

# Studying large-scale structure probes of modified gravity with COLA

Bartolomeo Fiorini,<sup>\*</sup> Kazuya Koyama and Albert Izzard

Institute of Cosmology & Gravitation, University of Portsmouth,  
Dennis Sciamia Building, Burnaby Road, Portsmouth, PO1 3FX, U.K.

E-mail: [bartolomeo.fiorini@port.ac.uk](mailto:bartolomeo.fiorini@port.ac.uk), [kazuya.koyama@port.ac.uk](mailto:kazuya.koyama@port.ac.uk),  
[albert.izzard@port.ac.uk](mailto:albert.izzard@port.ac.uk)

Received August 5, 2022

Revised September 15, 2022

Accepted November 20, 2022

Published December 19, 2022

**Abstract.** We study the effect of two Modified Gravity (MG) theories,  $f(R)$  and nDGP, on three probes of large-scale structure, the real space power spectrum estimator  $Q_0$ , bispectrum and voids, and validate fast approximate COLA simulations against full  $N$ -body simulations for the prediction of these probes. We find that using the first three even multipoles of the redshift space power spectrum to estimate  $Q_0$  is enough to reproduce the MG boost factors of the real space power spectrum for both halo and galaxy catalogues. By analysing the bispectrum and reduced bispectrum of Dark Matter (DM), we show that the strong MG signal present in the DM bispectrum is mainly due to the enhanced power spectrum. We warn about adopting screening approximations in simulations as this neglects non-linear contributions that can source a significant component of the MG bispectrum signal at the DM level, but we argue that this is not a problem for the bispectrum of galaxies in redshift space where the signal is dominated by the non-linear galaxy bias. Finally, we search for voids in our mock galaxy catalogues using the ZOBOV watershed algorithm. To apply a linear model for Redshift-Space Distortion (RSD) in the void-galaxy cross-correlation function, we first examine the effects of MG on the void profiles entering into the RSD model. We find relevant MG signals in the integrated-density, velocity dispersion and radial velocity profiles in the nDGP theory. Fitting the RSD model for the linear growth rate, we recover the linear theory prediction in an nDGP model, which is larger than the  $\Lambda$ CDM prediction at the  $3\sigma$  level. In  $f(R)$  theory we cannot naively compare the results of the fit with the linear theory prediction as this is scale-dependent, but we obtain results that are consistent with the  $\Lambda$ CDM prediction.

**Keywords:** cosmological parameters from LSS, cosmological simulations, galaxy clustering, modified gravity

**ArXiv ePrint:** [2208.01345](https://arxiv.org/abs/2208.01345)

<sup>\*</sup>Corresponding author.



---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Simulations and HOD tuning</b>	<b>3</b>
<b>3</b>	<b>Power spectrum</b>	<b>4</b>
3.1	Estimator of the real space power spectrum	4
3.2	Results for power spectrum	5
<b>4</b>	<b>Bispectrum</b>	<b>7</b>
4.1	Definitions and measurements	7
4.2	Results for bispectrum	8
<b>5</b>	<b>Voids</b>	<b>12</b>
5.1	Redshift-space distortion model in voids	13
5.2	Velocity model	14
5.3	Void finding and measurements	14
5.4	Results for voids	19
<b>6</b>	<b>Conclusions</b>	<b>24</b>
<b>A</b>	<b>Bispectrum of halos and galaxies in real space</b>	<b>26</b>

---

## 1 Introduction

Thanks to upcoming stage IV experiments such as the space-based Euclid mission [1] and the ground-based Rubin observatory’s Legacy Survey of Space and Time (LSST) [2] and the Dark Energy Spectroscopic Instrument (DESI) [3], we are entering an era of unprecedented detailed cosmological observations of the Large-Scale Structure (LSS). Joint multi-probe techniques are a promising direction to break the degeneracies in the theory parameter space, enabling tight constraints on cosmological parameters at intermediate redshift. In particular, since structure formation is tightly linked to the dynamics of gravity, it is key to detect any deviation from General Relativity (GR) in future analyses, so that we better understand gravity on cosmological scales and also to get unbiased inferences of cosmological parameters.

The simplest extensions to GR can be described by an extra degree of freedom in the form of a scalar field that mediates an additional force, the so-called fifth force. The stringent Solar System tests of gravity impose tight constraints on the strength of the fifth force [4] which needs to be weak in such environments, therefore Modified Gravity (MG) theories of cosmological interest incorporate screening mechanisms that suppress the effect of the fifth force in high-density regions [5, 6].

Cosmological simulations are indispensable to create model predictions on non-linear scales and to better constrain the relationship between galaxies and the underlying Dark Matter (DM) density field, encapsulated in the galaxy bias. The non-linear regime is also

where the screening mechanisms can leave their imprint. Full  $N$ -body simulations enable accurate predictions of the DM density field on non-linear scales but they are computationally very expensive. Their extension to MG are even more expensive because they have to solve the non-linear differential equation for the extra scalar field [7].

Approximate simulation methods are computationally cheaper alternatives to full  $N$ -body simulations. In particular, the COmoving Lagrangian Acceleration (COLA) method [8] combines second-order Lagrangian perturbation theory (2LPT) to model analytically the trajectories of DM particles on large scales, with an  $N$ -body method known as the Particle Mesh (PM) algorithm to take care of small scale physics. This separation between large scales and small scales allows COLA simulations to trade accuracy on small scales for a great reduction in computational cost while leaving the accuracy on large scales unaffected. The extension of the COLA method to MG theories requires the use of screening approximations to avoid a significant impact on their computational efficiency. Thanks to the use of screening approximations based on spherically symmetric solutions developed in [9], the COLA method was extended to MG theories while still being computationally efficient [10, 11]. Despite the limited accuracy of COLA in the deep non-linear regime, it has been shown to have a good performance in predicting the reaction of the matter power spectrum to changes in cosmological parameters, with sub-percent accuracy up to  $k \sim 1 h \text{ Mpc}^{-1}$  for a wide range of cosmologies around the reference one [12].

Several empirical prescriptions have been proposed in the literature to link the DM density field predicted by DM-only cosmological simulations to the distribution of galaxies. Among them, the widely used Halo Occupation Distribution (HOD) prescription has been extended and validated to produce galaxy mocks catalogues from COLA simulations in  $f(R)$  [13] and nDGP [14] gravity theories, reproducing full  $N$ -body results for the MG signals on the redshift-space multipoles of the galaxy power spectrum [15].

Given the large number of parameters that enter the modelling of galaxy clustering (e.g. cosmological parameters, MG parameters, HOD parameters), degeneracies in the high-dimensional theory parameter space can affect the constraining power of the LSS. However, the theory parameters may have different impacts on different summary statistics. An example of this is the degeneracy between massive neutrinos and  $\sigma_8$  in the power spectrum which is broken with bispectrum information [16, 17].

In this work, we exploit the results and simulated data sets from [15] to investigate the impact of MG in statistics beyond the redshift-space multipoles of the power spectrum, and by doing so we validate the COLA method to formulate theoretical predictions on these additional statistics in MG theories. As mentioned, COLA simulations developed in [11, 18] use a screening approximation that linearises the scalar field equation while including an environmental dependent screening factor. The validity of this approximation needs to be checked for higher-order statistics as they lack contributions from the non-linearity of the scalar field. We focus on two MG theories, the Hu-Sawicki  $f(R)$  theory [19] for  $n = 1$  and the normal branch of the 5D braneworld theory known as Dvali-Gabadadze-Porrati gravity (DGP) [14], because they are good representatives of the screening mechanisms they incorporate, chameleon [20, 21] and Vainshtein mechanism [22], respectively. After reviewing the simulations and catalogues at the base of this analysis, we test the effects of MG theories and assess COLA's accuracy on the power spectrum orthogonal to the line-of-sight in section 3, on the bispectrum in section 4 and on voids in section 5. We then conclude in section 6.

Models	GR, F5, N1
Realisations	5
Box size	$1024 h^{-1} \text{Mpc}$
$N_{\text{part}}$	$1024^3$
$M_{\text{part}}$	$7.7985 \cdot 10^{10} h^{-1} M_{\odot}$
Force grid	$3072^3$
Timesteps	30
Initial conditions	2LPT, $z = 49$

**Table 1.** PITER simulations.

Models	GR, F5, N1
Realisations	5
Box size	$1024 h^{-1} \text{Mpc}$
$N_{\text{part}}$	$1024^3$
$M_{\text{part}}$	$7.7985 \cdot 10^{10} h^{-1} M_{\odot}$
Domain grid	$1024^3$
Refinement criterion	8
Initial conditions	Zel'dovich, $z = 49$

**Table 2.** ELEPHANT simulations.

## 2 Simulations and HOD tuning

Theoretical predictions on the redshift-space power spectrum (and higher-order statistics) of galaxies can be formulated from mock galaxy catalogues produced using cosmological simulations with the HOD prescription.

In this work, we use data from  $z \simeq 0.5$  of the PITER and ELEPHANT simulations suites produced in [15] for the extension and validation of the pipeline for the production of galaxy mock catalogues in MG with the COLA method. The ELEPHANT simulations [23] are full  $N$ -body simulations in GR,  $f(R)$ , and nDGP gravity computed with the ECOSMOG code [24], which implements the adaptive mesh-refinement technique and solves the equation for the fifth force without approximations. The PITER simulation suite introduced in [15] is a set of COLA simulations designed to mimic the ELEPHANT simulations. Both PITER and ELEPHANT suites (COLA and  $N$ -body respectively from now on) follow the evolution of  $1024^3$  particles in a  $(1024 h^{-1} \text{Mpc})^3$  box and consist of 5 realisations in GR, F5 and N1 gravity.<sup>1</sup> F5 refers to the Hu-Sawicki  $f(R)$  model with  $n = 1$  and  $|f_{R0}| = 10^{-5}$  and N1 to the normal branch of the DGP model (nDGP) with  $H_0 r_c = 1$ . The cosmology characterising these simulations is

$$\begin{aligned} \Omega_{m,0} &= 0.281, & \Omega_{\Lambda,0} &= 0.719, & \Omega_{b,0} &= 0.046, \\ n_s &= 0.971, & \sigma_8 &= 0.842, & h &= 0.697, \end{aligned} \quad (2.1)$$

and additional technical details are summarised in table 1 and table 2 (see [15] for a comparison of the settings adopted in the two simulations suites and [15, 23] for more details on the simulations settings).

The halo catalogues were created from the snapshots of COLA and  $N$ -body simulations at redshift  $z = 0.5057$  using a friend-of-friend halo finder with linking length of 0.2 in units of the mean inter-particle distance. The halo catalogues were populated with galaxies by means of the HOD model proposed in [25] for the occupation numbers of central and satellite galaxies. To enable the accurate modelling of MG effects on the halo profiles and consequently on the phase-space distribution of satellite galaxies, in F5 theory, the concentration-mass relation and the velocity-dispersion profile for the Navarro-Frenk-White model [26] were modified based on fitting functions calibrated on  $N$ -body halo catalogues produced with the Rockstar halo-finder [27]. The HOD tuning was performed in MG theories and in COLA simulations to reproduce the monopole  $P_0$  and the quadrupole  $P_2$  of the galaxy power

<sup>1</sup>In each realisation the same initial seed is used for all gravity models, but COLA and  $N$ -body simulations have different initial seeds.

spectrum<sup>2</sup> in redshift-space up to  $k_{\text{max}} = 0.3 h \text{ Mpc}^{-1}$  estimated using  $N$ -body simulations in GR with the HOD parameters in eq. (26) of [28], originally obtained to reproduce the clustering of CMASS galaxies in the data release 9 of the BOSS survey.

All the halo samples drawn in this work follow the same abundance matching criteria as in [15]. That is, for each simulation, the most massive halos are selected while matching the reference abundance determined in the  $N$ -body catalogues in GR using a mass-cut at  $M_{\text{cut}} = 10^{13} h^{-1} M_{\odot}$ .

To increase the signal-to-noise ratio in summary statistics at the galaxy level, [15] relied on a set of 5 HOD realisations produced from each simulation realisation and, in addition to this, redshift-space statistics were estimated by averaging over the 3 lines-of-sight parallel to the box's major axes. We adopt the same technique for the galaxy statistics discussed in this work.

### 3 Power spectrum

The real space power spectrum

$$\langle \delta(\mathbf{k}) \delta(\mathbf{k}') \rangle \equiv (2\pi)^3 P^r(k) \delta_D(\mathbf{k} + \mathbf{k}') . \quad (3.1)$$

is not directly measurable in galaxy surveys because we cannot probe the real space position of galaxies. What we can directly measure is the redshift-space power spectrum  $P^s(k, \mu)$  where  $\mu \equiv \hat{z} \cdot \hat{k}$  and  $\hat{z}$  is the line-of-sight in the plane-parallel approximation. The redshift-space power spectrum can be projected on the orthonormal basis of Legendre polynomials

$$P_{\ell}(k) = \frac{2\ell + 1}{2} \int_{-1}^1 P^s(k, \mu) \mathcal{L}_{\ell}(\mu) d\mu , \quad (3.2)$$

where  $\mathcal{L}_{\ell}$  is the Legendre polynomial of order  $\ell$ , obtaining the multipole of the power spectrum  $P_{\ell}$ .

#### 3.1 Estimator of the real space power spectrum

Before studying statistics beyond the power spectrum, we first study the method to estimate the real space power spectrum from the combination of redshift-space multipoles proposed in [29].

From the definition of the power spectrum multipoles and given the orthogonality of the Legendre polynomials, the redshift-space power spectrum  $P^s(k, \mu)$  can be reconstructed using the multipoles  $P_{\ell}$

$$P^s(k, \mu) = \sum_{\ell} P_{\ell}(k) \mathcal{L}_{\ell}(\mu) . \quad (3.3)$$

On the other hand,  $P^s(k, \mu)$  can also be expanded in the monomial basis as

$$P^s(k, \mu) = \sum_n Q_n(k) \mu^n . \quad (3.4)$$

An interesting quantity arising from the latter expansion is the first moment  $Q_0 = P^s(k, 0)$  which, representing the power spectrum orthogonal to the line-of-sight, is unaffected by

---

<sup>2</sup>See section 3 for the definition of the power spectrum and its multipoles in redshift space.

Redshift-Space Distortions (RSD) and coincides with the real space power spectrum  $P^r(k)$ . Using eq. (3.3),  $Q_0$  can be expressed in term of the power spectrum multipoles  $P_\ell$

$$Q_0 = \sum_{\ell} P_{\ell}(k) \mathcal{L}_{\ell}(0). \quad (3.5)$$

The Kaiser model for RSD [30] predicts that linear velocity inflow towards over-densities produces power transfer from the real space power spectrum to the quadrupole and hexadecapole of the power spectrum in redshift-space

$$\delta^s(\mathbf{k}) = \left[1 + f\mu^2\right] \delta^r(\mathbf{k}). \quad (3.6)$$

On top of this, the Fingers of God (FoG) effect [31], due to the random non-linear motion of galaxies on small scales, further transfers power to higher multipoles. When the FoG effect can be neglected, the sum in eq. (3.5) can be restricted to the multipoles sourced by the Kaiser effect [29, 32]

$$Q_0 = \sum_{\ell=0,2,4} P_{\ell}(k) \mathcal{L}_{\ell}(0) = P_0 - \frac{1}{2}P_2 + \frac{3}{8}P_4. \quad (3.7)$$

Otherwise higher multipoles should be considered in the expansion, and since these are more affected by noise than lower multipoles, approaches such as the one of [29] can be used to mitigate the loss of precision.

