the F5 model [78,79]. As we state above, when the mass cut is the same, the halos in F5 form from smaller peaks in the initial density field, which are intrinsically less clustered. Another reason is that the bias is the ratio between halo clustering and matter clustering, and matter clustering is stronger in F5 than in GR. It is worth mentioning that the result for $\delta \lesssim 0$ indicates that if voids are traced by halos, they might be smaller and/or less abundant in the MG scenario, which is consistent with [80,81].

V_1 is the surface area of the excursion sets or the isodensity contours. The values of V_1 become smaller in the F5 model for nearly all density thresholds for the halo field. For a high enough threshold, one can assume that the excursion sets are isolated overdense regions enclosed by the isodensity contours. Hence, it is natural to expect that V_1 becomes smaller, given that the volume fraction occupied by these overdense regions becomes smaller. For a low enough density threshold, although one can still assume the isodensity contours are isolated, they no longer enclose the excursion sets, but rather the regions with density below the threshold, whose volume fraction is $1 - V_0$. Thus, the surface area is expected to become smaller when $\Delta V_0 > 0$. The trend of ΔV_1 is the opposite for the matter density field, but the logic is the same: the change of the surface area of the isodensity contours follows that of the volume fraction occupied by the regions they enclose.

V_2 is the integrated mean curvature over the surface area of the excursion sets,

$$V_2(\delta) = \frac{1}{\mathcal{V}} \int v_2^{(loc)}(\delta, \mathbf{x}) d^2 S(\mathbf{x}), \quad (9)$$

where \mathcal{V} is the volume of the box, and $v_2^{(loc)}(\delta, \mathbf{x})$ is the local mean curvature (average of the two principle curvatures) at a point \mathbf{x} on the surface of the excursion sets with density threshold specified by δ . From the definition, the MG effects on V_2 are determined by both its effects on the surface area, which we have already obtained, and on the mean curvature. As suggested in [28,32], we can

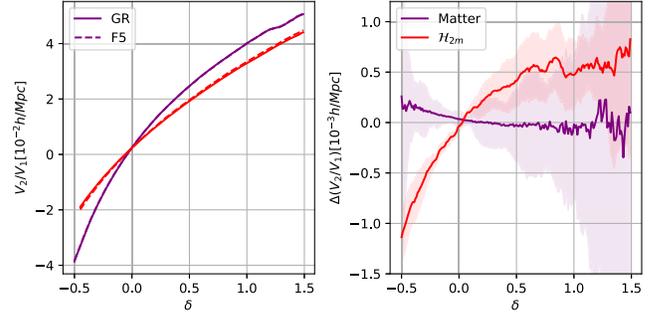


FIG. 9. Left: Ratio of V_2 to V_1 for F5 (dashed line) and GR (solid line) measured from the halo (red, \mathcal{H}_{2m}) and matter (purple) density field with $R_G = 15h^{-1}$ Mpc and at $z = 0$, as a function of δ . Right: The difference in this ratio between the two gravity models.

separate the two by rewriting V_2 as

$$V_2(\delta) = V_1(\delta) \langle v_2(\delta) \rangle. \quad (10)$$

Here, the ratio of V_2 to V_1 gives the surface-area-weighted average of the mean curvature $\langle v_2(\delta) \rangle$. The sign of V_2 represents whether the surface is overall convex or concave, and its absolute value $|V_2|$ determines how much the surface is curved. Since we define the positive direction of the surface pointing from lower- to higher-density regions, V_2 is negative when $\delta \lesssim -0.1$ and positive when $\delta \gtrsim -0.1$. By comparing the curves of V_2 for the two gravity models, we find that $|V_2|$ is larger in the F5 model when $-0.46 \lesssim \delta \lesssim 0.38$ for the halo field, and smaller elsewhere. The situation is different for the matter field, where $|V_2|$ is larger for nearly all density ranges in the MG scenario.

To obtain the new information brought by V_2 , we plot V_2/V_1 and the difference between it and the two gravity models in Fig. 9. The ratio is a global measure of the mean curvature of excursion sets. We take $\delta \lesssim -0.5$, where the values of both $|V_2|$ and V_1 are smaller in F5 for halos, as an example to help understand the curves. We can assume the underdense regions are isolated and spherelike structures in this density range, whose curvatures are inversely proportional to their radii. Compared to the GR case, the smaller $|V_2/V_1|$ in F5 for the matter field indicates the underdense structures (e.g., voids traced by dark matter particles) on average have larger radii (in other words, voids are larger and/or large voids are more abundant). The larger $|V_2/V_1|$ for the halo field indicates the opposite results: voids traced by dark matter halos are statistically smaller in the F5 model, which is consistent with findings in [80,81].

The Euler characteristic V_3 describes the connectedness of the structures, which equals the number of isolated structures minus the number of holes. We find that for halos, the excursion sets are more connected with $V_3 < 0$ when $-0.34 \lesssim \delta \lesssim 0.32$ and less connected with positive

V_3 elsewhere. From the curve of ΔV_3 , we find that in the ranges $-0.52 \lesssim \delta \lesssim -0.16$ and $0.08 \lesssim \delta \lesssim 0.7$, the structures are less connected in F5 with $\Delta V_3 > 0$, and more connected elsewhere.

D. Constraint from the MFs of halos

To evaluate the constraining power from the MFs of the halo number density field, we also use the Fisher matrix to estimate the predictive error $\sigma_{\log_{10}(|f_{R0}|)}$. Similarly to what we did for the matter density field, we divide the halo density field constructed from each original realization into 64 subsamples to estimate the covariance matrix. The constraints obtained from the halo field are also shown in Table I.

We find that for all these halo populations, the variation of the constraint with R_G is the same as that for the matter field. For each R_G , the constraints show dependence on the redshifts and halo populations. For example, the results obtained from \mathcal{H}_2 at $z = 0$ are slightly better than the constraint from the same halo population at high redshift. But the results from the halo population with the lowest number density provide us with the best constraint at $z = 1$. These findings are consistent with the results shown in Figs. 7 and 8.

