

# Cosmological constraints from Fourier phase statistics

Kamran Ali,<sup>1,2★</sup> Danail Obreschkow,<sup>1,2</sup> Cullan Howlett,<sup>1,2</sup> Camille Bonvin,<sup>3</sup>  
 Claudio Llinares,<sup>4</sup> Felipe Oliveira Franco<sup>3</sup> and Chris Power<sup>1,2</sup>

<sup>1</sup>International Centre for Radio Astronomy Research (ICRAR), University of Western Australia, 35 Stirling Highway, Crawley WA 6009, Australia

<sup>2</sup>ARC Centre of Excellence for All-sky Astrophysics (CAASTRO)

<sup>3</sup>Département de Physique Théorique and Center for Astroparticle Physics (CAP), University of Geneva, 24 quai Ernest Ansermet, CH-1211 Geneva, Switzerland

<sup>4</sup>Department of Physics, Institute for Computational Cosmology, Durham University, Durham DH1 3LE, UK

Accepted 2018 June 25. Received 2018 May 28; in original form 2017 December 26

## ABSTRACT

Most statistical inference from cosmic large-scale structure relies on two-point statistics, i.e. on the galaxy–galaxy correlation function (2PCF) or the power spectrum. These statistics capture the full information encoded in the Fourier amplitudes of the galaxy density field but do not describe the Fourier phases of the field. Here, we quantify the information contained in the line correlation function (LCF), a three-point Fourier phase correlation function. Using cosmological simulations, we estimate the Fisher information (at redshift  $z = 0$ ) of the 2PCF, LCF, and their combination, regarding the cosmological parameters of the standard  $\Lambda$ CDM model, as well as a warm dark matter model and the  $f(R)$  and Symmetron-modified gravity models. The galaxy bias is accounted for at the level of a linear bias. The relative information of the 2PCF and the LCF depends on the survey volume, sampling density (shot noise), and the bias uncertainty. For a volume of  $1 h^{-3} \text{Gpc}^3$ , sampled with points of mean density  $\bar{n} = 2 \times 10^{-3} h^3 \text{Mpc}^{-3}$ , and a bias uncertainty of 13 per cent, the LCF improves the parameter constraints by about 20 per cent in the  $\Lambda$ CDM cosmology and potentially even more in alternative models. Finally, since a linear bias only affects the Fourier amplitudes (2PCF), but not the phases (LCF), the combination of the 2PCF and the LCF can be used to break the degeneracy between the linear bias and  $\sigma_8$ , present in two-point statistics.

**Key words:** methods: numerical – large-scale structure of Universe.

## 1 INTRODUCTION

Cosmic large-scale structure (LSS) grows from primordial density fluctuations under the effect of gravity and dark energy. This structure hence contains useful information on the cosmological model, whether this is a specific form of the standard  $\Lambda$ CDM model or an alternative proposal. Modern galaxy redshift surveys (e.g. Alam et al. 2017; DES Collaboration 2017) decode the information contained in the LSS using spatial statistics, most commonly the two-point statistics, that is the isotropic two-point correlation function (2PCF) or its Fourier counterpart, the power spectrum.

While the structural information of the early Universe, seen in the cosmic microwave background (CMB), seems to be fully described by the two-point statistics, non-linear gravity-driven evolution causes a flow of information into higher order statistics as the Universe evolves (Scoccimarro 1997). This moti-

vates the search for efficient statistical estimators to probe the excess information in the LSS, which lies beyond the two-point statistics. This quest has gained much momentum with the prospect of further surveys such as DES (The Dark Energy Survey Collaboration 2005), TAIPAN (da Cunha et al. 2017), EUCLID (Laureijs et al. 2011), and SKA HI surveys (Dewdney et al. 2009).