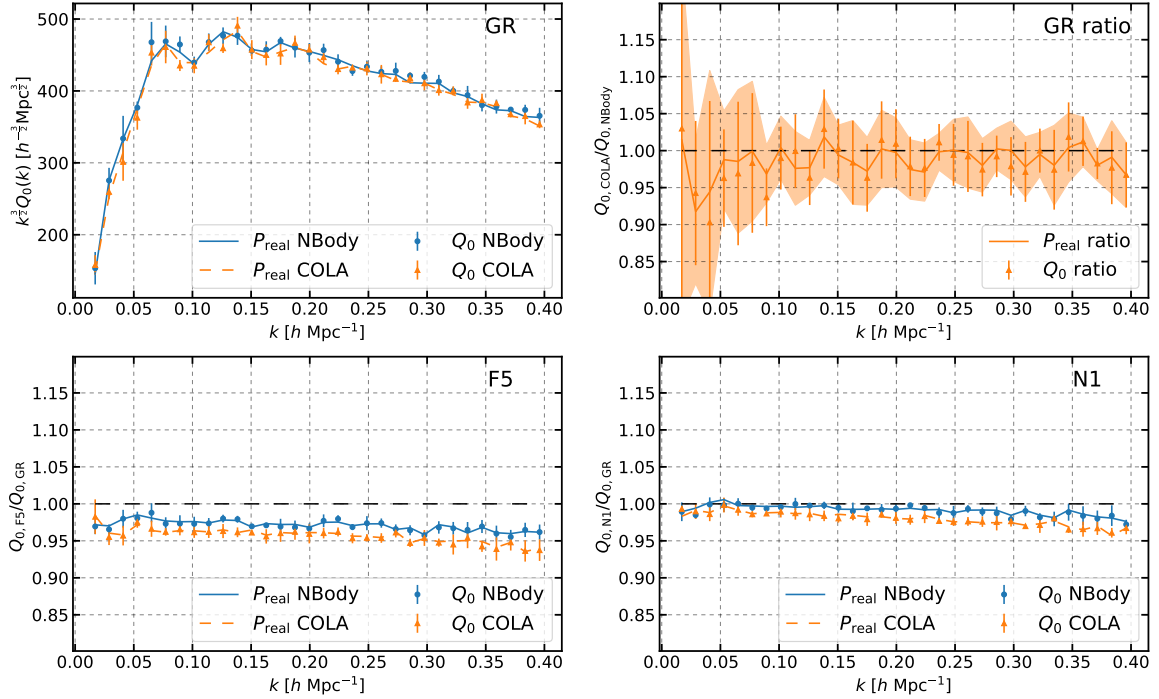
Since  $Q_0$  is unaffected by RSD, theoretical modelling is easier than for the redshift-space multipoles  $P_{\ell}$ , and this fact was used to extend the validity range of perturbation theory-based approaches such as EFTofLSS [29]. We limit our analysis to the  $Q_0$  estimate by means of the truncated sum of eq. (3.7), which we measure in our halo and galaxy catalogues and compare to the real space power spectrum to gain insight into the effect of MG on this observable and into the validity of COLA simulations to make theoretical predictions for  $Q_0$ .

### 3.2 Results for power spectrum

We compute the real space power spectrum and the multipoles in redshift-space in our halo and galaxy catalogues with the publicly available code `nbodykit`.<sup>3</sup> We assign the density to a  $512^3$  mesh grid with the Triangular Shaped Clouds (TSC) interpolation and we correct the resulting power spectra for shot noise and the effects of the interpolation used [33, 34]. We use bins of width  $\Delta k = 2k_f$  between  $k_{\min} = \frac{3}{2}k_f$  and  $k_{\max} = 0.4 h \text{ Mpc}^{-1}$ , where  $k_f$  is the fundamental frequency of the box. For the multipoles in redshift-space, we add RSD to the real space positions of tracers using in turn lines-of-sight parallel to each of the simulation box axes, and average the results over the three lines-of-sight. In the case of galaxy catalogues, we also average over the 5 HOD realisations for each simulation realisation. We then use the mean and standard deviation over the simulations' realisations to produce the signals and error bars that we plot in figure 1 and figure 2.

In figure 1 we compare  $Q_0$  to the real space power spectrum for halos. The top left panel shows that  $Q_0$  is in good agreement with  $P^r$  both in COLA and  $N$ -body halo catalogues. The top right panel shows that COLA and  $N$ -body are consistent within the variance for both  $Q_0$  and  $P^r$ . The bottom panels show that the MG boost factors for  $Q_0$  and  $P^r$  are consistent within the variance and that COLA is able to reproduce the  $N$ -body boost factors,

<sup>3</sup><https://nbodykit.readthedocs.io>.



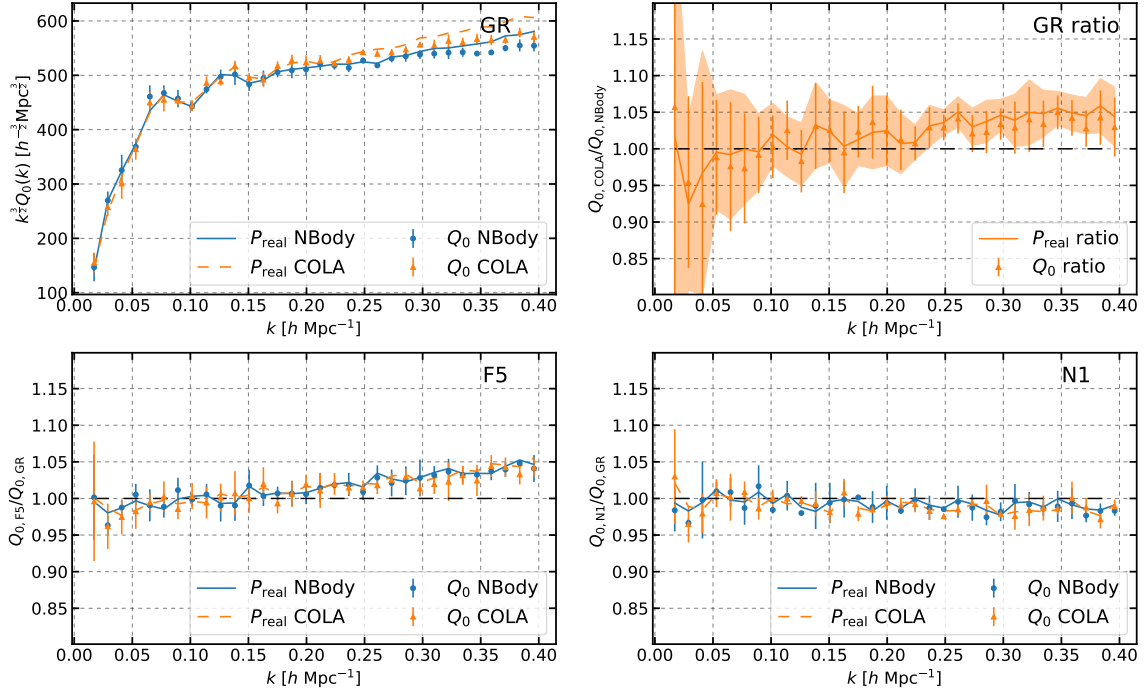
**Figure 1.** Comparison of the power spectrum orthogonal to the line-of-sight  $Q_0$  (dots and error-bars) and the real space power spectrum  $P^r$  (lines and shaded regions) for halos in COLA (in orange) and  $N$ -body simulations (in blue). *Top left:* full signal in GR. *Top right:* ratio of the GR signal between COLA and  $N$ -body simulations. *Bottom:* boost factors, the ratio of  $Q_0$  in F5 and N1 to that in GR.

the ratio of the power spectrum between MG and GR, with  $\sim 2\%$  accuracy at all scales in both F5 and N1. The MG signal in F5 is in general stronger than in N1 where the signal is compatible with GR within the variance.

The comparison between  $Q_0$  and  $P^r$  for the galaxy catalogues is shown in figure 2. The top left panel shows that  $Q_0$  is in good agreement with  $P^r$  up to  $k \sim 0.25 h \text{Mpc}^{-1}$  in COLA and up to  $k \sim 0.3 h \text{Mpc}^{-1}$  in  $N$ -body galaxy catalogues, where it starts deviating due to the truncation of the sum in eq. (3.7). The top right panel shows that COLA and  $N$ -body are consistent within the variance for both  $Q_0$  and  $P^r$  up to  $k \sim 0.24 h \text{Mpc}^{-1}$  where COLA starts deviating from  $N$ -body. The bottom panels show that the MG boost factors for  $Q_0$  and  $P^r$  are consistent within the variance at all scales as are COLA and  $N$ -body results. We note that, in spite of the limits of the truncation in eq. (3.7) and of the precision of COLA in reproducing the  $N$ -body power spectrum in GR, the MG boost factors are fully consistent at all scales. Overall the MG signal is quite weak in N1 while it reaches the 5% in F5 on small scales.

This suggests that the observable  $Q_0$  can be an interesting statistic to include in cosmological analysis to constrain MG theories and that COLA can be used to accurately predict the MG boost factors of  $Q_0$ . Furthermore, the good agreement of the MG boost factors of  $P^r$  and  $Q_0$  estimated using only the multipoles  $P_0, P_2, P_4$  suggests that higher order multipoles can be modelled using the GR theory without significantly impacting the MG signal.





**Figure 2.** Same as figure 1 but for galaxies.

## 4 Bispectrum

The bispectrum is the simplest higher-order summary statistics beyond the power spectrum and it holds non-gaussian information of the density field. The study of bispectrum also serves as a test for the screening approximation adopted in COLA simulations.

### 4.1 Definitions and measurements

We define the bispectrum as the three-point function in Fourier space for closed triangle configurations

$$\langle \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \delta(\mathbf{k}_3) \rangle \equiv (2\pi)^3 B(k_1, k_2, k_3) \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3). \quad (4.1)$$

and the reduced bispectrum as the ratio

$$Q(k_1, k_2, k_3) = \frac{B(k_1, k_2, k_3)}{P(k_1)P(k_2) + P(k_2)P(k_3) + P(k_3)P(k_1)}, \quad (4.2)$$

where  $P(k)$  is the power spectrum [35]. In the absence of primordial non-gaussianity, the bispectrum at the leading order in perturbation theory shows a  $P^2$  scaling which is removed in the reduced bispectrum to highlight the information content beyond the power spectrum [36].

To measure the bispectrum in our simulations we use the bispectrum estimator proposed in [32] as implemented in the publicly available library FML.<sup>4</sup> The density is interpolated on a  $360^3$  cells grid with a 4th order interpolation scheme known as Piece-wise Cubic Spline (PCS) with interlacing [32, 34]. We estimate the bispectrum in bins of size  $\Delta k = 3k_f$ , where

<sup>4</sup><https://github.com/HAWinther/FML>.



$k_f = 2\pi/L$  is the fundamental frequency of the box, centring the first bin in  $k_{\min} = 3k_f$ . The shot noise, which is subtracted from the bispectrum signal, is calculated as

$$B^{\text{SN}}(k_1, k_2, k_3) = \frac{1}{\bar{n}} (P(k_1) + P(k_2) + P(k_3)) + \frac{1}{\bar{n}^2}, \quad (4.3)$$

with  $\bar{n}$  being the number density of tracers and  $P(k)$  their power spectrum [37].

To get an insight into the dependence of the bispectrum signals on the triangle shape, we bin all the configurations with  $k_1 > 0.1 h \text{ Mpc}^{-1}$  in a grid of  $x_2 \equiv k_2/k_1$ ,  $x_3 \equiv k_3/k_1$ , and average the signals weighted with the number of fundamental triangles falling in each bin

$$\bar{B}(x_2^i, x_3^j) = \frac{\sum_{k_1=k_{\min}}^{k_{\max}} \sum_{k_2, k_3}^{\text{bin}} B(k_1, k_2, k_3) N^{\text{T}}(k_1, k_2, k_3)}{\sum_{k_1=k_{\min}}^{k_{\max}} \sum_{k_2, k_3}^{\text{bin}} N^{\text{T}}(k_1, k_2, k_3)}, \quad (4.4)$$

where  $N^{\text{T}}$  is the number of fundamental closed-triangles

$$N^{\text{T}}(k_1, k_2, k_3) = \prod_{i=1}^3 \int_{k_i} d^3 q_i \delta_{\text{D}}(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3) \sum_{n_1, n_2, n_3} \delta_{\text{D}}(\mathbf{q}_1 - n_1 k_f) \delta_{\text{D}}(\mathbf{q}_2 - n_2 k_f) \delta_{\text{D}}(\mathbf{q}_3 - n_3 k_f), \quad (4.5)$$

and similarly for the reduced bispectrum

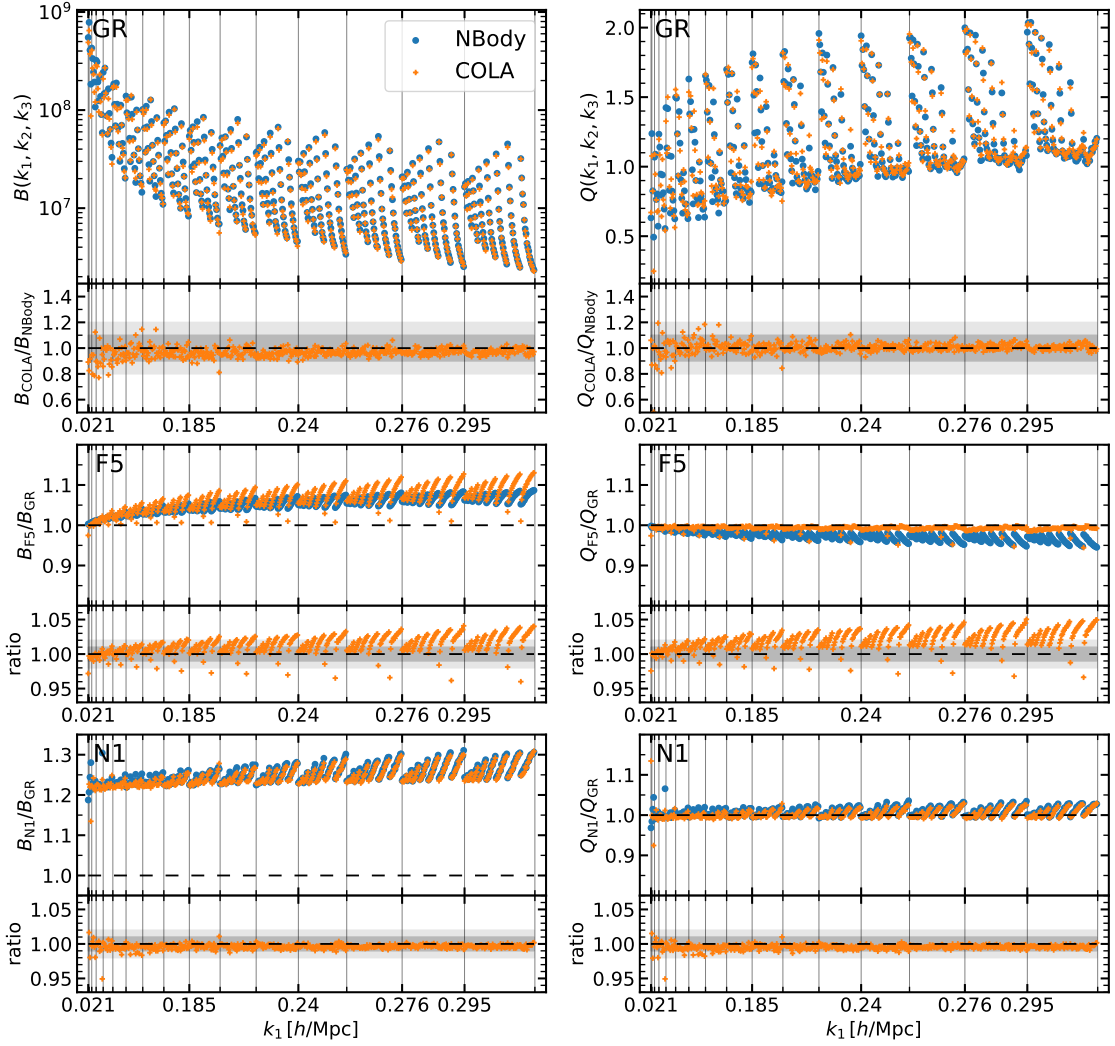
$$\bar{Q}(x_2^i, x_3^j) = \frac{\sum_{k_1=k_{\min}}^{k_{\max}} \sum_{k_2, k_3}^{\text{bin}} Q(k_1, k_2, k_3) N^{\text{T}}(k_1, k_2, k_3)}{\sum_{k_1=k_{\min}}^{k_{\max}} \sum_{k_2, k_3}^{\text{bin}} N^{\text{T}}(k_1, k_2, k_3)}. \quad (4.6)$$

## 4.2 Results for bispectrum

Since the bispectrum is a function of three variables, it is non-trivial to represent it in 2D plots, so we follow an approach often used in literature and represent all the configurations of the bispectrum in a single scatter plot by restricting our analysis to the configurations with  $k_1 \geq k_2 \geq k_3$  and organising them in ascending order of the values of the wavenumbers  $k_1$ ,  $k_2$ , and  $k_3$ , respectively. We show all the configurations of the bispectrum (and the reduced bispectrum) in figure 3 and figure 4 for DM and galaxies in redshift-space respectively. We show also the plots for halos and galaxies in real space in appendix A. In these figures, the left column shows the full bispectrum signal and the right column the reduced bispectrum. We split the results into three rows, the top one shows the full signal in GR, the middle one shows the boost factor in F5 (i.e. the ratio between F5 and GR) and the bottom one shows the boost factor in N1. To benchmark results with respect to full  $N$ -body, each row contains a sub-panel at the bottom showing the ratio between the COLA and  $N$ -body signals.