Similar forecasts for the constraints on the $f(R)$ modified gravity parameters have been made in the literature by using more traditional statistics—e.g., using the galaxy clustering and cluster abundance [55], galaxy clustering and weak lensing [56], and cosmic shear statistics [34,82]. Specifically, [55] conducts a Fisher analysis using σ_8 constraints derived from the abundance of the thermal Sunyaev-Zel'dovich effect in selected galaxy clusters from the Simons Observatory as well as linear and quasilinear redshift-space two-point galaxy correlation functions from a DESI-like experiment. For a log-based fiducial parameter value of $\log_{10}(|f_{R0}|) = -5$, paired with the parameter value $n = 1$, they find a predicted error $\sigma_{\log_{10}(|f_{R0}|)} = 0.12$ and $\sigma(n) = 0.36$ after combining the galaxy clustering and cluster abundance (see Table II in Ref. [55]). Reference [56] even attempts a step forward, trying to establish a nonlinear matter power spectrum emulator for modified gravity models based on N -body simulations. For a Euclid-like mission and a fiducial value of $\log_{10}(|f_{R0}|) = -5$, they obtain a constraint of $\sigma_{\log_{10}(|f_{R0}|)} = 0.2$ by using the redshift-space galaxy power spectrum with $k_{\max} = 0.25h/\text{Mpc}$, which can be tightened to $\sigma_{\log_{10}(|f_{R0}|)} = 0.04$ when combined with Euclid weak lensing (see Table III in Ref. [56]).

Both of these results look better than ours. However, the survey volume we forecast for is much smaller than theirs at $1.56 \times 10^{-2} h^{-3} \text{ Gpc}^3$, while that for Euclid is $\sim 68 h^{-3} \text{ Gpc}^3$, and that for DESI is $\sim 61 h^{-3} \text{ Gpc}^3$. According to our findings in Sec. IV A, the constraints we obtain will be improved by approximately a factor of 66 or 63, reaching the accuracy level of 10^{-3} should we assume the survey volume of Euclid

or DESI. Considering we have only one parameter in our forecast, these constraints will be degraded to some degree in a multiparameter forecast. However, our findings still highlight the strong constraining power of the MFs on MG parameters compared to the traditional statistics.

VI. REDSHIFT SPACE DISTORTION AND MFs OF HOD GALAXIES

In this section, we discuss the redshift space distortion (RSD), which is one of the most important effects for any cosmological analysis from the spectroscopic galaxy surveys. Based on the mock galaxy catalogue using the halo occupation distribution method, we briefly study the modified gravity signatures detected by the MFs of the galaxy distributions.

A. Redshift space distortion

Two-point redshift space clustering in modified gravity scenarios has been widely studied in [76,83–88]. For example, works in [76] find that the linear model cannot model the RSD well in both GR and modified gravity scenarios on small scales, and the linear distortion parameter β can hardly be used to limit the $f(R)$ gravity. But in [89], it is shown that if a proper estimator of β is chosen, a significant difference between MG and GR models can be extracted.

The effects of redshift space distortion on the MFs under the Gaussian limit are studied in [90], where it is found that by choosing a unique threshold ν_A to define the patterns, the MFs in redshift space can have a similar shape but different amplitude compared to those in real space, while the non-Gaussian case is also studied in detail in [30]. As they are one of the most important effects when applying the MFs to limit the MG parameters in the galaxy survey, we also study how the RSD effects affect our measured MFs.

We use a simple distant-observer approximation and choose the z axis as the line of sight to get the position \vec{s} of the halos/galaxies in redshift space,

$$\vec{s} = \vec{r} + (1+z)\vec{v}_{\parallel}/H(z), \quad (11)$$

where \vec{r} and \vec{v}_{\parallel} are the real space position and the line-of-sight component of their peculiar velocities, while $H(z)$ is the Hubble parameter at redshift z . We choose \mathcal{H}_2 as a representative and show its MFs at the two redshifts in Fig. 10.

We find that the RSD effects not only reduce the amplitude of the MFs with orders 1, 2, and 3, but also expand the curves of MFs from $\delta \sim 1$ toward both the higher and lower ends of thresholds. The reason for the former is that the MFs in redshift space are sensitive to the parameter fb^{-1} [30], where f is the growth rate and b is the bias. The latter is because the RSD enlarges the variance