Obvious candidates for the statistical estimators beyond the 2PCF are the three-point and higher order isotropic correlation functions. In fact the family of all  $N$ -point correlation functions constitutes a full (albeit highly redundant) description of the statistical information in the LSS. For instance the bispectrum (Fourier transform of the three-point correlation function) of the CMB sets an upper limit on the non-Gaussianity (Planck Collaboration 2016) of the primordial density field, thereby restricting the range of allowed cosmological models. At late times, the bispectrum provides additional constraints on the standard cosmological model (Gil-Marín et al. 2015). The same estimator, however, is unable to put stringent constraints on alternate cosmological models (Gil-Marín et al. 2011), until the effects of redshift

\*E-mail: kamranali\_9@hotmail.com

space distortions are incorporated into the analysis. The information pertaining to the redshifted structure allows differentiation between modified gravity and standard models (Sabui et al. 2016) at small scales. Including higher order (>3 point) correlation functions makes it possible to increase the minimum length scales that need to be considered to differentiate between cosmological models (Hellwing et al. 2013).

Alternative estimators to the standard  $N$ -point correlation functions exist and have previously been used to distinguish between standard and alternate cosmological models. For instance, halo shape statistics (Llinares, Mota & Winther 2014), the two-point function of different types of non-linearly rescaled density fields (Lombriser, Simpson & Mead 2015; White 2016; Llinares & McCullagh 2017) and void count and shape statistics (Cai, Padilla & Li 2015; Falck et al. 2017; Voivodic et al. 2017) also probe this non-linear regime and allows us to differentiate between modified gravity models. However, these alternative estimators, as well as the standard  $N$ -point correlation functions, are strongly correlated to the 2PCF. This is because these alternative estimators depend on the amplitudes of the Fourier field, already fully measured by the 2PCF. Therefore, total constraints from the 2PCF and such alternative estimators are often smaller than their independent addition would suggest.

Given the dependence of the 2PCF on the Fourier space amplitudes, it seems promising to introduce a second statistical measure that only depends on Fourier phases. One such measure is the so-called line correlation function (LCF, defined in Section 2.1) introduced by Obreschkow et al. (2013). Using a simplistic information analysis Obreschkow et al. speculated that the LCF is a promising estimator, especially when probing alternative cosmological models. Moreover, because the LCF measures the three-point statistics of the Fourier phases, irrespective of amplitudes, it is independent of linear bias (Wolstenhulme, Bonvin & Obreschkow 2015), the uncertainty of which plagues all LSS surveys. To study the effectiveness of the LCF in galaxy surveys, Eggemeier & Smith (2017) studied its correlation with the two-point estimator on different scales. This was further expanded upon by Byun et al. (2017), who compared the effectiveness of the LCF with other three-point estimators, including the bispectrum and found it to be a promising candidate for future surveys. Given that all the odd isotropic  $N$ -point functions contribute to the LCF (Wolstenhulme et al. 2015), we expect this estimator to play a more significant role in cosmological models which affect the non-linear regime of structure evolution due to gravity. Hence, it is an interesting avenue to apply the LCF to alternative gravity models.

In this paper, we investigate the additional information contained in the LCF, relative to the information in the 2PCF, based on cosmological  $N$ -body simulations. We account for the difference between observable galaxies and the underlying total density field via a linear bias. We start by defining the statistical estimators and the Fisher information (FI) matrix (FIM). We then set out the algorithm to compute the derivatives and covariances required for the FIM and investigate the effect of the uncertainty of the linear bias on the overall covariance matrix. In Section 3, we use  $N$ -body simulations to measure the information in the 2PCF, the LCF and their combined information in standard ( $\Lambda$ CDM), Warm Dark Matter (WDM),  $f(R)$ , and Symmetron cosmologies and end by exploring the effect of linear bias uncertainties on parameter estimation. Throughout this work, we only take into account the real-space 2PCF and LCF while similar analyses in redshift-space are left for future work. Section 4 concludes the paper with a brief discussion.

## 2 METHODS

This section introduces the tools used for analysing the cosmological information in  $N$ -body simulations. We start by defining the statistical estimators, i.e. the 2PCF and LCF, and outlining the method of measuring their relative information. We then describe the algorithm used to measure the derivatives and covariances of the 2PCF and LCF required for FIM. Finally, we quantify the effect of linear bias uncertainties on the covariance matrix of the statistical estimators.