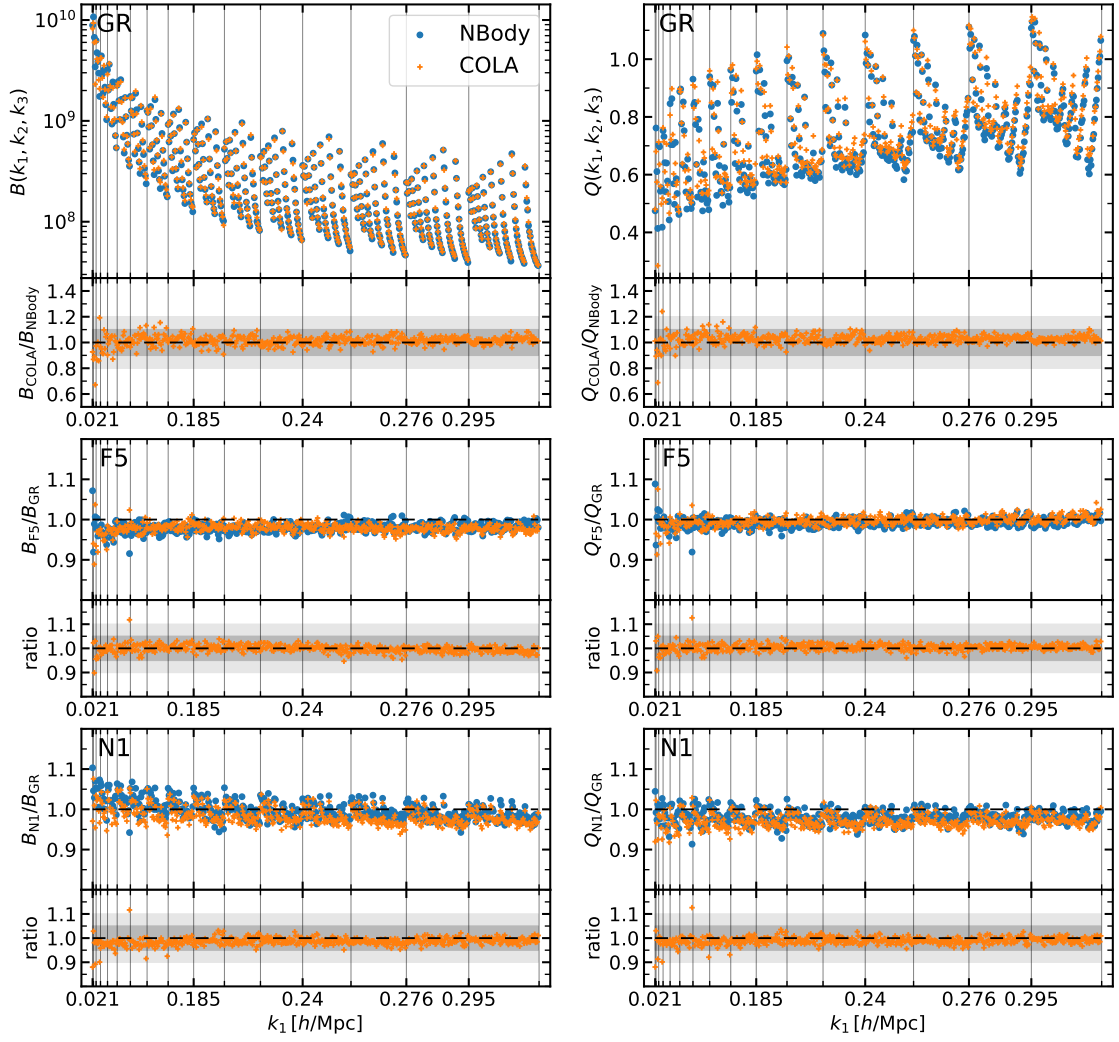
The MG boost factors of  $\bar{B}$  and  $\bar{Q}$  (from eq. (4.4) and eq. (4.6) respectively), that is, the ratio of  $\bar{B}$  and  $\bar{Q}$  between MG and GR, are used to produce figure 5 for DM and figure 6 for galaxies in redshift-space (see appendix A for equivalent plots of halos and galaxies in real space). In each figure, the left panels display the full bispectrum and the right panels the reduced bispectrum. We show the results for F5 in the top row and the results for N1 in the bottom row. The colour represents the amplitude of the MG signal, according to the colour bar.

At the DM level, figure 3 shows that COLA is able to reproduce  $N$ -body results in GR with  $\sim 10\%$  accuracy both for the bispectrum and the reduced bispectrum in agreement with the findings in [38]. Focusing on the MG boost factors in  $N$ -body, we notice that the strong MG signal of the bispectrum, reaching  $\sim 30\%$  in N1 and  $\sim 10\%$  in F5, is suppressed in the reduced bispectrum. This shows that most of the bispectrum signal is



**Figure 3.** Comparison of bispectrum (left column) and reduced bispectrum (right column) of the matter distribution in COLA (orange crosses) and  $N$ -body (blue dots) simulations. The configurations with  $k_1 \geq k_2 \geq k_3$  are displayed in ascending order of the values of the wavenumbers  $k_1$ ,  $k_2$ , and  $k_3$  respectively. The vertical lines denote the value of  $k_1$  for the configurations immediately to the right of each line. *Top*: full signal in GR. *Middle and Bottom*: boost factors in F5 and N1 respectively. In each panel, the bottom sub-panel shows the ratio between the COLA and  $N$ -body signals displayed in the top sub-panel.

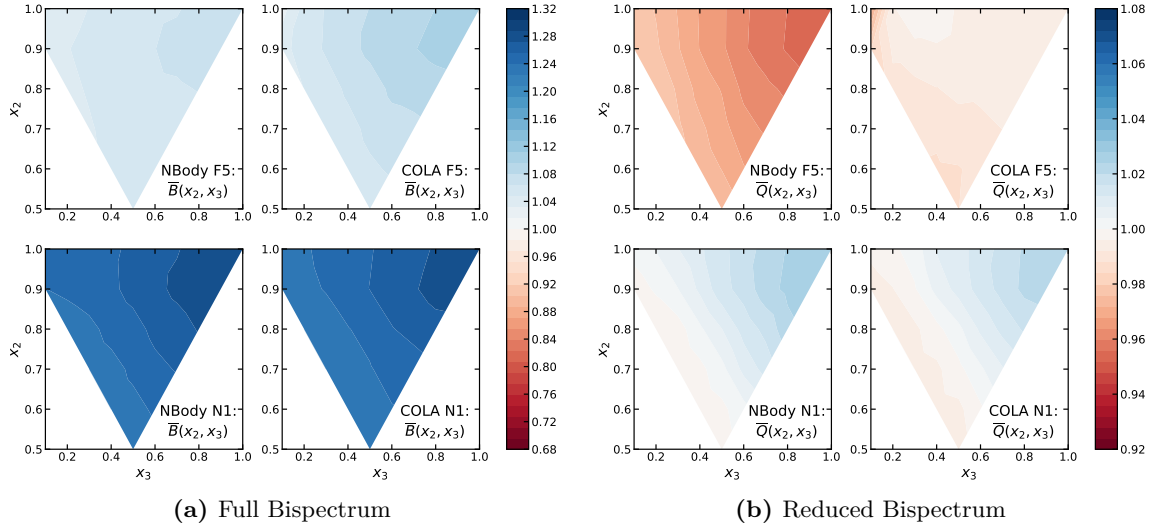
due to the enhancement of the power spectrum in MG. The MG boost factors of COLA do not reproduce the information content beyond the power spectrum as highlighted by the reduced bispectrum signal in the case of F5, while it reproduces the  $N$ -body signals with  $\sim 2\%$  accuracy in N1. The failure of COLA simulations in F5 gravity can be traced back to the use of the screening approximation, which allows us to recover with good accuracy the power spectrum but produces a DM bispectrum signal consistent with the results of [39] in the case of simulations without the chameleon mechanisms (see their figure 1). This is because the screening approximation used in COLA simulations linearises the scalar field equation and thus it lacks the contribution to the bispectrum from the non-linearity of the



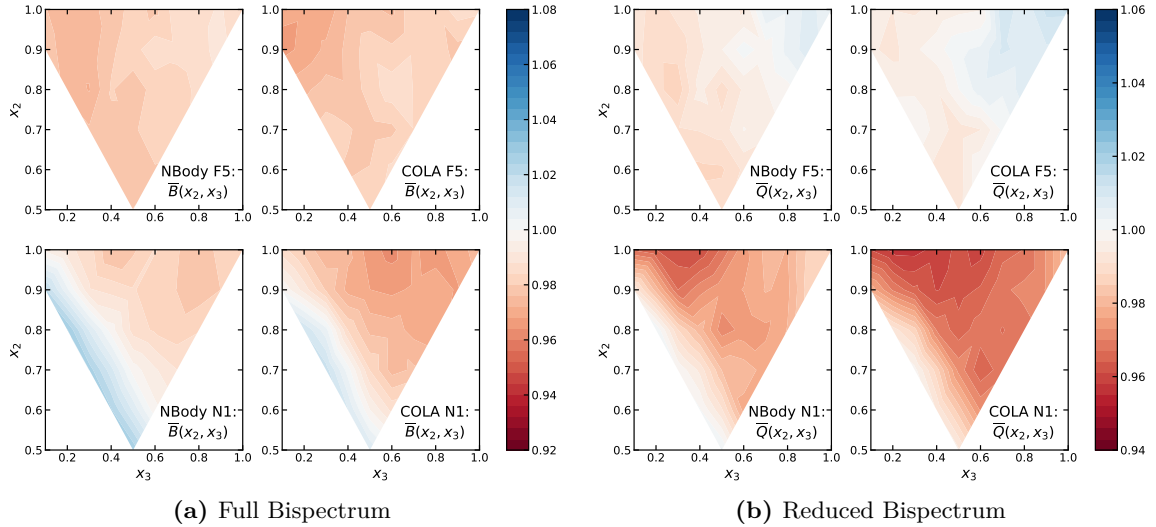
**Figure 4.** Same as figure 3 but for the monopole of bispectrum and reduced bispectrum of galaxies in redshift-space.

scalar field. This effect is stronger in F5, where the effect of screening is strong even on quasi-non-linear scales, while it is weak in N1. Concerning the configuration dependence of the DM bispectrum, figure 5 shows that the bispectrum boost factor signals of DM are stronger for equilateral configurations. COLA results reproduce the same configuration dependence of  $N$ -body in N1, but the reduced bispectrum boost factor in F5 clearly shows the failure of COLA simulations in reproducing non-trivial 3-points interactions, due to the screening approximation used in COLA.

Figure 4 shows the bispectrum monopole of the galaxy catalogues in redshift-space. The agreement between COLA and  $N$ -body is within 10% in GR for both the bispectrum and the reduced bispectrum. In the F5 and N1 gravity models, the bispectrum boost factors show deviations from GR at  $\sim 5\%$  level for some configurations, and the same is true for the reduced bispectrum. COLA reproduces  $N$ -body results for the boost factors with at least 5% accuracy. The much weaker MG signal of the redshift-space galaxy bispectrum monopole compared to the DM bispectrum is attributable to the HOD parameters tuning



**Figure 5.** Comparison of the configuration dependence of the F5 (top row) and N1 (bottom row) boost factors of bispectrum (on the left) and reduced-bispectrum (on the right) of DM in  $N$ -body (first and third columns) and COLA simulations (second and fourth columns). The colour bars show the amplitude of the boost factors with blue (red) denoting stronger (weaker) signal in MG with respect to GR. *In each panel:* the top right, top left and bottom corners of the triangle correspond to the equilateral, squeezed and folded configurations respectively. The squeezed configuration is missing from the figure due to the bin's width.



**Figure 6.** Same as figure 5 but for the monopole of the bispectrum and reduced bispectrum of galaxies in redshift-space.

procedure, which tries to reproduce the target multipoles of the galaxy power spectrum, in this case the GR one [15]. This tuning produces a power spectrum contribution to the bispectrum that is similar for GR and MG theories (as can be seen from the boost-factors of the bispectrum in the left column) and, on the other hand, dilutes the MG signal of the bispectrum coming from higher-order contributions of DM clustering due to the nonlinearity of galaxy bias (as evidenced by the boost-factors of the reduced bispectrum in the right

column). This influence of galaxy bias is also visible in figure 6 that shows the configuration dependence of galaxy bispectra boost factors substantially differs from what is found in the DM bispectra boost factors. This explains why the configuration dependence in COLA is qualitatively in agreement with the  $N$ -body result. The contribution from the non-linear galaxy bias dominates over the DM non-linearity hiding the inaccuracy of COLA simulations in reproducing the DM bispectra in F5.

The bispectra of halos and galaxies in real space are affected by biases in a similar fashion to the monopole of the galaxy bispectrum in redshift space. Likewise, they do not show strong MG signals. For an equivalent analysis on the bispectra of halos and galaxies in real space see appendix A.

Our results show that the MG bispectrum signal is mainly due to MG effects on the power spectrum and less to the higher order terms carrying specific information of the gravity model as shown in [39] for  $f(R)$  gravity. Extra care needs to be taken when using COLA simulations with screening approximations to compute the reduced bispectrum of dark matter, in particular in  $f(R)$  gravity models as the non-linearity of the scalar field due to screening is neglected. This is relevant for the computation of higher-order statistics in weak lensing, for example. On the other hand, in galaxy clustering, there are large contributions from non-linear galaxy biases to the bispectrum, and this inaccuracy in the reduced bispectrum in COLA is hidden. The bispectrum, when combined with the power spectrum, is useful to break degeneracies in the galaxy biases and COLA simulations can be used to model accurately galaxy bispectrum in MG models, despite the use of screening approximations.

## 5 Voids

Voids are low-density regions of space that are useful for cosmological analysis because the weaker gravitational field produces a phase-space distribution that is better described by linear models [40, 41]. Also, screening mechanisms do not operate efficiently in voids, thus voids can be used to test models of gravity [23, 42–48]. Observationally we cannot directly probe DM, so we have to rely on galaxies to identify voids, keeping in mind that galaxies are biased tracers of the underlying density field and that we only know the redshift-space position of galaxies.

Many void definitions have been proposed in the literature, differing in the properties (e.g. abundances and profiles) of the voids they identify. Some void definitions can reveal to be more suited than others in highlighting specific features of the underlying galaxy and DM distributions as discussed in many works (see for example [23, 41, 46, 49, 50]). In particular, the watershed void finding technique (implemented in various codes such as ZOBOV [51] and VIDE [52]) has been shown to identify (real-space) voids whose (stacked) galaxy density profiles in redshift-space can be approximated by linear models sensitive to the linear growth rate of structure [40, 41]. As MG theories may affect the linear growth rate  $f$ , we want to test if  $f$  can also be recovered in MG theories using the RSD model for the void-galaxy Cross-Correlation Function (CCF) in [41].

In subsection 5.1 and 5.2 we review the RSD model from [41, 53], and in subsection 5.3 we perform void finding on our simulation suites and measure the void quantities relevant to the RSD model. Finally, in subsection 5.4 we fit the multipoles of the void-galaxy CCF in redshift-space using the RSD model to recover the linear growth rate  $f$ . Comparing the measurements and results for the fit between COLA and  $N$ -body simulations, we validate COLA simulations for void analysis.

### 5.1 Redshift-space distortion model in voids

Let  $\xi^r(\mathbf{r})$  be the real-space CCF between voids and galaxies. Assuming a bijective differentiable mapping between real and redshift-space  $\mathbf{s} = \mathbf{f}(\mathbf{r})$ , the redshift-space CCF  $\xi^s(\mathbf{s})$  is related to the real-space CCF by

$$\int_V [1 + \xi^s(\mathbf{s})] d^3s = \int_V [1 + \xi^r(\mathbf{r})] d^3r. \quad (5.1)$$

Performing the change of variable  $\mathbf{r} \rightarrow \mathbf{s}$  on the right-hand side we obtain

$$\int_V [1 + \xi^s(\mathbf{s})] d^3s = \int_V [1 + \xi^r(\mathbf{r}(\mathbf{s}))] J_{\mathbf{r},\mathbf{s}} d^3s, \quad (5.2)$$

where  $J_{\mathbf{r},\mathbf{s}}$  is the determinant of the Jacobian matrix  $\frac{\partial \mathbf{r}}{\partial \mathbf{s}}$ . Requiring that this equation holds for every arbitrary volume  $V$  implies

$$[1 + \xi^s(\mathbf{s})] = [1 + \xi^r(\mathbf{r}(\mathbf{s}))] J_{\mathbf{r},\mathbf{s}}. \quad (5.3)$$

This relation can be manipulated to express the redshift-space CCF  $\xi^s$  in terms of  $\xi^r$ , once the mapping between real and redshift-space is fixed, determining the two unknowns  $\mathbf{r}(\mathbf{s})$  and  $J_{\mathbf{r},\mathbf{s}}$ . This mapping in the distant observer approximation can be described by

$$\mathbf{s} = \mathbf{r} + \frac{\mathbf{v} \cdot \hat{\mathbf{z}}}{aH} \hat{\mathbf{z}}, \quad (5.4)$$

where  $\mathbf{v}$  is the galaxy velocity,  $a$  is the scale factor, and  $H$  is the Hubble parameter. The void positions are by construction assumed to be invariant under redshift-space remapping, since the CCF in this model always refers to the correlation between real-space voids and galaxies in either real or redshift-space [41].