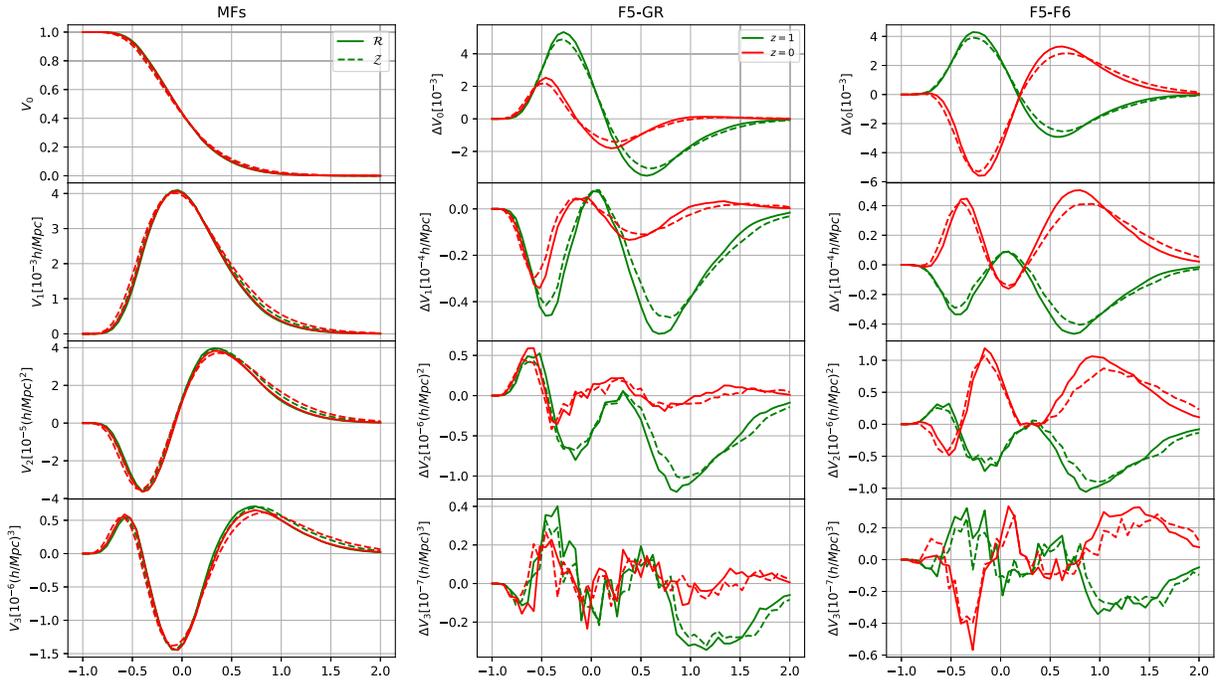


FIG. 10. Left: The MFs measured from \mathcal{H}_2 of the GR model in real (solid lines) and redshift (dashed lines) space, and at $z = 0$ (red) and $z = 1$ (green). Middle and right: Similar to the panels in Fig. 7.

of the field, and the MFs as a function of δ become wider as the variance of the field gets larger.

From the middle and right panels in Fig. 10, we find that the RSD effects slightly reduce the differences between models. The reason is that the MG-induced clustering

properties are degenerated to the RSD effects: The clustering in real space is weaker in $f(R)$, but the RSD effects which enlarge the clustering are relatively stronger. We also perform a Fisher matrix analysis to quantify the detected MG signatures in redshift space and show the results in

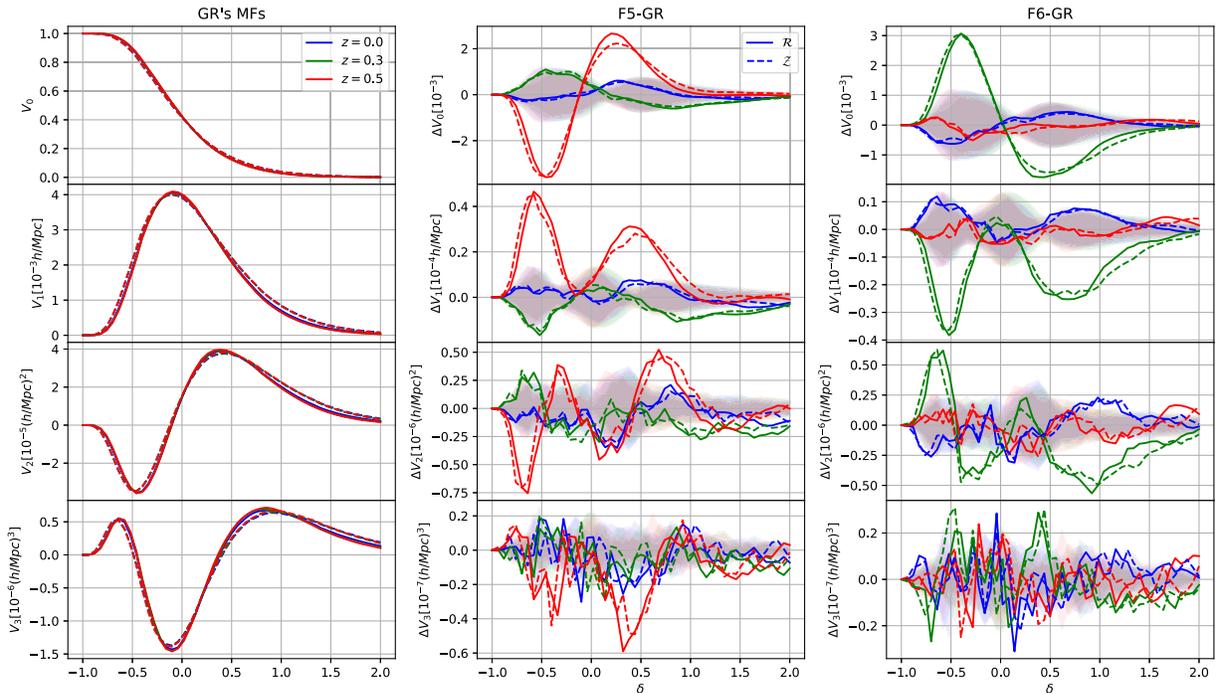


FIG. 11. Left: MFs measured from HOD galaxies of the GR model with $R_G = 15h^{-1}$ Mpc, in real (solid lines) and redshift (dashed lines) space. Middle and Right: The difference in MFs between F5/F6 and GR models.

Table I. The findings are consistent, in that the RSD effects slightly reduce the constraining power of the MFs.

B. HOD galaxies

We build the galaxy catalogue using the halo population distribution (HOD) models [91–93] with five parameters as in [93]. In this model, the mean number of galaxies in a halo with mass M_h is the sum of the mean number of central and satellite galaxies

$$\langle N(M_h) \rangle = \langle N_c \rangle + \langle N_s \rangle, \quad (12)$$

and the mean central galaxy occupation is given by

$$\langle N_c(M_h) \rangle = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{\log M_h - \log M_{\min}}{\sigma_{\log M}} \right) \right], \quad (13)$$

while the mean number of satellite galaxies is

$$\langle N_s(M_h) \rangle = \langle N_c(M_h) \rangle \times \left(\frac{M_h - M_0}{M_1} \right)^\alpha, \quad (14)$$

and $\langle N_s \rangle = 0$ if $M_h < M_0$. The central galaxies are placed at the center of their host halos, while the satellite galaxies are assumed to be radially distributed within radius r_{200c} , following the NFW profiles of their host halos.