### 2.1 Estimators

Cosmological  $N$ -body simulation boxes with periodic boundary conditions (see Section 3.1) are used to compute the matter density field  $\rho(\mathbf{r})$  at redshift  $z = 0$ . It is convenient to define the overdensity field,  $\delta(\mathbf{r})$ , as

$$\delta(\mathbf{r}) = \frac{\rho(\mathbf{r}) - \bar{\rho}}{\bar{\rho}}, \quad (1)$$

where  $\bar{\rho}$  is the average density within the simulation volume. The next step is to take the Fourier transform of this quantity, using the formalism of Obreschkow et al. (2013), and compute the power spectrum,  $P(\mathbf{k})$ , defined as

$$\langle \delta_{\mathbf{k}} \delta_{\mathbf{k}'} \rangle = (2\pi)^3 \delta_{\mathbf{D}}(\mathbf{k} + \mathbf{k}') P(\mathbf{k}); \quad (2)$$

where  $\langle \rangle$  and  $\delta_{\mathbf{D}}$  are the ensemble averages and Dirac Delta function, respectively. The 2PCF is then obtained using

$$\xi(r) = \frac{V}{(2\pi)^3} \int d^3k \frac{\sin(|\mathbf{k}|r)}{|\mathbf{k}|r} P(|\mathbf{k}|). \quad (3)$$

Since the 2PCF and the power spectrum are Fourier counterparts of one another they contain the same FI. Note that in real surveys this might not be the case given that the two statistics are computed using their own approximate estimators.

The LCF measures correlation between the Fourier phases,  $\epsilon(\mathbf{k}) = \delta(\mathbf{k})/|\delta(\mathbf{k})|$ . We here use the LCF definition of Wolstenhulme et al. (2015)

$$\ell(r) = \frac{V^3}{(2\pi)^9} \left(\frac{r^3}{V}\right)^{3/2} \iint_{|\mathbf{k}_1|, |\mathbf{k}_2|, |\mathbf{k}_3| \leq 2\pi/r} d^3\mathbf{k}_1 d^3\mathbf{k}_2 d^3\mathbf{k}_3 e^{i[|\mathbf{k}_1 \cdot \mathbf{s} + \mathbf{k}_2 \cdot (\mathbf{s} + \mathbf{r}) + \mathbf{k}_3 \cdot (\mathbf{s} - \mathbf{r})]} \langle \epsilon(\mathbf{k}_1) \epsilon(\mathbf{k}_2) \epsilon(\mathbf{k}_3) \rangle. \quad (4)$$

This LCF is computed using the `PROCORR` package, which follows the formalism of Obreschkow et al. (2013).

### 2.2 Fisher information matrix

The main aim of this paper is to measure the available information in the estimators and thereby constrain a set of parameters. To do so we use the FIM, which for a given Log-Likelihood function,  $\mathcal{L}$ , is defined by

$$\mathbf{F}_{ij} = \left\langle \frac{\partial \mathcal{L}(\mathbf{P}, \boldsymbol{\theta})}{\partial \theta_i} \frac{\partial \mathcal{L}(\mathbf{P}, \boldsymbol{\theta})}{\partial \theta_j} \right\rangle, \quad (5)$$

where  $\boldsymbol{\theta}$  consists of  $N$  model parameters,  $\theta_i \forall i \in 1, \dots, N$ , with  $\mathbf{P}$  being the estimator we measure. A useful feature of this matrix is that its inverse is an estimator of the covariance matrix of the model parameters (see Section 3.1). In the Laplace approximation (Gaussianity of likelihoods), the FIM simplifies to (Tegmark, Taylor

& Heavens 1997)

$$F_{ij} = \frac{\partial \mathbf{P}(\boldsymbol{\theta})}{\partial \theta_i} \mathbf{C}(\mathbf{P}, \boldsymbol{\theta})^{-1} \frac{\partial \mathbf{P}(\boldsymbol{\theta})}{\partial \theta_j} + \frac{1}{2} \text{Tr} \left[ \mathbf{C}(\mathbf{P}, \boldsymbol{\theta})^{-1} \frac{\partial \mathbf{C}(\mathbf{P}, \boldsymbol{\theta})}{\partial \theta_i} \mathbf{C}(\mathbf{P}, \boldsymbol{\theta})^{-1} \frac{\partial \mathbf{C}(\mathbf{P}, \boldsymbol{\theta})}{\partial \theta_j} \right], \quad (6)$$