Assuming a spherically symmetric model for the galaxy velocity  $\mathbf{v} = v(r)\hat{\mathbf{r}}$ , the Jacobian determinant is given by

$$J_{\mathbf{r},\mathbf{s}} = \left[ 1 + \frac{v(r)}{aHr} + \frac{z^2 \frac{d}{dr} v(r)}{aHr^2} - \frac{z^2 v(r)}{aHr^3} \right]^{-1}, \quad (5.5)$$

with  $r = \|\mathbf{r}\| = \sqrt{x^2 + y^2 + z^2}$ , which can be rewritten as

$$J_{\mathbf{r},\mathbf{s}} = \left[ 1 + \frac{\bar{v}(r)}{r} + \mu_r^2 \frac{d\bar{v}(r)}{dr} - \frac{\mu_r^2 \bar{v}(r)}{r} \right]^{-1}. \quad (5.6)$$

where we defined  $\bar{v} \equiv \frac{v}{aH}$  and  $\mu_r \equiv \frac{z}{r}$  for simplicity of notation.

In addition to the spherically symmetric model for the galaxy velocity we consider a random velocity component along the line-of-sight,<sup>5</sup>  $\mathbf{v} = v_r(r)\hat{\mathbf{r}} + v_{\parallel}\hat{\mathbf{z}}$ , where  $P(v_{\parallel}) = \mathcal{N}(0, \sigma(r))$  as proposed in [41]. This makes the mapping between real and redshift-space depends also on  $v_{\parallel}$

$$\begin{aligned} \mathbf{s}(\mathbf{r}, v_{\parallel}) &= \mathbf{r} + \frac{(v_r(r)\hat{\mathbf{r}} + v_{\parallel}\hat{\mathbf{z}}) \cdot \hat{\mathbf{z}}}{aH} \hat{\mathbf{z}} \\ &= \mathbf{r} + \frac{(v_r(r)\mu_r + v_{\parallel})}{aH} \hat{\mathbf{z}}, \end{aligned} \quad (5.7)$$

<sup>5</sup>A more realistic model would be considering a random velocity component with random orientation (and dispersion 3 times larger), but the two models are equivalent for what concerns RSD.



but does not explicitly affect the Jacobian, as  $v_{\parallel}$  is an integration variable independent of the real-space radius  $r$ . We solve eq. (5.7) for  $\mathbf{r}$  iteratively to obtain  $\mathbf{r}(s, v_{\parallel})$  and integrate over  $v_{\parallel}$  to compute the theoretical prediction for the redshift-space CCF

$$\xi_{\text{th}}^s(s, \mu_s) = \int (1 + \xi^r) J_{\mathbf{r},s} e^{-\frac{v_{\parallel}^2}{2\sigma_v^2}} dv_{\parallel} - 1, \quad (5.8)$$

where  $\xi^r$ ,  $J_{\mathbf{r},s}$  and  $\sigma_v$  are expressed in terms of  $s$  and  $\mu_s$ . Finally we project  $\xi_{\text{th}}^s(s, \mu_s)$  on the Legendre polynomials to compute monopole and quadrupole of the void-galaxy CCF in redshift-space

$$\begin{aligned} \xi_{0,\text{th}}^s(s) &= \int_0^1 \xi_{\text{th}}^s(s, \mu_s) d\mu_s, \\ \xi_{2,\text{th}}^s(s) &= \frac{5}{2} \int_0^1 \xi_{\text{th}}^s(s, \mu_s) (3\mu_s^2 - 1) d\mu_s. \end{aligned} \quad (5.9)$$

In the comparisons of the void-galaxy CCF in redshift-space shown in subsection 5.3, we include these theoretical predictions without any further assumption to test the accuracy of this approach.

## 5.2 Velocity model

Galaxy velocities correlate with void positions, as galaxies are attracted towards overdensities far from the void centre. In an isotropic universe, despite that voids have in general very irregular shapes, the average matter distribution of stacked voids is isotropic and therefore spherically symmetric. The average velocity of galaxies around voids depends on the matter distribution and it is also spherically symmetric. This velocity can be estimated in simulations and it has been found to be well described down to scales of  $\sim 30$  Mpc/ $h$  by the simple linear model

$$v_r^{\text{model}}(r) = -\frac{1}{3} f a H r \Delta(r), \quad (5.10)$$

where  $f$  is the linear growth rate and  $\Delta(r) \equiv \frac{3}{r^3} \int_0^r \delta(r') r'^2 dr'$  is the integrated matter density profile [41, 54]. We will show in subsection 5.3 a comparison between this velocity model and the radial-velocity profile estimated from our simulations that confirms the same trend of [41] for the accuracy of the velocity model.

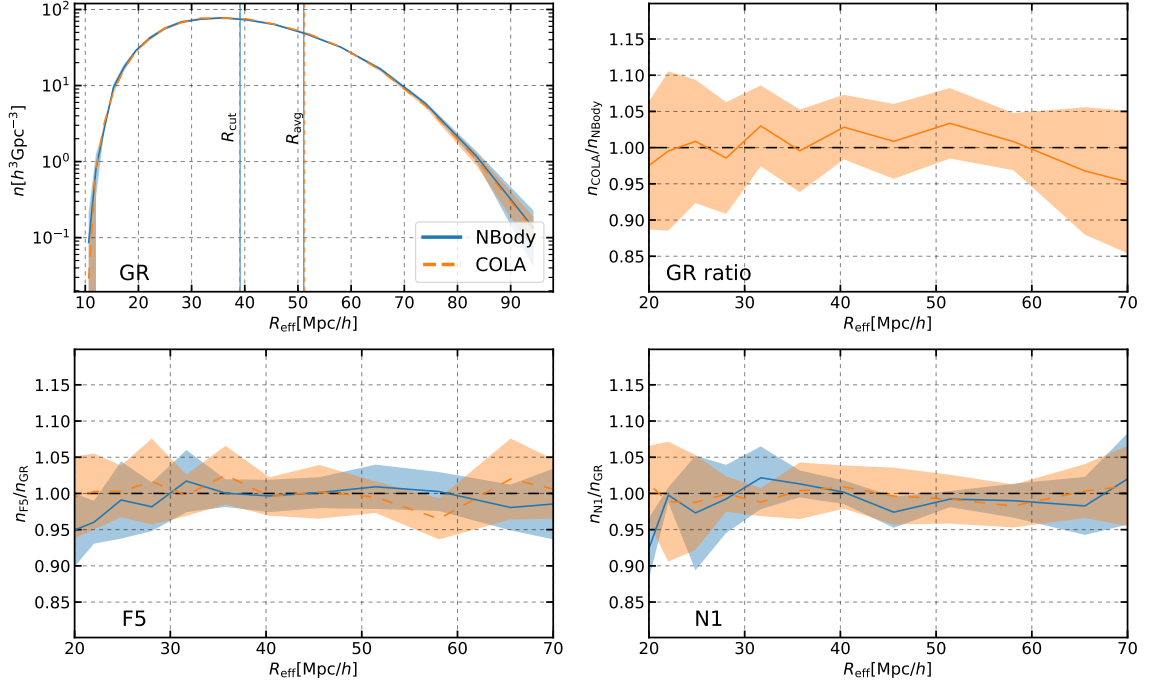
## 5.3 Void finding and measurements

To perform void-finding for the real-space galaxy mocks, we make use of the publicly available code **Revolver**<sup>6</sup> which implements two void-finding techniques: **ZOBOV** [51] and **voxel**. With the aim of performing a similar analysis to [41] we choose **ZOBOV** void-finding. This algorithm finds voids as the watershed basins of the density field of galaxies in real-space estimated using the Voronoi tessellation [51]. The void effective radius is determined by the void total volume  $V$  via the formula  $R_{\text{eff}} = (3V/4\pi)^{1/3}$ . We use this definition of void radius to count the voids in 20 bins logarithmically spaced between 10 and 100  $h^{-1}$  Mpc to obtain the void size function displayed in figure 7. This shows that the void-size functions in GR, N1 and F5 are all compatible within the variance, as are the void-size functions of COLA and  $N$ -body simulations.

We select a sub-sample of voids whose radius is larger than the median radius of the catalogue, which is sufficient to select voids that are uncorrelated and statistically isotropic [41].

<sup>6</sup><https://github.com/seshnadathur/Revolver>.





**Figure 7.** Comparison of the void size function in COLA (in orange) and  $N$ -body simulations (in blue). *Top left:* full signal in GR. The vertical lines denote the minimum and average void radii after applying the median cut. *Top right:* ratio of the GR signal between COLA and  $N$ -body simulations. *Bottom:* boost factors in F5 and N1. The panels involving ratios are shown in the range  $[20, 70] h^{-1} \text{ Mpc}$  as the signal-to-noise ratio deteriorates outside of this range.

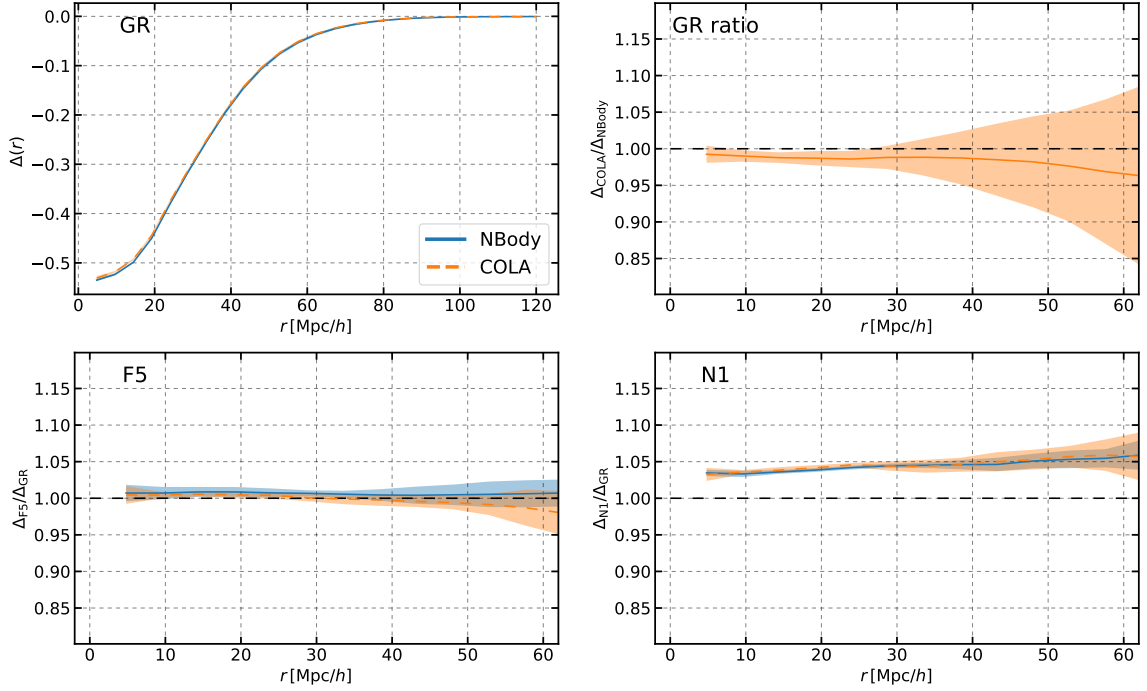
This results in void catalogues with a minimum void radius  $R_{\text{cut}} \simeq 39 h^{-1} \text{ Mpc}$  and an average void radius  $R_{\text{avg}} \simeq 51 h^{-1} \text{ Mpc}$  across all the simulations. We add the vertical lines in the top-left panel of figure 7 to highlight these two scales in the void-size function. The void centre is identified as the centre of the largest empty sphere inside the void volume. This is done by finding the circumcentre of the tetrahedron described by the 4 galaxies closest to the under-density peak [55]. Given the definition of the void centre, it is possible to study several quantities characterising the profile of voids. In particular, we focus on:

- the integrated matter density profile,  $\Delta(r) \equiv \frac{3}{r^3} \int_0^r \delta(r') r'^2 dr'$ , where  $\delta(r)$  is the matter density-contrast profile of stacked voids;
- the galaxy radial-velocity profile,  $v_r(r) = \langle \mathbf{v}_g \cdot \hat{r} \rangle$ ;
- the galaxy velocity-dispersion profile  $\sigma_{v_{\parallel}}(r) = \sqrt{\langle |\mathbf{v}_g|^2 \rangle - v_r^2} / \sqrt{3}$ ;
- the CCF of voids and galaxies in real and redshift-space,  $\xi^r$  and  $\xi^s$  respectively.

These quantities, relevant to the RSD model of the void-galaxy CCF discussed in subsection 5.1, are estimated using 25 linearly spaced radial bins from the void centre to  $120 h^{-1} \text{ Mpc}$ . To measure the integrated-density profile and the CCF we use the publicly available code `pyCUTE`<sup>7</sup> based on `CUTE`<sup>8</sup> [56]. In the case of real-space (redshift-space) statistics, the signal is first averaged over the 5 HOD realisations (and over the three lines of sight),

<sup>7</sup><https://github.com/seshnadathur/pyCUTE>.

<sup>8</sup><https://github.com/damonge/CUTE>.



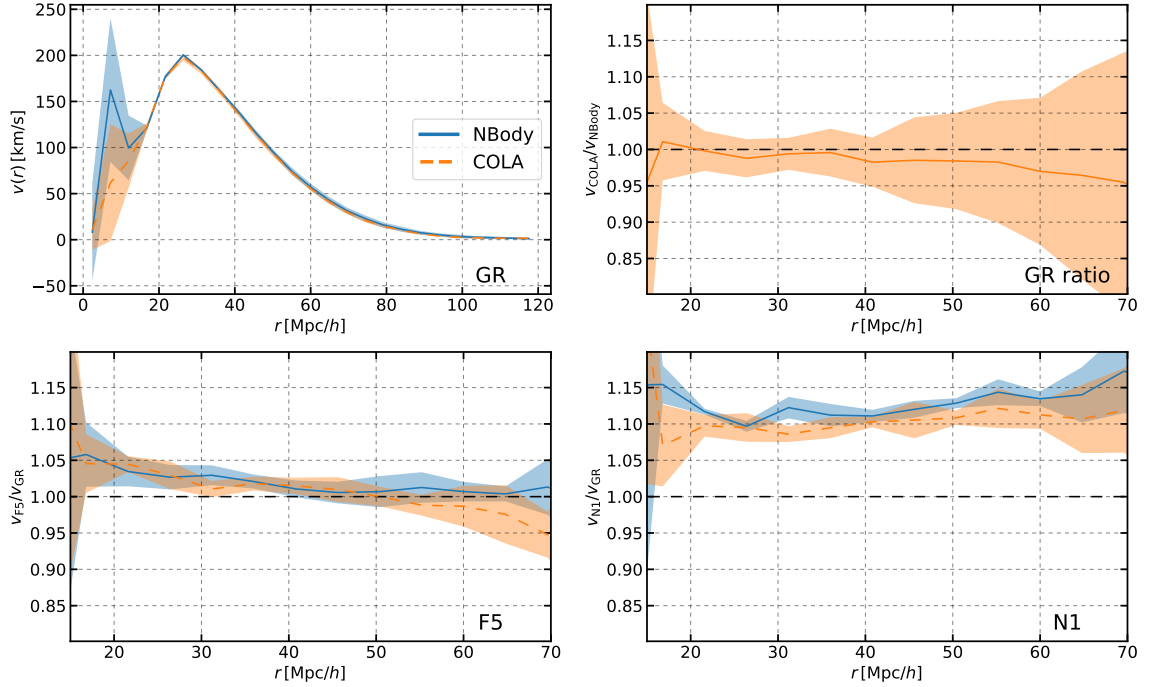
**Figure 8.** Comparison of the integrated matter density profile of voids in COLA (in orange) and  $N$ -body simulations (in blue). *Top left:* full signal in GR. *Top right:* ratio of the GR signal between COLA and  $N$ -body simulations. *Bottom:* boost factors in F5 and N1. The panels involving ratios are shown up to  $r \simeq 60 h^{-1}$  Mpc as the signal-to-noise ratio deteriorates for larger separations.

and then the mean and standard deviations over the 5 boxes are used respectively for the lines and shaded regions in the figures.