The five independent parameters $\{M_{\min}, M_0, M_1, \sigma_{\log M}, \alpha\}$ for GR are chosen from the BOSS-CMASS-DR9 sample at $z = 0.5$ [94]. The parameters for $f(R)$ models are tuned to match the number density and real-space clustering in the GR case. We refer the interested reader to [76] for the values of these parameters and the detailed mock procedures. We show the MFs measured from the mock galaxies with $R_G = 15h^{-1}$ Mpc at three redshifts (0, 0.3, and 0.5) in Fig. 11.

Given the process of tuning the HOD parameters to align with the identical real-space clustering properties of GR and MG models, together with the degeneracy mentioned above, we may expect no MG signature in the two-point statistics measured from the HOD galaxies. But as is shown in Fig. 11, where we show the errors measured from the fluctuations of the five ELEPHANT HOD mocks, there are significant differences in the MFs between MG and GR models at some specific redshifts—e.g., $z = 0.5$ for F5 and $z = 0.3$ for F6. The finding proves the MFs' power to capture the higher-order information and their potential to be used as a probe of modified gravity in the future galaxy surveys.

VII. CONCLUSION

Examining models of modified gravity aimed at addressing the cosmic acceleration problem is crucial for advancing our understanding of both gravity and cosmic acceleration. These theories imprint distinct signatures on the large-scale structure of the Universe. In an effort to fully capitalize on upcoming precise large-scale structure surveys and derive stringent constraints on these theories,

we explore the potential of using the Minkowski functionals of the large-scale structure as a novel probe for modified gravity. Employing N -body simulations, we investigate how the signal of modified gravity in the MFs of the large-scale structures is affected by smoothing scale, redshift, tracer bias, and redshift space distortion. To quantify the constraining power, we use the Fisher matrix, exploring how the constraints on modified gravity parameters change with various factors. Our findings are summarized below:

- (1) By varying the survey volume V , we explicitly check how the constraint changes, and we find that it roughly scales with V by $\sigma_{\log_{10}(f_{R0})} \propto 1/\sqrt{V}$.
- (2) We analyze how the MG signals in the MFs of LSS change with smoothing scale and redshift, and we find that the signals are stronger for smaller smoothing scale and lower redshift. At the same time, we find that the forecasted error monotonically increases with R_G at a given redshift, and for a given R_G , the constraint is better at a lower redshift.
- (3) Because the MG effects in the MFs of LSS are more pronounced on small scales and at low redshifts, the combined constraint from different smoothing scales and redshifts is mostly dominated by the results from small smoothing scales and low redshifts. This suggests that we should try to push toward small scales and focus on low redshifts in future analysis.
- (4) To study how the tracer bias affects the results, we construct halo catalogues with different number densities at two different redshifts. We find that the MG signatures are strongly affected by the mass- and redshift-dependent halo bias.
- (5) We also study how the redshift space distortion affects our results. Due to the degeneracy between the bias and the redshift space distortion, the RSD effects will slightly reduce the MG signatures encoded in the MFs.
- (6) Although the HOD parameters are tuned to match the real-space clustering properties between GR and $f(R)$, we still find significant MG signatures contained in the MFs of the HOD galaxies, which indicates the MFs' power to extract high-order information.

Note that the constraints in this work are forecasted for a survey volume of $V \simeq 1.56 \times 10^{-2} h^{-3} \text{ Gpc}^3$, which is much smaller compared to the next-generation galaxy redshift surveys. We expect much better constraints with a volume as large as the next-generation surveys. Our best constraint from an individual scale is obtained for $R_G = 5h^{-1}$ Mpc, which is completely accessible for the currently ongoing DESI survey. Our analyses in this work are all based on N -body simulations. Systematical effects such as irregular survey geometry, which will affect our results, have not been taken into account. Although our studies focus on the MG parameter, they can be instructive for the general application of MFs as a cosmological probe.

ACKNOWLEDGMENTS

This work is supported by National Natural Science Foundation of China Grants No. 12173036 and No. 11773024, by National Key R&D Program of China Grants No. 2021YFC2203100 and No. 2022YFF0503404, by the China Manned Space Project “Probing dark energy, modified gravity and cosmic structure formation by CSST multi-cosmological measurements” and Grant No. CMS-CSST-2021-B01, by Fundamental Research Funds for Central Universities Grants No. WK3440000004 and No. WK3440000005, by Cyrus Chun Ying Tang Foundations, and by the 111 Project for “Observational and Theoretical Research on Dark Matter and Dark Energy” (B23042). B.L. is supported by the European Research Council (ERC) through a starting Grant (No. ERC-StG-716532 PUNCA), and by UK Science and Technology Facilities Council (STFC) Consolidated Grants No. ST/I00162X/1 and No. ST/P000541/1. C.B.-H. is supported by the Chilean National Agency of Research and Development (ANID) through Grant FONDECYT/Postdoctorado No. 3230512.

This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility [95]. The equipment was funded by BEIS via STFC capital Grants No. ST/K00042X/1, No. ST/P002293/1, No. ST/R002371/1, and No. ST/S002502/1, Durham University and STFC operation Grant No. ST/R000832/1. DiRAC is part of the UK National e-Infrastructure.

APPENDIX A: CONVERGENCE TEST

One may be concerned about whether the Fisher matrix converges because of our limited number of samples. As shown in Eq. (4), the Fisher matrix contains two terms: the derivative term and the covariance term. Hence, we perform the convergence test separately. For the covariance term, we fix the number of subsamples for derivative estimation and each time select N subsamples from $N_{\max} = 1920$ sub-boxes for covariance estimation. For the derivative term, we fix the number of the subsamples for covariance estimation and randomly select N subsamples from $N_{\max} = 320$ sub-boxes each time without replacement, then estimate the derivative term from them. We obtain these subsamples by dividing each ELEPHANT GR realization into 64 sub-boxes with $L_{\text{box}} = 256h^{-1}$ Mpc. We note that, due to the additivity, there should be little difference between the mean values of the MFs estimated from the original five GR realizations and their subsamples. We show in Fig. 12 how the errors from the MFs converge when the number of samples is increased. We find that when $N_{\text{covariance}} > 300$ and $N_{\text{derivative}} > 200$, the error converges [with $|\sigma(N)/\sigma(N_{\max}) - 1.0| < 0.05$].