where  $\mathbf{C}$  is the covariance matrix of the likelihood. The variation of the covariance matrix (trace term in equation 6) turns out to be subdominant relative to the first term. In the case of the 2PCF, this was shown by explicit numerical calculations. For instance, in the particular case of the Symmetron-modified gravity model, Llinares & McCullagh (2017) found that the covariance matrix of the power spectrum (hence the 2PCF) does not vary significantly relative to that of  $\Lambda$ CDM. For the LCF, Eggemeier & Smith (2017) show that the trace term vanishes identically at lowest order, since the Gaussian part of the covariance matrix is independent of the cosmology. This lowest order solution agrees with full numerical computations over a wide range of length scales  $r \geq 40 h^{-1}$  Mpc. Hence, in subsequent sections, we will ignore the trace term and evaluate the derivatives of the 2PCF and the LCF relative to the model parameters at a fiducial cosmology.

An alternative information measure, commonly, used in cosmology literature is the cumulative signal to noise of the relevant statistic (e.g. Sefusatti & Vernizzi 2011; Bonvin, Hui & Gaztanaga 2016). This metric, defined in terms of the raw estimator instead of its derivative, i.e.  $\text{SNR} = \mathbf{P}\mathbf{C}^{-1}\mathbf{P}$ , quantifies the possibility of measuring the estimator in a given cosmological volume. However, in general the Fisher methodology provides a more robust way of estimating the parameters.

### 2.3 Derivatives

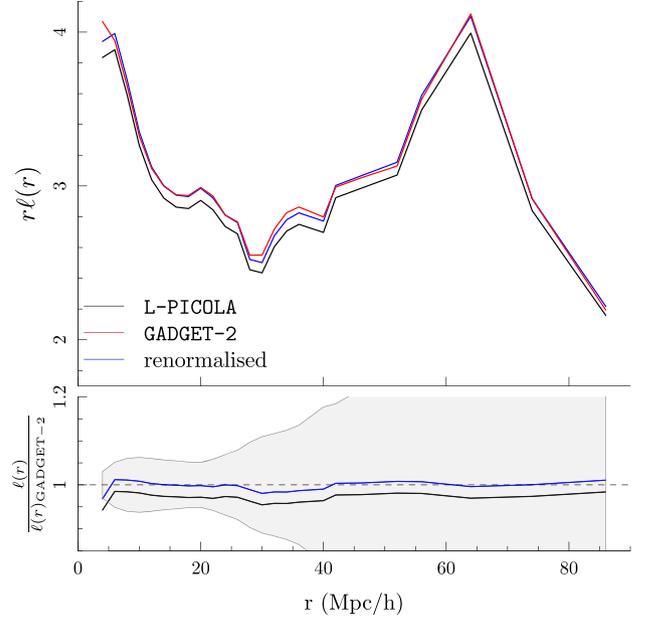
To evaluate the FI of the LCF and the 2PCF, we require the expectation values of their derivatives with respect to the cosmological parameters  $\theta_i$  under investigation. In practice these expectations of the derivatives can be computed with a finite difference method using two simulation boxes per parameter, one centred at a fiducial cosmology,  $\theta_0$ , and one with the parameter in question varied by a small difference  $\Delta\theta$ ,

$$\frac{\partial \mathbf{P}(\theta_0)}{\partial \theta} = \frac{\mathbf{P}(\theta_0) - \mathbf{P}(\theta_0 - \Delta\theta)}{\Delta\theta}. \quad (7)$$

Here, the two simulations (with  $\theta_0$  and one for  $\theta_0 - \Delta\theta$ ) rely on the same initial random seed to minimize the randomness of the derivatives. We further reduce this randomness by applying a slight smoothing filter (two-point running average) to the LCF, which suffers from the limited number of modes more than the 2PCF. In Section 3.5, where we fit the bias in  $\Lambda$ CDM only at large scales (with only few modes per box), the derivatives are averaged over 50 random realizations.