A comparison of the integrated matter density profile measured in our simulations is shown in figure 8. The signal is characterised by a good signal-to-noise ratio in the void interior (due to the amplitude of the signal and the abundance of DM tracers) which deteriorates above  $\sim 60 h^{-1}$  Mpc due to the signal approaching the large-scales limit of 0. COLA and  $N$ -body results are very consistent (within the variance) in all gravity models. The MG signal in the integrated-density profile is weak in the case of F5 (compatible with GR) while it shows a  $\sim 5\%$  enhancement in the case of N1 gravity.

The average radial velocity of galaxies as a function of the separation between the galaxies and the void centre is represented in figure 9. The bins below  $15 h^{-1}$  Mpc are heavily affected by shot noise due to the low galaxy counts. On large scales, the signal approaching 0 is responsible for the large scatter in the ratios plots, which are therefore shown in the range  $[15, 70] h^{-1}$  Mpc. In GR, COLA is in agreement with  $N$ -body within the variance at all scales. A small scale-dependent MG signal is present in F5, with velocities larger than in GR for small separations and approaching GR results on large scales. The boost factor in N1 shows a  $\sim 10\%$  enhancement of the radial-velocity profile with respect to GR. The different behaviours of the velocity profiles in N1 and F5 are consistent with the different nature of their fifth force, long-ranged in N1, and short-ranged in F5.

Figure 10 displays the dispersion of the random galaxy velocity component along the line of sight. The large scatter in the first bins is again due to the low number of galaxies in the proximity of the void centre. The velocity-dispersion shows a mild dependence on separations



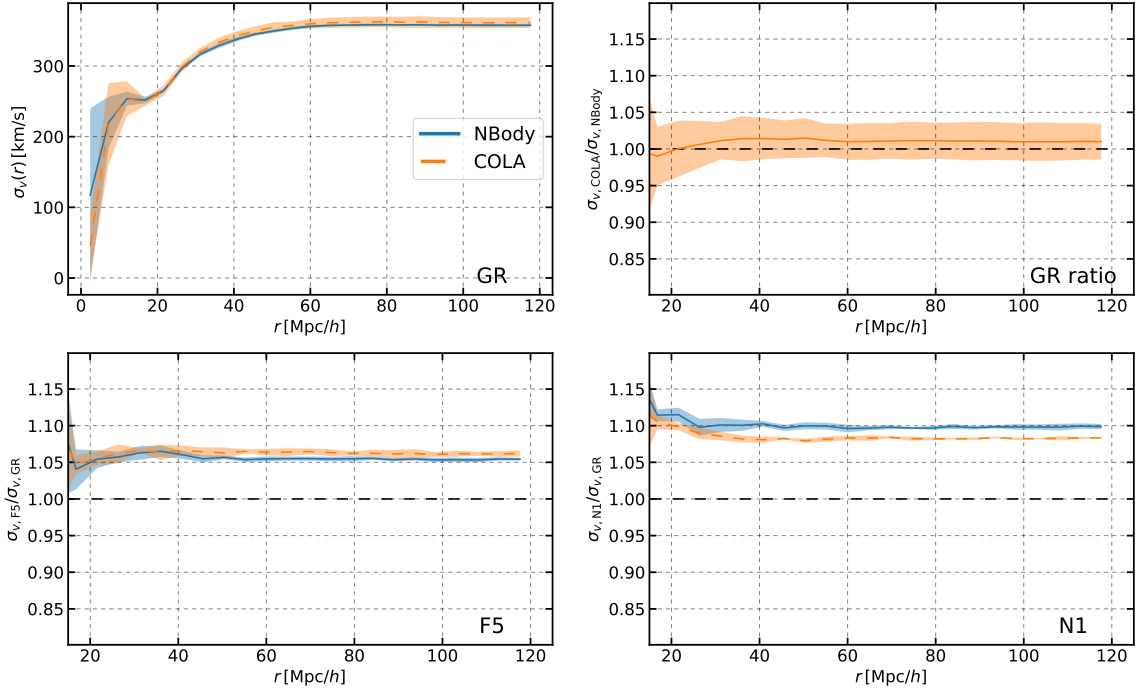
**Figure 9.** Comparison of the galaxy radial-velocity profile around voids in COLA (in orange) and  $N$ -body simulations (in blue). *Top left:* full signal in GR. *Top right:* ratio of the GR signal between COLA and  $N$ -body simulations. *Bottom:* boost factors in F5 and N1. The panels involving ratios are displayed in the range  $[20, 70] h^{-1} \text{ Mpc}$  as the signal-to-noise ratio deteriorates outside of the range.

at small separations and plateaus at  $\sim 370 \text{ km/s}$ . COLA results are in agreement with the  $N$ -body results in GR within the variance at all scales. The boost factors in F5 and N1 show a clear enhancement with respect to GR, of  $\sim 5\%$  and  $\sim 10\%$  respectively. The accuracy of COLA in reproducing the boost factors is  $1\%$  in F5 and  $2\%$  in N1.

The real-space void-galaxy CCF in figure 11 shows that COLA and  $N$ -body results are in agreement within the variance in all gravity models. Note that the top right panel shows the difference between COLA and  $N$ -body in GR instead of the usual ratio because the signal goes to zero on large scales. Similarly, the bottom panels show the differences between MG and GR results instead of the boost factors for the same reason. In both F5 and N1, there is no clear MG signal.

In figure 12 and figure 13 we compare the monopole and the quadrupole of the void-galaxy CCF in redshift-space. In the top panels, the monopoles of the CCF (“data”, black dots) are consistent within the variance (shaded regions) with the ones predicted using eq. (5.9) and the radial-velocity measured in the simulations (“RSD model”, orange lines). This proves the accuracy of the RSD model discussed in section 5.1. On the other hand, using the radial-velocity from eq. (5.10)<sup>9</sup> in eq. (5.9) (“RSD + vel. model”, blue lines) produces theoretical predictions that are still a good description of the data but with some discrepancies arising in the void interior that are traceable to the inaccuracies of the velocity model in reproducing the radial-velocity profile of galaxies in voids. The inaccuracies of the velocity model are visible in figure 14 where, comparing the velocity model with the radial-velocity profile estimated from  $N$ -body simulations in GR, we confirm the results of [41] that the

<sup>9</sup>The value of the linear growth rate  $f$  is taken from the linear theory.

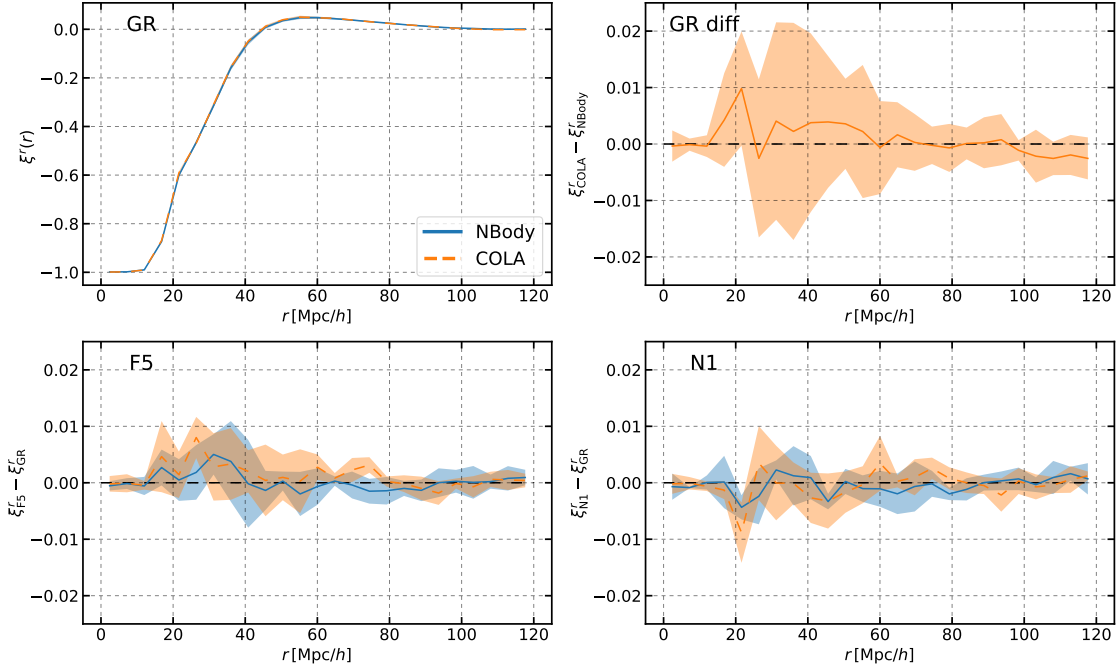


**Figure 10.** Comparison of the galaxy velocity-dispersion profile around voids in COLA (in orange) and  $N$ -body simulations (in blue). *Top left:* full signal in GR. *Top right:* ratio of the GR signal between COLA and  $N$ -body simulations. *Bottom:* boost factors in F5 and N1. The panels involving ratios are displayed down to  $15 h^{-1}$  Mpc as the signal-to-noise ratio deteriorates for smaller separations.

model is a good description of the galaxy radial-velocity profile above  $\sim 30 h^{-1}$  Mpc with some discrepancies arising below that distance from the void centre.

The bottom two rows of figure 12 and figure 13 show the impact of MG on monopole and quadrupole of the void-galaxy CCF in redshift space. In the case of F5, the signal is compatible with GR within the variance. Interestingly, N1 shows a clear enhancement of the distortions with respect to GR. This signal is well reproduced by using eq. (5.9) with the radial-velocity model computed in terms of  $f^{N1}$  from linear theory and  $\Delta^{N1}(r)$  and  $\sigma_{v_{\parallel}}^{N1}(r)$  estimated from N1 simulations (green lines). However, even ignoring the effects of MG on  $\Delta(r)$  and  $\sigma_{v_{\parallel}}(r)$  (i.e. using  $\Delta^{\text{GR}}(r)$  and  $\sigma_{v_{\parallel}}^{\text{GR}}(r)$  from GR simulations) but using the growth factor  $f^{N1}$  from N1 linear theory (blue lines) produces theoretical predictions that catch most of the signal (although representing a worse fit to the data). This is somehow surprising as we have seen from figure 8 and figure 10 that N1 gravity has impacts on both the integrated-density profile and the velocity-dispersion profile. To gain insight into this aspect, we add to the comparison the theoretical predictions formulated using the growth factor  $f^{N1}$  from N1 linear theory and, in one case  $\Delta^{\text{GR}}(r)$  and  $\sigma_{v_{\parallel}}^{N1}(r)$  (red lines), while in the other  $\Delta^{N1}(r)$  and  $\sigma_{v_{\parallel}}^{\text{GR}}(r)$  (violet lines). Comparing the two cases we can see that they have somehow competing effects on the redshift-space CCF and this explains why the main effect of N1 gravity is due to the enhancement of the linear growth factor.

For the monopole and quadrupole of the void-galaxy CCF in redshift-space (figure 12 and figure 13 respectively), results in COLA (right column) are consistent with the ones from  $N$ -body (left column) in all gravity models.



**Figure 11.** Comparison of the galaxy-void real-space CCF in COLA (in orange) and  $N$ -body simulations (in blue). *Top left:* full signal in GR. *Top right:* difference of the GR signal between COLA and  $N$ -body simulations. *Bottom:* difference between the CCF in MG and GR for F5 and N1 gravity. The comparison uses differences instead of usual ratios since the signal crosses 0.

#### 5.4 Results for voids

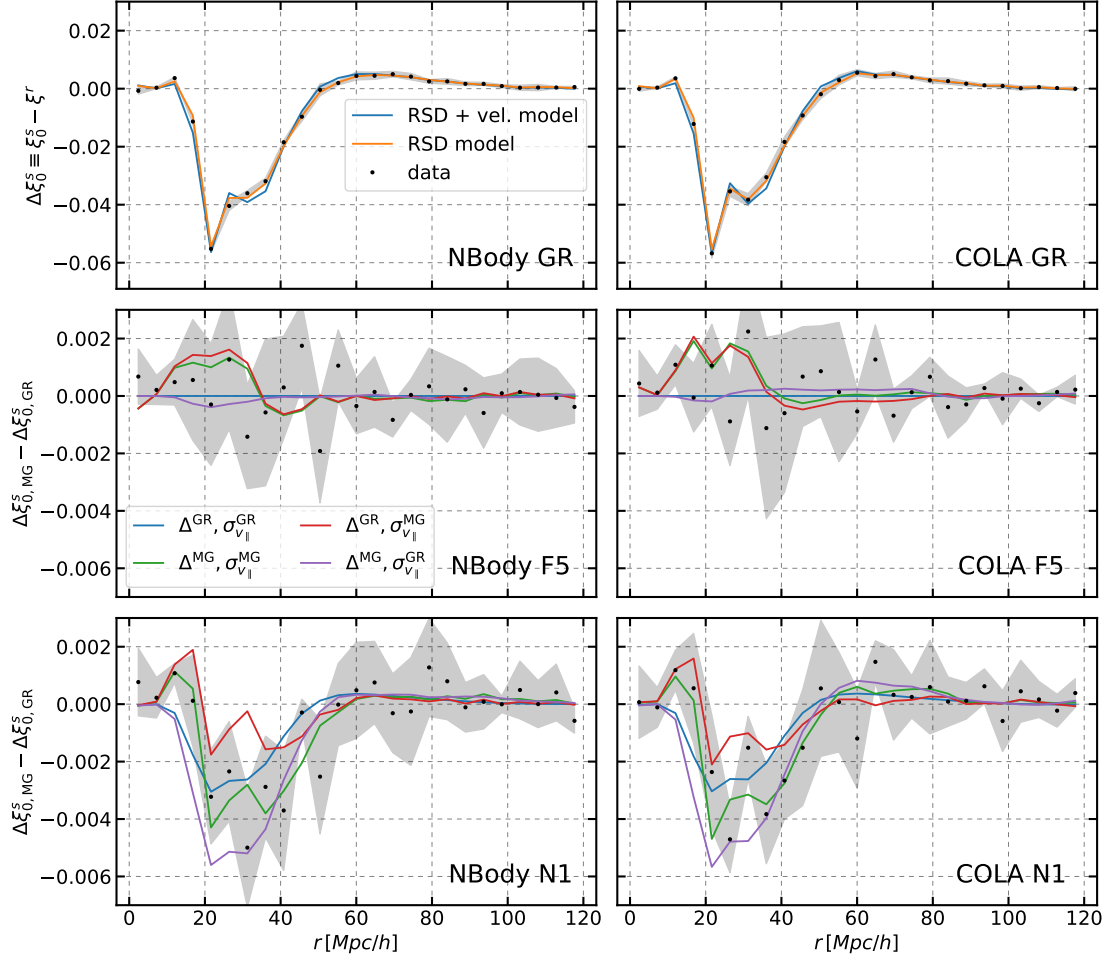
Using the RSD model discussed in subsection 5.1 with the velocity model of subsection 5.2, we leave the linear growth rate  $f$  in eq. (5.10) as a free parameter and we test if we can recover the growth rate predictions from linear theory. We apply the delete-1 jackknife re-sampling technique [57–59] to  $n_{\text{sv}} = 125$  sub-volumes in which we split each realisation and average over the  $n_r = 5$  realisations to estimate the covariance matrix for the multipoles of the CCF in redshift-space

$$\text{Cov}_{\ell, \ell', i, j} = \frac{1}{n_r} \sum_{m=1}^{n_r} \left[ \frac{n_{\text{sv}} - 1}{n_{\text{sv}}} \sum_{k=1}^{n_{\text{sv}}} \left( \xi_k^s - \bar{\xi}^s \right)_{\ell, i} \left( \xi_k^s - \bar{\xi}^s \right)_{\ell', j} \right]_m, \quad (5.11)$$

where  $\bar{\xi}^s = \frac{1}{n_{\text{sv}}} \sum_{k=1}^{n_{\text{sv}}} \xi_k^s$  is the average over the jackknife realisations of the CCF  $\xi_k^s$  estimated using 5 HOD realisations and 3 lines of sight. The indices  $\ell, \ell'$  run over the multipoles  $\{0, 2\}$ , while  $i, j$  run over the radial bins. Given the periodic boundary conditions of the simulation box, we use the natural 2-point estimator