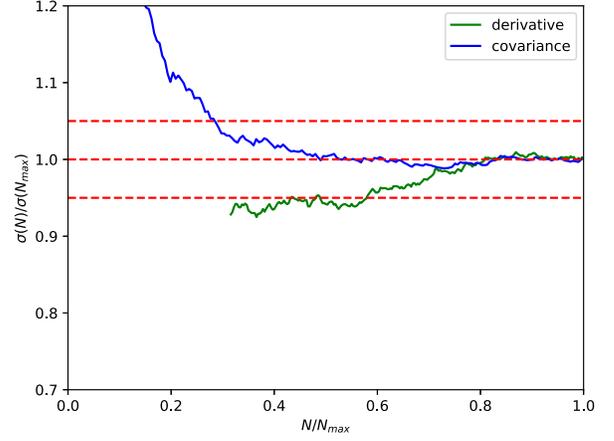


FIG. 12. Convergence of the error $\sigma_{\log_{10}(|f_{R0}|)}$ from the MFs ($R_G = 5h^{-1}$ Mpc). The blue line shows the ratio of the constraint obtained from the covariance matrix estimated from N sub-boxes to that obtained from $N_{\max} = 1920$ sub-boxes, while the number of sub-boxes used to estimate the derivative term is fixed. The green line shows the ratio of the constraint obtained with the derivative estimated from N sub-boxes to that obtained from $N_{\max} = 320$ sub-boxes, while the number of sub-boxes used to estimate the covariance term is fixed.

APPENDIX B: COVARIANCE DEPENDENCY ON THE PARAMETERS

It is a common choice to assume that the covariance matrix of the observables is independent of the cosmological and modified gravity parameters. However, since we choose F5 as our fiducial model but estimate the covariance using the GR realisations, concerns are raised regarding the validity of such a choice. In order to address these concerns, we partition the five ELEPHANT GR realizations into 320 sub-boxes and apply the same procedure to the five F5 realizations. Subsequently, we estimate the covariance matrix at $z = 0$ for each R_G from the respective subsamples.

The outcomes of this analysis are presented in Table III, and we also show the related values obtained from the RAMSES realizations in the table. The differences caused by the different cosmological or modified gravity parameters are minor, which proves the validity of our choice.

TABLE III. To test our assumption that the covariance matrix is independent with the cosmological and modified gravity parameters, we divide the five ELEPHANT realizations into 320 sub-boxes and obtain the Fisher matrix results from them. The table shows values of $\sigma_{\log(|f_{R0}|)}$ in the two cases.

$R_G (h^{-1} \text{ Mpc})$	5	10	15	20
RAMSES, GR	0.28	0.73	1.69	2.96
ELEPHANT, GR	0.27	0.70	1.52	2.56
ELEPHANT, F5	0.27	0.68	1.54	2.59