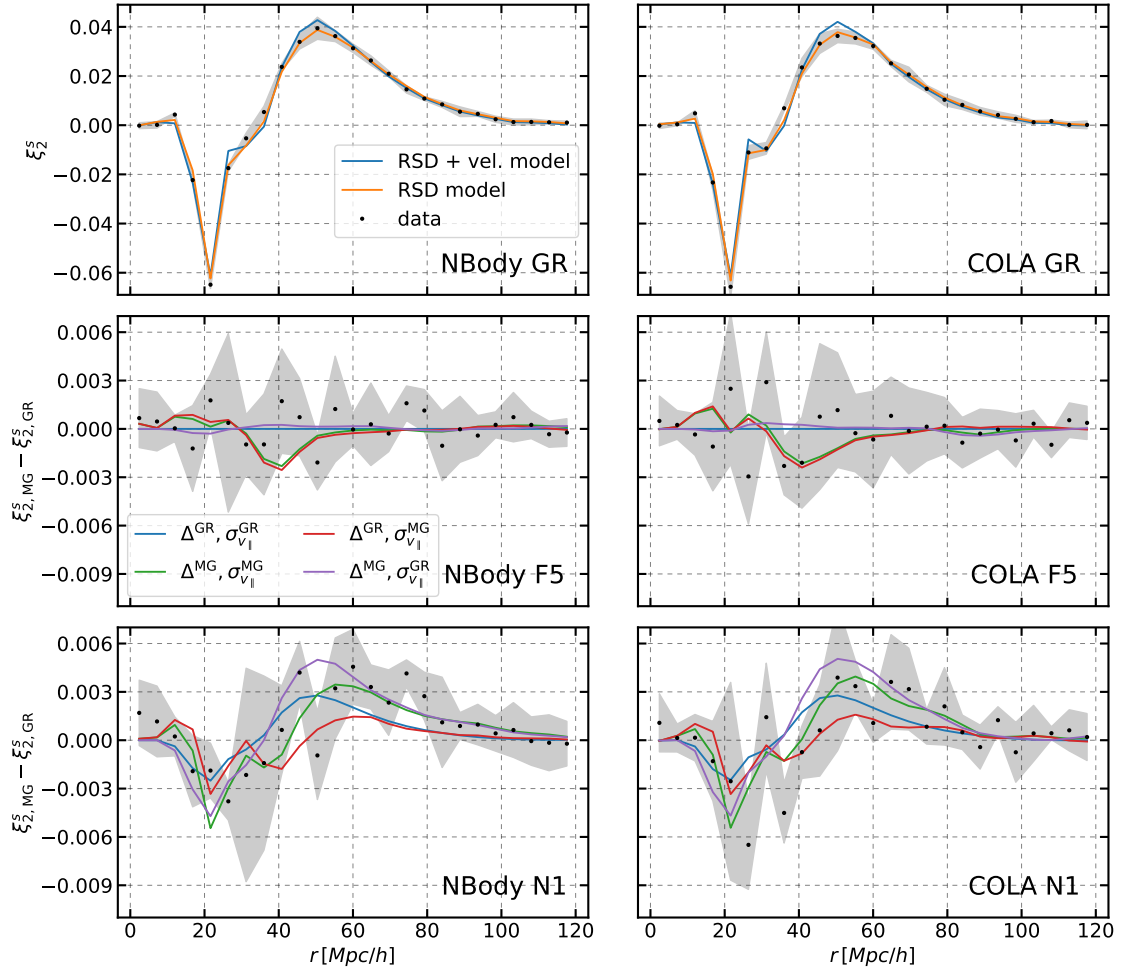
$$\hat{\xi}_{\text{N}} = \frac{DD}{RR} - 1. \quad (5.12)$$

The correlation matrix obtained from the covariance matrix (figure 15) shows strong off-diagonal correlations between different bins in the monopole while the quadrupole is almost diagonal. From a visual comparison of our correlation matrix with the one in figure 6 of [41] we notice that in our case the off-diagonal correlations in the quadrupole are smaller. This is due to the fact that we average the signal over three lines of sight while [41] was using only

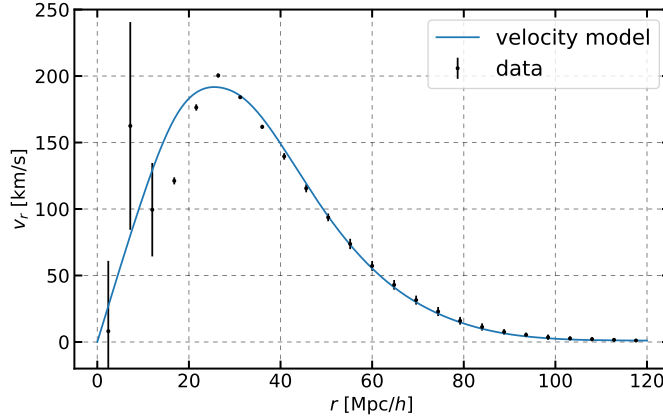


**Figure 12.** Comparison of the monopole of the CCF of voids and galaxies in redshift-space between  $N$ -body (left column) and COLA (right column) simulations. *Top panels:* the difference between the redshift-space monopole of the CCF and the real-space CCF,  $\Delta\xi_0^s \equiv \xi_0^s - \xi^r$ , measured in simulations (black dots and shaded region) is compared with the theory predictions (lines) formulated using the radial-velocity profile from simulations ("RSD model", orange) and using the velocity model of eq. (5.10) ("RSD + vel. model", blue). *Middle and bottom panels:* the difference between  $\Delta\xi_0^s$  in MG and GR (black dots and shaded regions) is compared to the theory predictions (lines) formulated using different combinations of inputs: integrated-density and velocity-dispersion from GR simulations (blue), integrated-density from MG and velocity-dispersion from GR simulations (violet), integrated-density from GR and velocity-dispersion from MG simulations (red), both integrated-density and velocity-dispersion from MG simulations (green). In all cases, we use the real-space CCF from GR simulations in the theory predictions, even in the MG panels to reduce the noise on the MG signals. The velocity model of eq. (5.10) is evaluated with the reference growth rate from linear theory (see table 3) in GR and N1 panels. In the case of F5, we use the reference growth rate from the GR linear theory.

1 line of sight. Using only 1 line of sight likely produces a more realistic error estimate for a real survey but, since our goal in this work is to test the consistency of the RSD model in beyond- $\Lambda$ CDM models and to validate approximate simulations, we choose to leverage on the average over three lines of sight (as well as the average over 5 HOD realisations) to increase the signal-to-noise ratio.

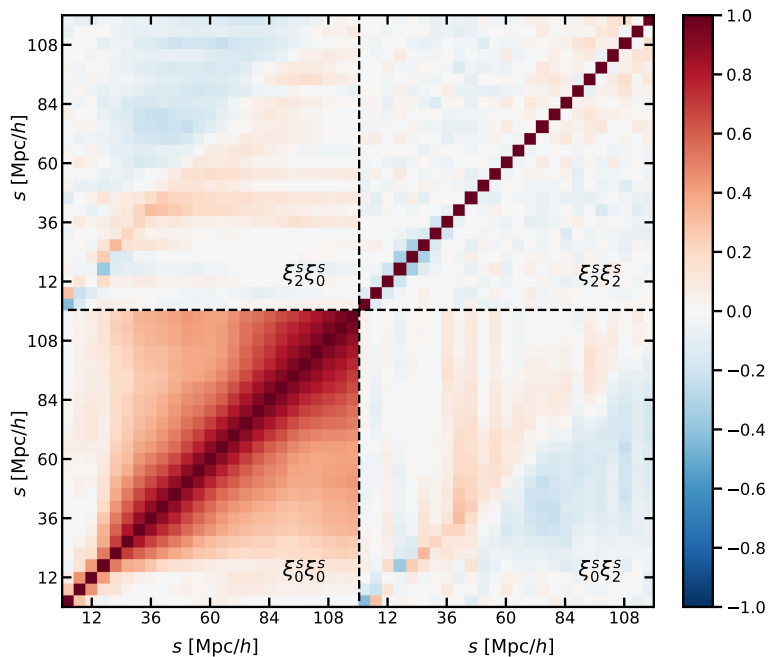


**Figure 13.** Same as figure 12 but for the quadrupole.



**Figure 14.** Comparison of the galaxy radial-velocity profile in GR  $N$ -body simulations with the linear model in eq. (5.10) using the fiducial growth rate  $f = 0.736$ , as a function of the distance from the void centre.





**Figure 15.** Correlation of the covariance matrix estimated from jackknife re-sampling in  $N$ -body simulations in GR. The black dashed lines are used to distinguish between the monopole (bottom/left) and quadrupole (top/right) terms. The correlation matrices for the other gravity models as well as the ones for  $N$ -body simulations appear visually equivalent.

The recent work [60], has proposed the use of the Landy-Szalay 2-point estimator [61]

$$\hat{\xi}_{\text{LS}} = \frac{DD - 2DR + RR}{RR}, \quad (5.13)$$

to deal with the mask applied by construction by the jackknife re-sampling, as well as a correction to account for proper weighting of the “cross-pairs” (pairs with only one point falling in the sub-volume masked by the jackknife), which has been shown to produce more realistic covariance matrix estimates for the galaxy correlation function. In light of these results, we compute the jackknife covariance using the publicly available code `pycorr`<sup>10</sup> and notice that:

- Using the Landy-Szalay estimator produces smaller off-diagonal correlations in the covariance matrix for the monopole and larger errors on scales smaller than  $\sim 20 h^{-1} \text{ Mpc}$ . The amplitude of the errors on these scales seems to depend on the number density of the random catalogues used (having tested for  $20\times$  and  $50\times$  the average number density of our galaxy and void catalogues) and shrinks for larger number density. This is compatible with shot-noise effects.
- Using the cross-pairs correction shrinks the errors on large scales in the monopole but does not significantly affect the correlations of the covariance matrix.

Since this technique has not been tested yet against mocks in the specific case of the void-galaxy CCF, we make the conservative choice of adopting the same technique originally used

<sup>10</sup><https://github.com/cosmodesi/pycorr>.

		GR		F5		N1	
		COLA	<i>N</i> -body	COLA	<i>N</i> -body	COLA	<i>N</i> -body
$f_{\text{ref}}$		0.736		–		0.777	
GR theory	$f_{\text{fit}}$	0.721	0.726	0.717	0.726	0.772	0.784
	$\sigma_f$	0.012	0.012	0.013	0.013	0.013	0.013
	$\Delta f/\sigma_f$	–1.2	–0.8	–	–	–0.4	0.5
	$\chi^2/\text{d.o.f.}$	0.96	1.08	0.88	0.92	1.53	1.69
MG theory	$f_{\text{fit}}$	–	–	0.735	0.734	0.763	0.774
	$\sigma_f$	–	–	0.013	0.013	0.013	0.013
	$\Delta f/\sigma_f$	–	–	–	–	–1.1	–0.2
	$\chi^2/\text{d.o.f.}$	–	–	0.67	0.77	0.78	0.95

**Table 3.** Results of the optimisation of eq. (5.14) and eq. (5.15) using theory predictions from GR simulations (top) and using theory predictions from MG simulations (bottom). COLA and *N*-body results are compared side by side in each gravity model (different columns). The additional row at the top of the table shows the linear theory predictions for the linear growth rate at redshift  $z = 0.5057$  in GR and N1. The linear theory growth rate in F5 is scale-dependent so we omit it from the table.

to validate the RSD model in voids [41], i.e. using the natural 2-point estimator in eq. (5.12) and removing all the pairs involving an object falling in the masked sub-volume.

With the covariance obtained with these settings, we perform the fit for the linear growth rate  $f$  by minimising the objective function

$$\chi^2 = \sum_{\ell, \ell'} \sum_{i, j} (\xi_{\ell, i}^s - \xi_{\ell, i}^{\text{th}}) \mathbf{Cov}_{\ell, \ell', i, j}^{-1} (\xi_{\ell', j}^s - \xi_{\ell', j}^{\text{th}}), \quad (5.14)$$

where  $\xi^s$  is obtained by averaging over the 3 lines of sight, the 5 HOD realisations and the 5 realisations for each gravity model in COLA and *N*-body simulations. The theoretical predictions are based on  $\xi^r$ ,  $\Delta$  and  $\sigma_{v_{\parallel}}$  estimated from the average of the signal over all the realisations in each gravity model in COLA and *N*-body simulations. We estimate the error on the best fit growth rate by finding the solutions to

$$\chi^2(f \pm \sigma_f) = \chi^2(f_{\text{fit}}) + 1. \quad (5.15)$$

The results of the fit are shown in table 3. In GR, both COLA and *N*-body recover the reference linear growth rate within  $\sim 1$  standard deviation. In the case of N1 and F5 theories we investigate the impact of MG on the RSD model by performing the fit with different theoretical inputs:

- using  $(\xi^r, \Delta, \sigma_{v_{\parallel}})^{\text{GR}}$  estimated in GR simulations (“GR theory”);
- using  $(\xi^r, \Delta, \sigma_{v_{\parallel}})^{\text{MG}}$  estimated in MG simulations (“MG theory”).

In both cases, we recover the reference growth rates within  $\sim 1$  standard deviation in N1. In F5, we do not have a reference growth rate to compare our results with since the linear theory solution is scale-dependent, but the best fit values are consistent with the reference value in GR. Interestingly, comparing best-fit values obtained using “MG theory” and “GR theory”, we notice a shift in the growth rate of  $\sim 1\sigma_f$  in both COLA and *N*-body simulations,

towards larger values in the case of F5 and towards smaller values in the case of N1 using the GR theory. We also find an increase of the reduced  $\chi^2$  in the GR theory. If this shift is in fact due to the effects of MG on the components entering the theoretical model, we might need to take into account these effects to get an accurate determination of the linear growth rate with  $\sigma_f \lesssim 0.01$  precision with the void-galaxy CCF. At that level of precision, however, the inaccuracies of the velocity model of subsection 5.2 (see also figure 14) could become the limiting factor [50]. In such a scenario, simulation-based emulation of the void radial-velocity profile may be a way to avoid systematic errors of theoretical origin, and the COLA method could certainly be considered a viable option given its computational efficient nature and its accuracy in reproducing full  $N$ -body results as proven by this work. With less precision on the growth rate estimate, e.g.  $\sigma_f \simeq 0.02$ , voids analysis can still help constrain nDGP theories by providing independent information on the amplitude of the linear growth rate. In the particular case of N1, MG effects on the integrated-density profile and on the velocity-dispersion can be ignored in the light of the cancellation discussed in the previous subsection, but in general, this cancellation depends on the specific gravity model (and screening mechanism) and it needs to be carefully checked whether the same conclusion applies to other theories of gravity.

We performed the fit also using the covariance matrices corrected for the cross-pairs and/or computing the jackknife CCF with the Landy-Szalay estimator and confirmed that we obtained the same results.

## 6 Conclusions

In this work, we used the simulation output of [15] to investigate the effects of MG on the estimator for the real space power spectrum  $Q_0$ , bispectrum and voids while validating COLA simulations for these probes. In section 3.1 we compared the power spectrum orthogonal to the line-of-sight estimated by means of the truncated sum of eq. (3.7),  $Q_0$ , with the real space power spectrum  $P^r$  in our halo and galaxy catalogues. For halos, we showed in figure 1 that  $Q_0$  agreed within the variance with the real space power spectrum in the full range of scales considered and for all gravity models. By comparing COLA and  $N$ -body results, we found that the two were in agreement within the variance in GR and within  $\sim 2\%$  in the MG boost factors, the ratio between GR and MG. For galaxies, we showed in figure 2 that the agreement between  $Q_0$  and  $P^r$  deteriorated after  $k \sim 0.25 h \text{ Mpc}^{-1}$  in COLA and after  $k \sim 0.3 h \text{ Mpc}^{-1}$  in  $N$ -body, meaning that higher multipoles need to be considered to recover  $P^r$  at smaller scales. However, we found that the MG boost factors of  $Q_0$  and  $P^r$  were in excellent agreement at all scales, indicating that most of the MG information is contained in the first three even multipole moments. We determined that the accuracy of COLA is within the variance up to  $k \sim 0.24 h \text{ Mpc}^{-1}$  in GR and at all scales in the MG boost factors.

We examined the bispectrum of DM and galaxies in redshift space in section 4, to understand how much of the MG bispectrum information filters through the HOD tuning. We compared results for bispectrum and reduced bispectrum to discriminate the part of the bispectrum signal due to the power spectrum from the one due to non-linearity. At the DM level, we showed in figure 3 that there were significant MG signals in the bispectrum, however, most of the MG signals in the bispectrum boost factors in F5 and N1 are due to the enhanced MG power spectrum as seen from the fact that the signal was much weaker in the reduced bispectrum. We found that COLA achieved  $\sim 10\%$  accuracy in reproducing the  $N$ -body results in GR and  $\sim 2\%$  in the N1 boost factor. In F5, we discovered that COLA simulations

did not accurately reproduce the  $N$ -body bispectrum boost factors. We concluded that they lacked the contribution coming from non-linearity in the screening mechanism, which is neglected in the screening approximation used in COLA by linearising the scalar field equation. We highlighted this also by showing the configuration dependence of the matter bispectrum in figure 5, where COLA failed to reproduce the configuration dependence of  $N$ -body simulation results in F5, while it captured the  $N$ -body configuration dependence in N1 in which screening is weak on quasi non-linear scales. For galaxies in redshift-space, we have confirmed with figure 4 that HOD tuning can hide the strong MG signal in the DM bispectrum coming from the power spectrum, with the resulting MG signals deviating only  $\sim 5\%$  from GR. By comparing COLA and  $N$ -body results, we found that COLA's accuracy is better than 10% in GR and  $\sim 5\%$  in the MG boost factors. By looking at the configuration dependence of the boost factors for the bispectrum of galaxies in redshift space, we highlighted that this was dissimilar to the one observed in DM and concluded that the MG signal was dominated by non-linearity coming from galaxy bias. Due to this, the inaccuracy of COLA simulations in F5 for DM is largely hidden and COLA reproduced the configuration dependence of  $N$ -body simulations well.