- [1] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley *et al.*, Observation of gravitational waves from a binary black hole merger, *Phys. Rev. Lett.* **116**, 061102 (2016).
- [2] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams *et al.*, Gravitational waves and gamma-rays from a binary neutron star merger: GW170817 and GRB 170817A, *Astrophys. J. Lett.* **848**, L13 (2017).
- [3] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich *et al.*, Observational evidence from supernovae for an accelerating universe and a cosmological constant, *Astron. J.* **116**, 1009 (1998).
- [4] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro *et al.*, Measurements of Ω and Λ from 42 high-redshift supernovae, *Astrophys. J.* **517**, 565 (1999).
- [5] Z. Hou, C. L. Reichardt, K. T. Story, B. Follin, R. Keisler, K. A. Aird *et al.*, Constraints on cosmology from the cosmic microwave background power spectrum of the 2500 deg² SPT-SZ survey, *Astrophys. J.* **782**, 74 (2014).
- [6] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi *et al.* (Planck Collaboration), Planck 2018 results: VI. Cosmological parameters, *Astron. Astrophys.* **641**, A6 (2020).
- [7] A. Amon, D. Gruen, M. A. Troxel, N. MacCrann, S. Dodelson, A. Choi *et al.*, Dark energy survey year 3 results: Cosmology from cosmic shear and robustness to data calibration, *Phys. Rev. D* **105**, 023514 (2022).
- [8] P. Zarrouk, E. Burtin, H. Gil-Marín, A. J. Ross, R. Tojeiro, I. Pâris *et al.*, The clustering of the SDSS-IV extended baryon oscillation spectroscopic survey DR14 quasar sample: Measurement of the growth rate of structure from the anisotropic correlation function between redshift 0.8 and 2.2, *Mon. Not. R. Astron. Soc.* **477**, 1639 (2018).
- [9] A. De Felice and S. Tsujikawa, $f(R)$ theories, *Living Rev. Relativity* **13**, 3 (2010).
- [10] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, Modified gravity and cosmology, *Phys. Rep.* **513**, 1 (2012).
- [11] S. Nojiri, S. D. Odintsov, and V. K. Oikonomou, Modified gravity theories on a nutshell: Inflation, bounce and late-time evolution, *Phys. Rep.* **692**, 1 (2017).
- [12] S. M. Carroll, V. Duvvuri, M. Trodden, and M. S. Turner, Is cosmic speed-up due to new gravitational physics?, *Phys. Rev. D* **70**, 043528 (2004).
- [13] S. M. Carroll, A. de Felice, V. Duvvuri, D. A. Easson, M. Trodden, and M. S. Turner, Cosmology of generalized modified gravity models, *Phys. Rev. D* **71**, 063513 (2005).
- [14] W. Hu and I. Sawicki, Models of $f(R)$ cosmic acceleration that evade solar system tests, *Phys. Rev. D* **76**, 064004 (2007).
- [15] D. G. York, J. Adelman, J. Anderson, E. John, S. F. Anderson, J. Annis, N. A. Bahcall *et al.*, The Sloan digital sky survey: Technical summary, *Astron. J.* **120**, 1579 (2000).
- [16] S. Alam, M. Ata, S. Bailey, F. Beutler, D. Bizyaev, J. A. Blazek *et al.*, The clustering of galaxies in the completed SDSS-III baryon oscillation spectroscopic survey: Cosmological analysis of the DR12 galaxy sample, *Mon. Not. R. Astron. Soc.* **470**, 2617 (2017).
- [17] M. J. Drinkwater, R. J. Jurek, C. Blake, D. Woods, K. A. Pimblett, K. Glazebrook *et al.*, The WiggleZ dark energy survey: Survey design and first data release, *Mon. Not. R. Astron. Soc.* **401**, 1429 (2010).
- [18] A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L. E. Allen *et al.* (DESI Collaboration), The DESI experiment part I: Science, targeting, and survey design, [arXiv:1611.00036](https://arxiv.org/abs/1611.00036).
- [19] Y. Gong, X. Liu, Y. Cao, X. Chen, Z. Fan, R. Li, X.-D. Li, Z. Li, X. Zhang, and H. Zhan, Cosmology from the Chinese space station optical survey (CSS-OS), *Astrophys. J.* **883**, 203 (2019).
- [20] L. Amendola, S. Appleby, A. Avgoustidis, D. Bacon, T. Baker, M. Baldi *et al.*, Cosmology and fundamental physics with the Euclid satellite, *Living Rev. Relativity* **21**, 2 (2018).
- [21] K. R. Mecke, T. Buchert, and H. Wagner, Robust morphological measures for large-scale structure in the Universe, *Astron. Astrophys.* **288**, 697 (1994).
- [22] C. Hikage, E. Komatsu, and T. Matsubara, Primordial non-Gaussianity and analytical formula for Minkowski functionals of the cosmic microwave background and large-scale structure, *Astrophys. J.* **653**, 11 (2006).
- [23] G. Pratten and D. Munshi, Non-Gaussianity in large-scale structure and Minkowski functionals, *Mon. Not. R. Astron. Soc.* **423**, 3209 (2012).
- [24] J. M. Kratochvil, E. A. Lim, S. Wang, Z. Haiman, M. May, and K. Huffenberger, Probing cosmology with weak lensing Minkowski functionals, *Phys. Rev. D* **85**, 103513 (2012).
- [25] D. Munshi, T. Namikawa, J. D. McEwen, T. D. Kitching, and F. R. Bouchet, Morphology of weak lensing convergence maps, *Mon. Not. R. Astron. Soc.* **507**, 1421 (2021).
- [26] M. Shirasaki, N. Yoshida, T. Hamana, and T. Nishimichi, Probing primordial non-Gaussianity with weak-lensing Minkowski functionals, *Astrophys. J.* **760**, 45 (2012).
- [27] Y. Liu, Y. Yu, H.-R. Yu, and P. Zhang, Neutrino effects on the morphology of cosmic large-scale structure, *Phys. Rev. D* **101**, 063515 (2020).
- [28] W. Liu, A. Jiang, and W. Fang, Probing massive neutrinos with the Minkowski functionals of large-scale structure, *J. Cosmol. Astropart. Phys.* **07** (2022) 045.
- [29] W. Liu, A. Jiang, and W. Fang, Probing massive neutrinos with the Minkowski functionals of the galaxy distribution, *J. Cosmol. Astropart. Phys.* **09** (2023) 037.
- [30] S. Codis, C. Pichon, D. Pogosyan, F. Bernardeau, and T. Matsubara, Non-Gaussian Minkowski functionals and extrema counts in redshift space, *Mon. Not. R. Astron. Soc.* **435**, 531 (2013).
- [31] T. Matsubara, C. Hikage, and S. Kuriki, Minkowski functionals and the nonlinear perturbation theory in the large-scale structure: Second-order effects, *Phys. Rev. D* **105**, 023527 (2022).
- [32] A. Jiang, W. Liu, W. Fang, and W. Zhao, Effects of peculiar velocities on the morphological properties of large-scale structure, *Phys. Rev. D* **105**, 103028 (2022).
- [33] C. Ling, Q. Wang, R. Li, B. Li, J. Wang, and L. Gao, Distinguishing general relativity and $f(R)$ gravity with the gravitational lensing Minkowski functionals, *Phys. Rev. D* **92**, 064024 (2015).
- [34] M. Shirasaki, T. Nishimichi, B. Li, and Y. Higuchi, The imprint of $f(R)$ gravity on weak gravitational lensing: II.

- Information content in cosmic shear statistics, *Mon. Not. R. Astron. Soc.* **466**, 2402 (2017).
- [35] W. Fang, B. Li, and G.-B. Zhao, New probe of departures from general relativity using Minkowski functionals, *Phys. Rev. Lett.* **118**, 181301 (2017).
- [36] G. Dvali, G. Gabadadze, and M. Porrati, 4D gravity on a brane in 5D Minkowski space, *Phys. Lett. B* **485**, 208 (2000).
- [37] B. Li, G.-B. Zhao, R. Teyssier, and K. Koyama, ECOSMOG: An efficient code for simulating modified gravity, *J. Cosmol. Astropart. Phys.* **01** (2012) 051.
- [38] B. Li, G.-B. Zhao, and K. Koyama, Exploring Vainshtein mechanism on adaptively refined meshes, *J. Cosmol. Astropart. Phys.* **05** (2013) 023.
- [39] R. Teyssier, Cosmological hydrodynamics with adaptive mesh refinement. A new high resolution code called RAMSES, *Astron. Astrophys.* **385**, 337 (2002).
- [40] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley *et al.*, Nine-year Wilkinson microwave anisotropy probe (WMAP) observations: Cosmological parameter results, *Astrophys. J. Suppl. Ser.* **208**, 19 (2013).
- [41] S. Prunet, C. Pichon, D. Aubert, D. Pogosyan, R. Teyssier, and S. Gottloeber, Initial conditions for large cosmological simulations, *Astrophys. J. Suppl. Ser.* **178**, 179 (2008).
- [42] C. Barrera-Hinojosa, B. Li, and Y.-C. Cai, Looking for a twist: Detecting the cosmological gravitomagnetic effect via weak lensing-kSZ cross correlations, *Mon. Not. R. Astron. Soc.* **510**, 3589 (2022).
- [43] H. Hadwiger, *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie* (Springer, Berlin, 1957).
- [44] T. Radó *et al.*, H. hadwiger, vorlesungen über inhalt, oberfläche und isoperimetrie, *Bull. Am. Math. Soc.* **65**, 20 (1959).
- [45] J. Schmalzing and T. Buchert, Beyond genus statistics: A unifying approach to the morphology of cosmic structure, *Astrophys. J. Lett.* **482**, L1 (1997).
- [46] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, The statistics of peaks of Gaussian random fields, *Astrophys. J.* **304**, 15 (1986).
- [47] H. Tomita, Curvature invariants of random interface generated by Gaussian fields, *Prog. Theor. Phys.* **76**, 952 (1986).
- [48] C. Hikage, J. Schmalzing, T. Buchert, Y. Suto, I. Kayo, A. Taruya, M. S. Vogeley, F. Hoyle, J. Richard Gott, and J. Brinkmann, Minkowski functionals of SDSS galaxies I: Analysis of excursion sets, *Publ. Astron. Soc. Jpn.* **55**, 911 (2003).
- [49] C. Park, P. Pranav, P. Chingangbam, R. Van De Weygaert, B. Jones, G. Vegter, I. Kim, J. Hidding, and W. A. Hellwing, BETTI numbers of Gaussian fields, *J. Korean Astron. Soc.* **46**, 125 (2013).
- [50] P. Pranav, R. van de Weygaert, G. Vegter, B. J. T. Jones, R. J. Adler, J. Feldbrugge, C. Park, T. Buchert, and M. Kerber, Topology and geometry of Gaussian random fields I: On Betti numbers, Euler characteristic, and Minkowski functionals, *Mon. Not. R. Astron. Soc.* **485**, 4167 (2019).
- [51] T. Matsubara, Statistics of smoothed cosmic fields in perturbation theory I: Formulation and useful formulas in second-order perturbation theory, [arXiv:astro-ph/0006269](https://arxiv.org/abs/astro-ph/0006269).
- [52] M. Lippich and A. G. Sánchez, MEDUSA: Minkowski functionals estimated from Delaunay tessellations of the three-dimensional large-scale structure, *Mon. Not. R. Astron. Soc.* **508**, 3771 (2021).
- [53] M. Tegmark, A. N. Taylor, and A. F. Heavens, Karhunen-Loève eigenvalue problems in cosmology: How should we tackle large data sets?, *Astrophys. J.* **480**, 22 (1997).
- [54] N. Ramachandra, G. Valogiannis, M. Ishak, and K. Heitmann (LSST Dark Energy Science Collaboration), Matter power spectrum emulator for $f(R)$ modified gravity cosmologies, *Phys. Rev. D* **103**, 123525 (2021).
- [55] R. Liu, G. Valogiannis, N. Battaglia, and R. Bean, Constraints on $f(R)$ and normal-branch Dvali-Gabadadze-Porrati modified gravity model parameters with cluster abundances and galaxy clustering, *Phys. Rev. D* **104**, 103519 (2021).
- [56] H. A. Winther, S. Casas, M. Baldi, K. Koyama, B. Li, L. Lombriser, and G.-B. Zhao, Emulators for the nonlinear matter power spectrum beyond Λ CDM, *Phys. Rev. D* **100**, 123540 (2019).
- [57] C.-Z. Ruan, C. Cuesta-Lazaro, A. Eggemeier, C. Hernández-Aguayo, C. M. Baugh, B. Li, and F. Prada, Towards an accurate model of small-scale redshift-space distortions in modified gravity, *Mon. Not. R. Astron. Soc.* **514**, 440 (2022).
- [58] M. Shirasaki, T. Hamana, and N. Yoshida, Probing cosmology with weak lensing selected clusters: II. Dark energy and $f(R)$ gravity models, *Publ. Astron. Soc. Jpn.* **68**, 4 (2016).
- [59] P. Zivick, P. M. Sutter, B. D. Wandelt, B. Li, and T. Y. Lam, Using cosmic voids to distinguish $f(R)$ gravity in future galaxy surveys, *Mon. Not. R. Astron. Soc.* **451**, 4215 (2015).
- [60] W. Liu, A. Jiang, and W. Fang (to be published).
- [61] J. Hartlap, P. Simon, and P. Schneider, Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix, *Astron. Astrophys.* **464**, 399 (2007).
- [62] S. Colombi, F. R. Bouchet, and R. Schaeffer, Large scale structure statistics: Finite volume effects, *Astron. Astrophys.* **281**, 301 (1994).
- [63] I. Szapudi and S. Colombi, Cosmic error and statistics of large-scale structure, *Astrophys. J.* **470**, 131 (1996).
- [64] B. P. Moster, R. S. Somerville, J. A. Newman, and H.-W. Rix, A cosmic variance cookbook, *Astrophys. J.* **731**, 113 (2011).
- [65] W. Hu and A. V. Kravtsov, Sample variance considerations for cluster surveys, *Astrophys. J.* **584**, 702 (2003).
- [66] C. E. Cunha, D. Huterer, M. T. Busha, and R. H. Wechsler, Sample variance in photometric redshift calibration: Cosmological biases and survey requirements, *Mon. Not. R. Astron. Soc.* **423**, 909 (2012).
- [67] S. P. Driver and A. S. G. Robotham, Quantifying cosmic variance, *Mon. Not. R. Astron. Soc.* **407**, 2131 (2010).
- [68] F. Lacasa and M. Kunz, Inadequacy of internal covariance estimation for super-sample covariance, *Astron. Astrophys.* **604**, A104 (2017).
- [69] A. Ducout, F. R. Bouchet, S. Colombi, D. Pogosyan, and S. Prunet, Non-Gaussianity and Minkowski functionals: Forecasts for Planck, *Mon. Not. R. Astron. Soc.* **429**, 2104 (2013).
- [70] P. S. Behroozi, R. H. Wechsler, and H.-Y. Wu, The ROCKSTAR phase-space temporal halo finder and the velocity offsets of cluster cores, *Astrophys. J.* **762**, 109 (2013).