In section 5 we analysed the voids obtained from COLA and  $N$ -body galaxy catalogues with the ZOBOV void-finder. In order to test the applicability of the linear model of [41] to the case of MG, we carefully tracked the MG effects in each key component involved in the modelling. We showed in figures 8, 9 and 10 that significant MG signals were present in the integrated-density, radial-velocity and velocity-dispersion profiles in N1, while in F5 only the velocity-dispersion profile showed a substantial enhancement compared to GR. We have not found any significant MG effect on the real space void-galaxy CCF, but by examining the monopole and quadrupole of the void-galaxy CCF in redshift space, we discovered that the RSD in N1 were significantly stronger than in GR, while in F5 they were not substantially different from the GR case. To understand the role of the linear growth rate, the integrated density and the velocity dispersion in the MG signal of the redshift-space void-galaxy CCF, we compared the predictions formulated using different combinations of theory components estimated from GR and MG simulations in figure 12 and figure 13. We concluded that the enhancements of the integrated-density profile and velocity-dispersion profile observed in N1 partially cancelled out and the linear growth rate played the leading role in enhancing the RSD in the void-galaxy CCF. Using the linear growth rate as a free parameter of the theory, we fitted the monopole and the quadrupole of the void-galaxy CCF in redshift space. We found that the best fit values were recovered within  $\sim 1\sigma$  of the reference value computed from the linear theory in GR and N1, even when using the integrated density and velocity dispersion from GR simulations in the theory predictions. Concerning the accuracy of COLA in reproducing  $N$ -body results, we determined that COLA and  $N$ -body were consistent within the variance for all the void summary statistics analysed in GR and for the MG boost factors with the exception of the velocity-dispersion boost factor in N1 where COLA's accuracy is at  $\sim 2\%$  level. We have found that the best fit values of the linear growth rate in COLA were consistent within  $\sim 1\sigma$  with  $N$ -body. As for  $N$ -body simulations, it was possible to recover the linear growth rate reference values within  $\sim 1\sigma$ , even when integrated density and velocity dispersion are taken from GR simulations. We note that we used voids identified in real space galaxy mocks so that the real space void positions were known. When we use real observational data, they need to be recovered by applying a density-field reconstruction method [62]. The reconstruction requires an assumption on the gravitational theory and there will be an additional MG effect in this process.

The use of multi-probes analysis will be key in breaking the degeneracies of the theory parameter space, enabling tight constraints of gravity on cosmological scales with next-generation experiments. The findings of this work are important to understanding the potential of  $Q_0$ , bispectrum and void-galaxy CCF in reducing the freedom of the galaxy bias model and in increasing the sensitivity to the gravity model. Furthermore, the validation of COLA simulations for these probes gives us a computationally cheaper alternative to full  $N$ -body simulations to formulate reliable theoretical predictions, bearing in mind the limitations of the approximations used in COLA simulations. This opens up the possibility to train beyond- $\Lambda$ CDM cosmological emulators for these additional observables in sight of Stage IV surveys.

## Acknowledgments

We thank Seshadri Nadathur and Sladana Radinović for useful insights and stimulating discussions on the void analysis. We thank Baojiu Li for providing the data from ELEPHANT simulations. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this work. AI and KK are supported by the U.K. Science and Technology Facilities Council (grant numbers ST/S000550/1 and ST/W001225/1). Numerical computations were done on the Sciama High Performance Compute (HPC) cluster which is supported by the ICG, SEPNet and the University of Portsmouth.

**Data availability.** Supporting research data are available on reasonable request from the corresponding author.

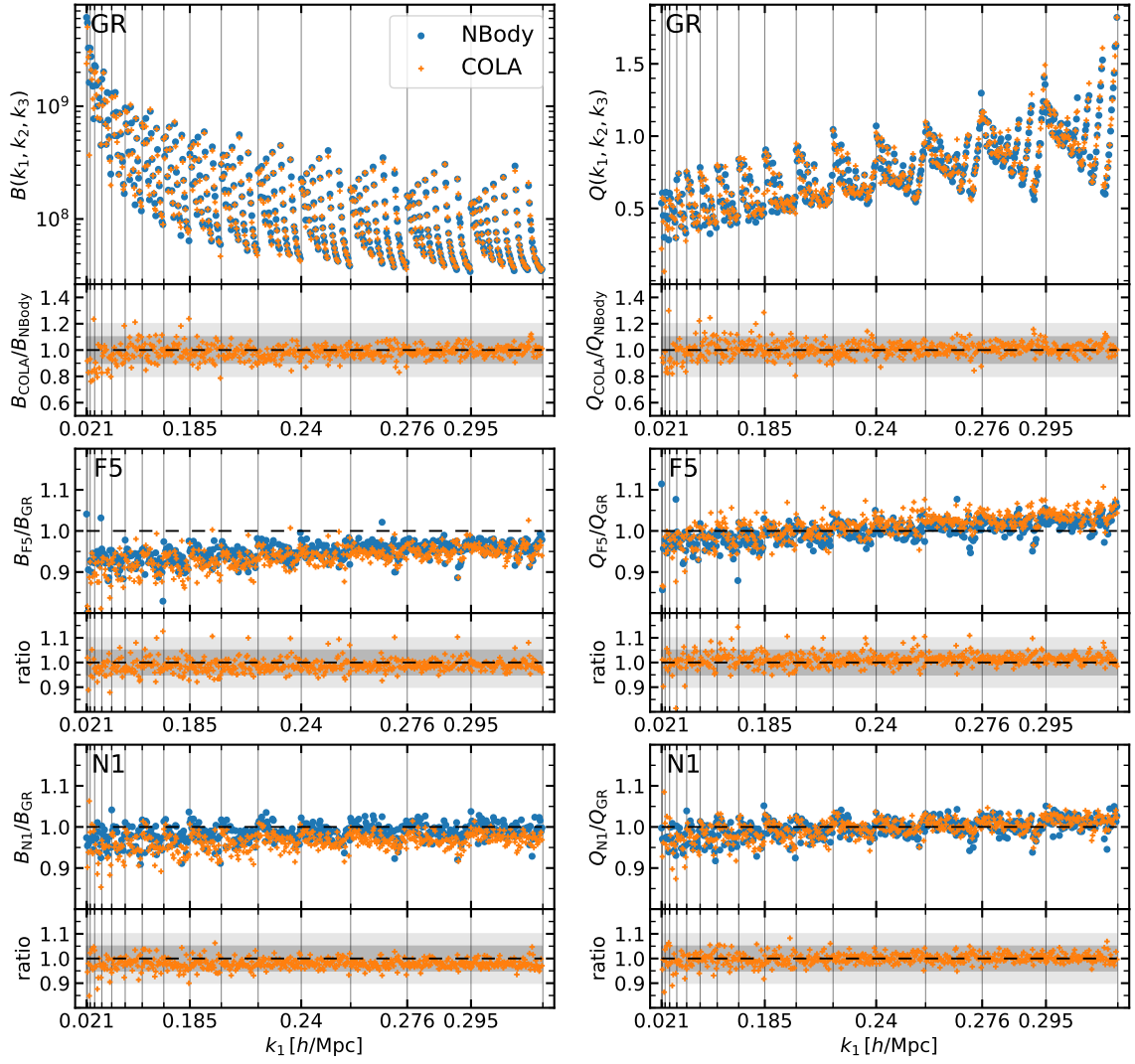
## A Bispectrum of halos and galaxies in real space

For completeness, we add to the analysis of section 4 the results for the bispectra of halos and galaxy in real space.

In the case of halo bispectrum, figure 16 shows that COLA is able to reproduce the  $N$ -body signals for both bispectrum and reduced bispectrum of halos with 10% accuracy in most of the configurations. The MG signal reaches  $\sim 10\%$  in the case of the full bispectrum in F5, while it is limited to  $\sim 5\%$  in the other cases. For the boost factors, COLA's accuracy is better than  $\sim 5\%$  for most of the configurations.

The configuration dependence of the bispectra boost factors of halos is shown in figure 18. Due to the bias of halos, the configuration dependence of DM in figure 5 is not reflected in the configuration dependence of the halo bispectra boost factors. Thanks to this, the configuration dependence of COLA in F5 is qualitatively consistent with the one of  $N$ -body halo catalogues for both full and reduced bispectrum, despite the limitations in reproducing the matter bispectra discussed in section 4.

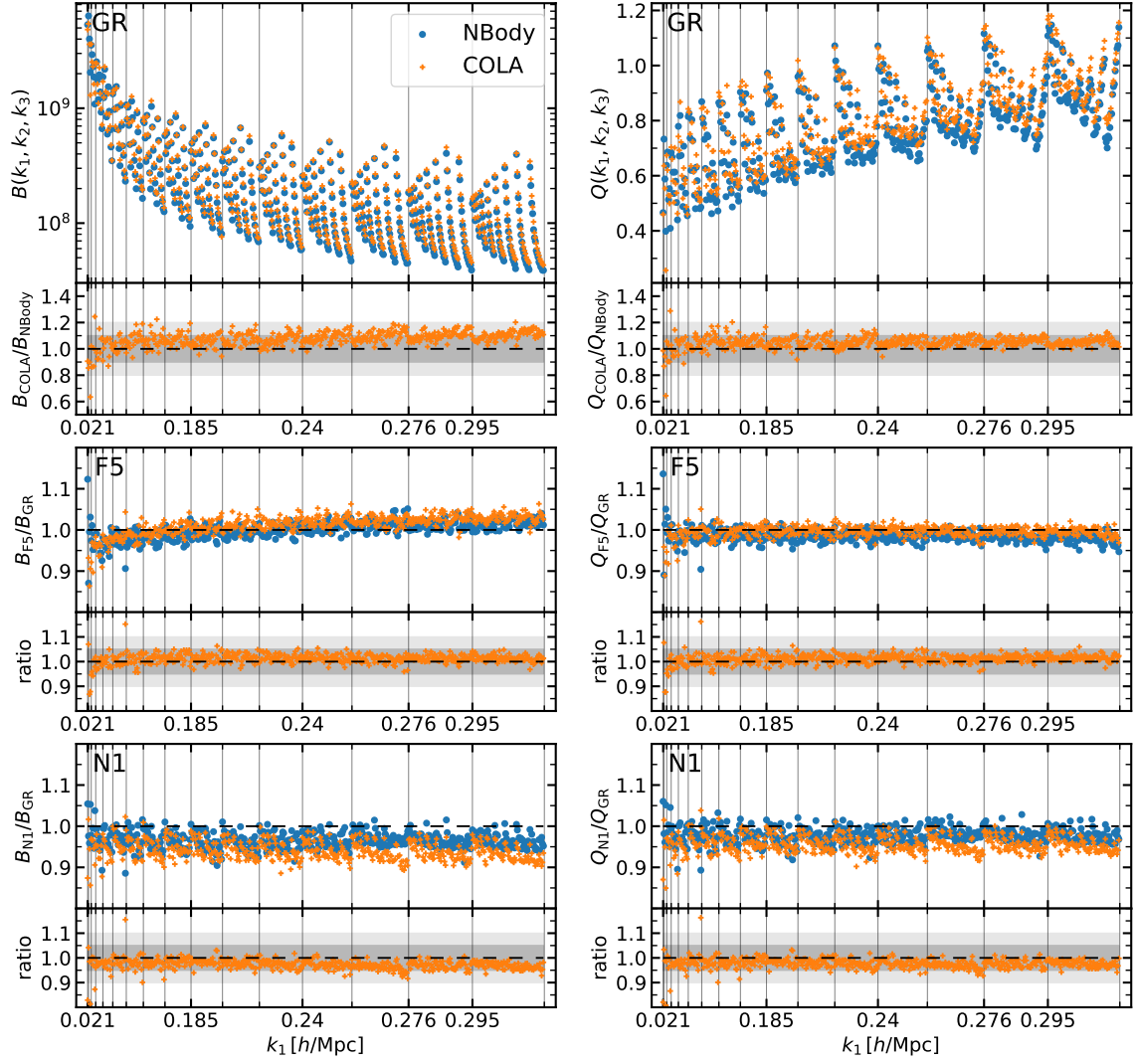
The galaxy bispectrum in real-space is shown in figure 17. In GR, COLA is able to reproduce the  $N$ -body signals with 20% accuracy for the full bispectrum and 10% accuracy for the reduced bispectrum. The MG signals reach  $\sim 10\%$  in the case of the full bispectrum in N1, while it is  $\sim 5\%$  or less in the other cases. Despite the lower accuracy in reproducing the full bispectrum in GR, COLA's boost factors are 5% accurate for most of the configurations. Figure 19 shows the configuration dependence of the bispectra boost factors of galaxies in real space. Again, the bias model dominates the MG signal hiding the clear MG boost factors of DM (figure 5). The strongest deviations from GR consistently appear towards the



**Figure 16.** Comparison of bispectrum (left column) and reduced bispectrum (right column) of halos in COLA (orange crosses) and  $N$ -body (blue dots) simulations. The configurations with  $k_1 \geq k_2 \geq k_3$  are displayed in ascending order of the values of the wavenumbers  $k_1$ ,  $k_2$ , and  $k_3$  respectively. The vertical lines denote the value of  $k_1$  for the configurations immediately to the right of each line. *Top*: full signal in GR. *Middle and Bottom*: boost factors in F5 and N1 respectively. In each panel, the bottom sub-panel shows the ratio between the COLA and  $N$ -body signals displayed in top sub-panel.

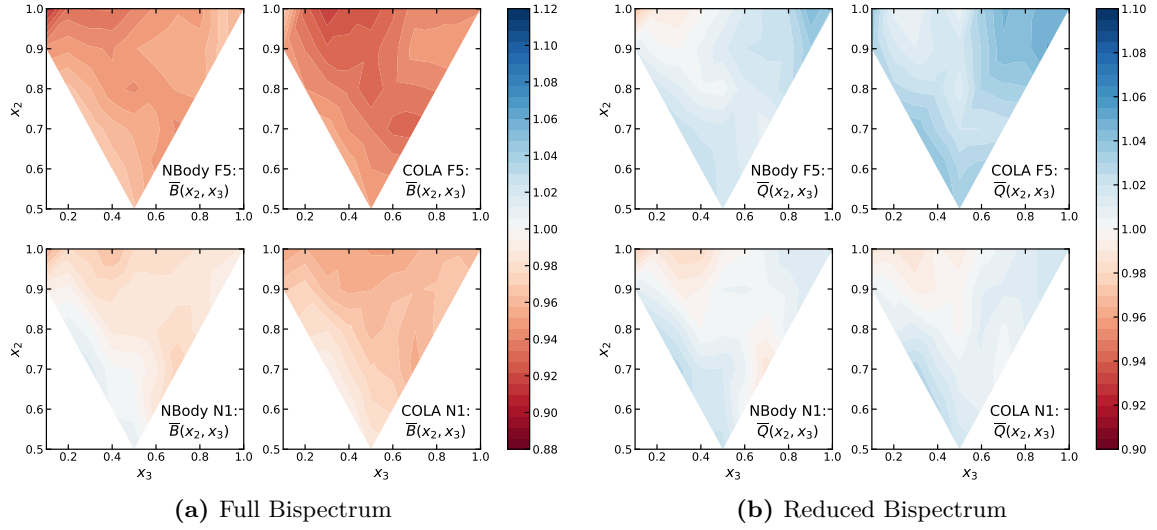
equilateral configuration (top right corner), in both full and reduced bispectrum independently of the gravity model. COLA tends to over-predict the MG signal of the bispectra in N1. Nonetheless, the qualitative description of the configuration dependence is consistent between COLA and  $N$ -body.