- [71] J.-h. He, A. J. Hawken, B. Li, and L. Guzzo, Effective dark matter halo catalog in $f(R)$ gravity, *Phys. Rev. Lett.* **115**, 071306 (2015).
- [72] B. Li, G.-B. Zhao, and K. Koyama, Haloes and voids in $f(R)$ gravity, *Mon. Not. R. Astron. Soc.* **421**, 3481 (2012).
- [73] P. Arnalte-Mur, W. A. Hellwing, and P. Norberg, Real- and redshift-space halo clustering in $f(R)$ cosmologies, *Mon. Not. R. Astron. Soc.* **467**, 1569 (2017).
- [74] C. Hernández-Aguayo, C. M. Baugh, and B. Li, Marked clustering statistics in $f(R)$ gravity cosmologies, *Mon. Not. R. Astron. Soc.* **479**, 4824 (2018).
- [75] F. Hoyle, M. S. Vogeley, I. Gott, J. Richard, M. Blanton, M. Tegmark, D. H. Weinberg, N. Bahcall, J. Brinkmann, and D. York, Two-dimensional topology of the Sloan digital sky survey, *Astrophys. J.* **580**, 663 (2002).
- [76] C. Hernández-Aguayo, J. Hou, B. Li, C. M. Baugh, and A. G. Sánchez, Large-scale redshift space distortions in modified gravity theories, *Mon. Not. R. Astron. Soc.* **485**, 2194 (2019).
- [77] H. A. Winther, K. Koyama, M. Manera, B. S. Wright, and G.-B. Zhao, COLA with scale-dependent growth: Applications to screened modified gravity models, *J. Cosmol. Astropart. Phys.* **08** (2017) 006.
- [78] F. Schmidt, M. Lima, H. Oyaizu, and W. Hu, Nonlinear evolution of $f(R)$ cosmologies: III. Halo statistics, *Phys. Rev. D* **79**, 083518 (2009).
- [79] C. Arnold, P. Fosalba, V. Springel, E. Puchwein, and L. Blot, The modified gravity light-cone simulation project: I. Statistics of matter and halo distributions, *Mon. Not. R. Astron. Soc.* **483**, 790 (2019).
- [80] Y.-C. Cai, N. Padilla, and B. Li, Testing gravity using cosmic voids, *Mon. Not. R. Astron. Soc.* **451**, 1036 (2015).
- [81] S. Contarini, F. Marulli, L. Moscardini, A. Veropalumbo, C. Giocoli, and M. Baldi, Cosmic voids in modified gravity models with massive neutrinos, *Mon. Not. R. Astron. Soc.* **504**, 5021 (2021).
- [82] B. Bose, M. Cataneo, T. Tröster, Q. Xia, C. Heymans, and L. Lombriser, On the road to per cent accuracy IV: ReACT—Computing the non-linear power spectrum beyond λ CDM, *Mon. Not. R. Astron. Soc.* **498**, 4650 (2020).
- [83] L. Xu, Constraints on $f(R)$ gravity through the redshift space distortion, *Phys. Rev. D* **91**, 063008 (2015).
- [84] E. Jennings, C. M. Baugh, B. Li, G.-B. Zhao, and K. Koyama, Redshift-space distortions in $f(R)$ gravity, *Mon. Not. R. Astron. Soc.* **425**, 2128 (2012).
- [85] A. Raccanelli, D. Bertacca, D. Pietrobon, F. Schmidt, L. Samushia, N. Bartolo, O. Doré, S. Matarrese, and W. J. Percival, Testing gravity using large-scale redshift-space distortions, *Mon. Not. R. Astron. Soc.* **436**, 89 (2013).
- [86] C. Arnold, M. Leo, and B. Li, Realistic simulations of galaxy formation in $f(R)$ modified gravity, *Nat. Astron.* **3**, 945 (2019).
- [87] B. Bose, K. Koyama, W. A. Hellwing, G.-B. Zhao, and H. A. Winther, Theoretical accuracy in cosmological growth estimation, *Phys. Rev. D* **96**, 023519 (2017).
- [88] B. Bose and K. Koyama, A perturbative approach to the redshift space power spectrum: Beyond the standard model, *J. Cosmol. Astropart. Phys.* **08** (2016) 032.
- [89] J. E. García-Farieta, W. A. Hellwing, S. Gupta, and M. Bilicki, Probing gravity with redshift-space distortions: Effects of tracer bias and sample selection, *Phys. Rev. D* **103**, 103524 (2021).
- [90] T. Matsubara, Statistics of isodensity contours in redshift space, *Astrophys. J.* **457**, 13 (1996).
- [91] J. A. Peacock and R. E. Smith, Halo occupation numbers and galaxy bias, *Mon. Not. R. Astron. Soc.* **318**, 1144 (2000).
- [92] A. A. Berlind and D. H. Weinberg, The halo occupation distribution: Toward an empirical determination of the relation between galaxies and mass, *Astrophys. J.* **575**, 587 (2002).
- [93] Z. Zheng, A. L. Coil, and I. Zehavi, Galaxy evolution from halo occupation distribution modeling of DEEP2 and SDSS galaxy clustering, *Astrophys. J.* **667**, 760 (2007).
- [94] L. Anderson, E. Aubourg, S. Bailey, D. Bizyaev, M. Blanton, A. S. Bolton *et al.*, The clustering of galaxies in the SDSS-III baryon oscillation spectroscopic survey: Baryon acoustic oscillations in the data release 9 spectroscopic galaxy sample, *Mon. Not. R. Astron. Soc.* **427**, 3435 (2012).
- [95] Dirac High Performance Computing Facility: The UK's High Performance Super Computer facility for the STFC Particle Physics, Astrophysics, Cosmology, Solar System and Planetary Science and Nuclear Physics theory community, www.dirac.ac.uk.