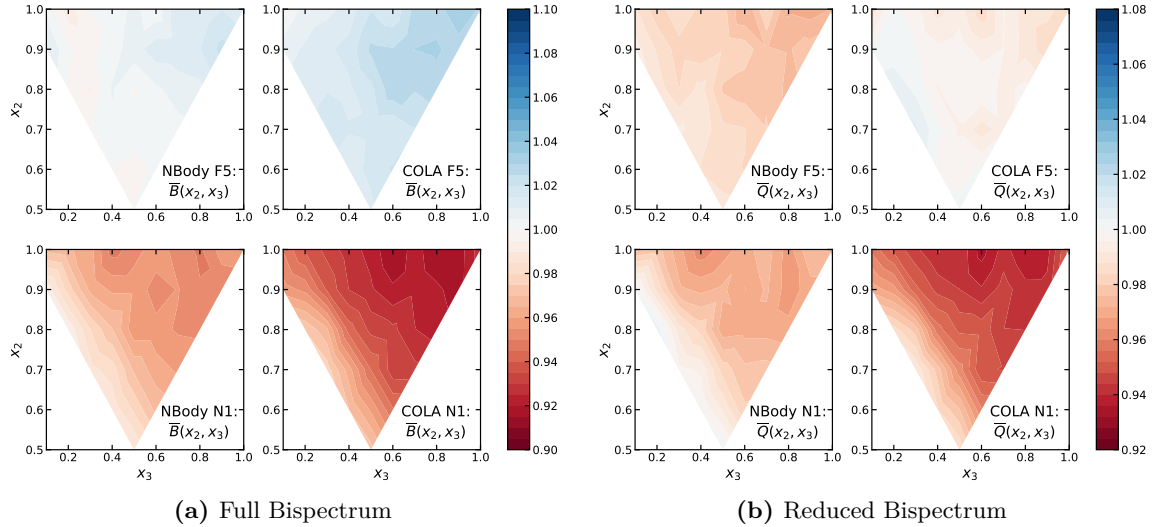


**Figure 17.** Same as figure 16 but for galaxies.





**Figure 18.** Comparison of the configuration dependence of the F5 (top row) and N1 (bottom row) boost factors of bispectrum (on the left) and reduced-bispectrum (on the right) of halos in  $N$ -body (first and third columns) and COLA simulations (second and fourth columns). The colour-bars show the amplitude of the boost factors with blue (red) denoting stronger (weaker) signal in MG with respect to GR. *In each panel:* the top right, top left and bottom corners of the triangle correspond to the equilateral, squeezed and folded configurations respectively. The squeezed configuration is missing from the figure due to the bin's width.



**Figure 19.** Same as figure 18 but for galaxies.

## References

- [1] EUCLID collaboration, *Euclid Definition Study Report*, [arXiv:1110.3193](#) [INSPIRE].
- [2] LSST SCIENCE and LSST PROJECT collaborations, *LSST Science Book, Version 2.0*, [arXiv:0912.0201](#) [INSPIRE].
- [3] DESI collaboration, *Overview of the DESI Legacy Imaging Surveys*, *Astron. J.* **157** (2019) 168 [[arXiv:1804.08657](#)] [INSPIRE].

- [4] C.M. Will, *The Confrontation between General Relativity and Experiment*, *Living Rev. Rel.* **17** (2014) 4 [[arXiv:1403.7377](#)] [[INSPIRE](#)].
- [5] K. Koyama, *Cosmological Tests of Modified Gravity*, *Rept. Prog. Phys.* **79** (2016) 046902 [[arXiv:1504.04623](#)] [[INSPIRE](#)].
- [6] A. Joyce, B. Jain, J. Khoury and M. Trodden, *Beyond the Cosmological Standard Model*, *Phys. Rept.* **568** (2015) 1 [[arXiv:1407.0059](#)] [[INSPIRE](#)].
- [7] H.A. Winther et al., *Modified Gravity N-body Code Comparison Project*, *Mon. Not. Roy. Astron. Soc.* **454** (2015) 4208 [[arXiv:1506.06384](#)] [[INSPIRE](#)].
- [8] S. Tassev, M. Zaldarriaga and D. Eisenstein, *Solving Large Scale Structure in Ten Easy Steps with COLA*, *JCAP* **06** (2013) 036 [[arXiv:1301.0322](#)] [[INSPIRE](#)].
- [9] H.A. Winther and P.G. Ferreira, *Fast route to nonlinear clustering statistics in modified gravity theories*, *Phys. Rev. D* **91** (2015) 123507 [[arXiv:1403.6492](#)] [[INSPIRE](#)].
- [10] G. Valogiannis and R. Bean, *Efficient simulations of large scale structure in modified gravity cosmologies with comoving Lagrangian acceleration*, *Phys. Rev. D* **95** (2017) 103515 [[arXiv:1612.06469](#)] [[INSPIRE](#)].
- [11] 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*, *JCAP* **08** (2017) 006 [[arXiv:1703.00879](#)] [[INSPIRE](#)].
- [12] G. Brando, B. Fiorini, K. Koyama and H.A. Winther, *Enabling matter power spectrum emulation in beyond- $\Lambda$ CDM cosmologies with COLA*, *JCAP* **09** (2022) 051 [[arXiv:2203.11120](#)] [[INSPIRE](#)].
- [13] A.A. Starobinsky, *A New Type of Isotropic Cosmological Models Without Singularity*, *Phys. Lett. B* **91** (1980) 99 [[INSPIRE](#)].
- [14] G.R. Dvali, G. Gabadadze and M. Porrati, *4 – D gravity on a brane in 5 – D Minkowski space*, *Phys. Lett. B* **485** (2000) 208 [[hep-th/0005016](#)] [[INSPIRE](#)].
- [15] B. Fiorini, K. Koyama, A. Izard, H.A. Winther, B.S. Wright and B. Li, *Fast generation of mock galaxy catalogues in modified gravity models with COLA*, *JCAP* **09** (2021) 021 [[arXiv:2106.05197](#)] [[INSPIRE](#)].
- [16] C. Hahn, F. Villaescusa-Navarro, E. Castorina and R. Scoccimarro, *Constraining  $M_\nu$  with the bispectrum. Part I. Breaking parameter degeneracies*, *JCAP* **03** (2020) 040 [[arXiv:1909.11107](#)] [[INSPIRE](#)].
- [17] C. Hahn and F. Villaescusa-Navarro, *Constraining  $M_\nu$  with the bispectrum. Part II. The information content of the galaxy bispectrum monopole*, *JCAP* **04** (2021) 029 [[arXiv:2012.02200](#)] [[INSPIRE](#)].
- [18] B.S. Wright, H.A. Winther and K. Koyama, *COLA with massive neutrinos*, *JCAP* **10** (2017) 054 [[arXiv:1705.08165](#)] [[INSPIRE](#)].
- [19] W. Hu and I. Sawicki, *Models of  $f(R)$  Cosmic Acceleration that Evade Solar-System Tests*, *Phys. Rev. D* **76** (2007) 064004 [[arXiv:0705.1158](#)] [[INSPIRE](#)].
- [20] J. Khoury and A. Weltman, *Chameleon fields: Awaiting surprises for tests of gravity in space*, *Phys. Rev. Lett.* **93** (2004) 171104 [[astro-ph/0309300](#)] [[INSPIRE](#)].
- [21] J. Khoury and A. Weltman, *Chameleon cosmology*, *Phys. Rev. D* **69** (2004) 044026 [[astro-ph/0309411](#)] [[INSPIRE](#)].
- [22] A.I. Vainshtein, *To the problem of nonvanishing gravitation mass*, *Phys. Lett. B* **39** (1972) 393 [[INSPIRE](#)].
- [23] M. Cautun et al., *The Santiago-Harvard-Edinburgh-Durham void comparison — I. SHEDding light on chameleon gravity tests*, *Mon. Not. Roy. Astron. Soc.* **476** (2018) 3195 [[arXiv:1710.01730](#)] [[INSPIRE](#)].

- [24] B. Li, G.-B. Zhao, R. Teyssier and K. Koyama, *ECOSMOG: An Efficient Code for Simulating Modified Gravity*, *JCAP* **01** (2012) 051 [[arXiv:1110.1379](#)] [[INSPIRE](#)].
- [25] Z. Zheng, A.L. Coil and I. Zehavi, *Galaxy Evolution from Halo Occupation Distribution Modeling of DEEP2 and SDSS Galaxy Clustering*, *Astrophys. J.* **667** (2007) 760 [[astro-ph/0703457](#)] [[INSPIRE](#)].
- [26] J.F. Navarro, C.S. Frenk and S.D.M. White, *The Structure of cold dark matter halos*, *Astrophys. J.* **462** (1996) 563 [[astro-ph/9508025](#)] [[INSPIRE](#)].
- [27] 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** (2013) 109 [[arXiv:1110.4372](#)] [[INSPIRE](#)].
- [28] M. Manera et al., *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: a large sample of mock galaxy catalogues*, *Mon. Not. Roy. Astron. Soc.* **428** (2012) 1036 [[arXiv:1203.6609](#)] [[INSPIRE](#)].
- [29] M.M. Ivanov, O.H.E. Philcox, M. Simonović, M. Zaldarriaga, T. Nischimichi and M. Takada, *Cosmological constraints without nonlinear redshift-space distortions*, *Phys. Rev. D* **105** (2022) 043531 [[arXiv:2110.00006](#)] [[INSPIRE](#)].
- [30] N. Kaiser, *Clustering in real space and in redshift space*, *Mon. Not. Roy. Astron. Soc.* **227** (1987) 1 [[INSPIRE](#)].
- [31] J.C. Jackson, *Fingers of God: A critique of Rees' theory of primordial gravitational radiation*, *Mon. Not. Roy. Astron. Soc.* **156** (1972) 1P [[arXiv:0810.3908](#)] [[INSPIRE](#)].
- [32] R. Scoccimarro, *Fast Estimators for Redshift-Space Clustering*, *Phys. Rev. D* **92** (2015) 083532 [[arXiv:1506.02729](#)] [[INSPIRE](#)].
- [33] Y.P. Jing, *Correcting for the alias effect when measuring the power spectrum using FFT*, *Astrophys. J.* **620** (2005) 559 [[astro-ph/0409240](#)] [[INSPIRE](#)].
- [34] E. Sefusatti, M. Crocce, R. Scoccimarro and H. Couchman, *Accurate Estimators of Correlation Functions in Fourier Space*, *Mon. Not. Roy. Astron. Soc.* **460** (2016) 3624 [[arXiv:1512.07295](#)] [[INSPIRE](#)].
- [35] R. Scoccimarro, *The bispectrum: from theory to observations*, *Astrophys. J.* **544** (2000) 597 [[astro-ph/0004086](#)] [[INSPIRE](#)].
- [36] J.N. Fry, *The Galaxy correlation hierarchy in perturbation theory*, *Astrophys. J.* **279** (1984) 499 [[INSPIRE](#)].
- [37] S. Matarrese, L. Verde and A.F. Heavens, *Large scale bias in the universe: Bispectrum method*, *Mon. Not. Roy. Astron. Soc.* **290** (1997) 651 [[astro-ph/9706059](#)] [[INSPIRE](#)].
- [38] M. Colavincenzo et al., *Comparing approximate methods for mock catalogues and covariance matrices — III: bispectrum*, *Mon. Not. Roy. Astron. Soc.* **482** (2018) 4883.
- [39] H. Gil-Marín, F. Schmidt, W. Hu, R. Jimenez and L. Verde, *The Bispectrum of  $f(R)$  Cosmologies*, *JCAP* **11** (2011) 019 [[arXiv:1109.2115](#)] [[INSPIRE](#)].
- [40] Y.-C. Cai, A. Taylor, J.A. Peacock and N. Padilla, *Redshift-space distortions around voids*, *Mon. Not. Roy. Astron. Soc.* **462** (2016) 2465 [[arXiv:1603.05184](#)] [[INSPIRE](#)].
- [41] S. Nadathur and W.J. Percival, *An accurate linear model for redshift space distortions in the void-galaxy correlation function*, *Mon. Not. Roy. Astron. Soc.* **483** (2019) 3472 [[arXiv:1712.07575](#)] [[INSPIRE](#)].
- [42] 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. Roy. Astron. Soc.* **451** (2015) 4215 [[arXiv:1411.5694](#)] [[INSPIRE](#)].
- [43] Y.-C. Cai, N. Padilla and B. Li, *Testing Gravity using Cosmic Voids*, *Mon. Not. Roy. Astron. Soc.* **451** (2015) 1036 [[arXiv:1410.1510](#)] [[INSPIRE](#)].

- [44] B. Falck, K. Koyama, G.-B. Zhao and M. Cautun, *Using Voids to Unscreen Modified Gravity*, *Mon. Not. Roy. Astron. Soc.* **475** (2018) 3262 [[arXiv:1704.08942](#)] [[INSPIRE](#)].
- [45] P. Cataldi, S. Pedrosa, N. Padilla, S. Landau, C. Arnold and B. Li, *Fingerprints of modified gravity on galaxies in voids*, *Mon. Not. Roy. Astron. Soc.* **515** (2022) 5358 [[arXiv:2207.12917](#)] [[INSPIRE](#)].
- [46] E. Paillas et al., *The Santiago-Harvard-Edinburgh-Durham void comparison II: unveiling the Vainshtein screening using weak lensing*, *Mon. Not. Roy. Astron. Soc.* **484** (2019) 1149 [[arXiv:1810.02864](#)] [[INSPIRE](#)].
- [47] E.L.D. Perico, R. Voivodic, M. Lima and D.F. Mota, *Cosmic voids in modified gravity scenarios*, *Astron. Astrophys.* **632** (2019) A52 [[arXiv:1905.12450](#)] [[INSPIRE](#)].
- [48] S. Contarini, F. Marulli, L. Moscardini, A. Veropalumbo, C. Giocoli and M. Baldi, *Cosmic voids in modified gravity models with massive neutrinos*, *Mon. Not. Roy. Astron. Soc.* **504** (2021) 5021 [[arXiv:2009.03309](#)] [[INSPIRE](#)].
- [49] J.M. Colberg et al., *The Aspen-Amsterdam Void Finder Comparison Project*, *Mon. Not. Roy. Astron. Soc.* **387** (2008) 933 [[arXiv:0803.0918](#)] [[INSPIRE](#)].
- [50] E. Massara et al., *Velocity profiles of matter and biased tracers around voids*, [arXiv:2206.14120](#) [[INSPIRE](#)].
- [51] M.C. Neyrinck, *ZOBOV: a parameter-free void-finding algorithm*, *Mon. Not. Roy. Astron. Soc.* **386** (2008) 2101 [[arXiv:0712.3049](#)] [[INSPIRE](#)].
- [52] P.M. Sutter et al., *VIDE: The Void IDentification and Examination toolkit*, *Astron. Comput.* **9** (2015) 1 [[arXiv:1406.1191](#)] [[INSPIRE](#)].
- [53] A. Woodfinden, S. Nadathur, W.J. Percival, S. Radinović, E. Massara and H.A. Winther, *Measurements of cosmic expansion and growth rate of structure from voids in the Sloan Digital Sky Survey between redshift 0.07 and 1.0*, *Mon. Not. Roy. Astron. Soc.* **516** (2022) 4307 [[arXiv:2205.06258](#)] [[INSPIRE](#)].
- [54] P.J.E. Peebles, *Principles of physical cosmology*, Princeton University Press (1994).
- [55] S. Nadathur and S. Hotchkiss, *The nature of voids: I. Watershed void finders and their connection with theoretical models*, *Mon. Not. Roy. Astron. Soc.* **454** (2015) 2228 [[arXiv:1504.06510](#)] [[INSPIRE](#)].
- [56] D. Alonso, *CUTE solutions for two-point correlation functions from large cosmological datasets*, [arXiv:1210.1833](#) [[INSPIRE](#)].
- [57] C.F.J. Wu, *Jackknife, bootstrap and other resampling methods in regression analysis*, *Ann. Stat.* **14** (1986) 1261.
- [58] P. Norberg, C.M. Baugh, E. Gaztanaga and D.J. Croton, *Statistical Analysis of Galaxy Surveys — I. Robust error estimation for 2-point clustering statistics*, *Mon. Not. Roy. Astron. Soc.* **396** (2009) 19 [[arXiv:0810.1885](#)] [[INSPIRE](#)].
- [59] P. Norberg, E. Gaztanaga, C.M. Baugh and D.J. Croton, *Statistical Analysis of Galaxy Surveys-IV: An objective way to quantify the impact of superstructures on galaxy clustering statistics*, *Mon. Not. Roy. Astron. Soc.* **418** (2011) 2435 [[arXiv:1106.5701](#)] [[INSPIRE](#)].
- [60] F.G. Mohammad and W.J. Percival, *Creating jackknife and bootstrap estimates of the covariance matrix for the two-point correlation function*, *Mon. Not. Roy. Astron. Soc.* **514** (2022) 1289 [[arXiv:2109.07071](#)] [[INSPIRE](#)].
- [61] S.D. Landy and A.S. Szalay, *Bias and variance of angular correlation functions*, *Astrophys. J.* **412** (1993) 64 [[INSPIRE](#)].
- [62] S. Nadathur, P. Carter and W. Percival, *A Zeldovich reconstruction method for measuring redshift space distortions using cosmic voids*, *Mon. Not. Roy. Astron. Soc.* **482** (2019) 2459 [[arXiv:1805.09349](#)] [[INSPIRE](#)].