### 2.4 Covariance matrix

The second ingredient for the FIM computation is the covariance matrix of the respective statistical estimator. We use 500 simulation boxes that are generated using the COmoving Lagrangian Acceleration (COLA) method (Tassev, Zaldarriaga & Eisenstein 2013) implemented in the L-PICOLA code (Howlett, Manera & Percival 2015). The simulations consist of  $256^3$  particles enclosed in a periodic box of length  $L = 512 h^{-1}$  Mpc with the same background cosmology as in Howlett et al. (2015). The LSS evolved by the COLA technique provides an accurate estimate of the covariances



**Figure 1.** Comparison of the LCF of simulation boxes evolved, by the two  $N$ -body codes L-PICOLA and GADGET-2, from the same initial density field. L-PICOLA tends to slightly but systematically underestimate the LCF by  $\approx 3$  per cent which needs to be accounted for before computing the covariance matrix. The shaded region shows the standard deviation of LCF indicating the significance of the systematic correction at different scales.

of two-point statistics to scales of  $k \leq 0.3 h \text{ Mpc}^{-1}$ . Hence, in our regime of interest,  $r \geq 10 h^{-1}$  Mpc, we can safely use L-PICOLA to estimate the covariance matrix of the 2PCF.

The three-point statistics on the other hand depend strongly on the non-linear growth of gravity and, hence, might not be easily reproducible by the COLA code. To gauge the accuracy of the COLA solver, we compare a set of simulation boxes which are evolved from the same initial density field but by two different  $N$ -body codes, GADGET-2 and L-PICOLA. As shown in Fig. 1, L-PICOLA slightly, but systematically underestimates the LCF by about 3 per cent. We approximately correct this small inaccuracy by rescaling the L-PICOLA LCF by 3 per cent in computing the covariance matrix. This correction is more significant at smaller scales, but the difference persists even at larger scales and might be due to L-PICOLA being unable to accurately reproduce the non-linear regime of structure growth since the LCF depends entirely on the correlation between Fourier phases, the growth of which is a purely non-linear phenomenon. We here assume that the applied correction also holds for the covariance matrix estimated through the L-PICOLA simulations. Since, the LCF depends on three-point statistics, the resulting covariance matrix should depend on four, five, and six-point functions which might differ by a different factor in the COLA method. We expect these corrections to be small relative to the diagonal terms in the covariance matrix; hence they do not significantly alter the FIM.

Given 500 L-PICOLA realizations of the 2PCF and the LCFs, we estimate the full covariance matrix

$$\hat{\mathbf{C}} = \begin{bmatrix} \hat{\mathbf{C}}_{\xi\xi} & \hat{\mathbf{C}}_{\xi\ell} \\ \hat{\mathbf{C}}_{\ell\xi} & \hat{\mathbf{C}}_{\ell\ell} \end{bmatrix}, \quad (8)$$

where the four sub-matrices are defined by

$$\begin{aligned} (\hat{C}_{\xi\xi})_{ij} &= \text{cov}_{N_{\text{sim}}}(\xi(r_i), \xi(r_j)) \\ (\hat{C}_{\ell\ell})_{ij} &= \text{cov}_{N_{\text{sim}}}(\ell(r_i), \ell(r_j)) \\ (\hat{C}_{\xi\ell})_{ij} &= (\hat{C}_{\ell\xi})_{ji} = \text{cov}_{N_{\text{sim}}}(\xi(r_i), \ell(r_j)). \end{aligned} \quad (9)$$

The operator  $\text{cov}_{N_{\text{sim}}}$  is the estimator of the covariance matrix from  $N_{\text{sim}}$  simulation boxes

$$\text{cov}_{N_{\text{sim}}}(x, y) = \frac{1}{N_{\text{sim}} - 1} \sum_{k=1}^{N_{\text{sim}}} (x_k - \langle x \rangle)(y_k - \langle y \rangle). \quad (10)$$

The covariance matrix of cosmological surveys is, in general, made up of three terms (see Bonvin et al. (2016): a Poisson (or shot noise) contribution which depends on the density of the observed galaxies. The second major contribution comes from the fact that we observe a single realization of the observable Universe. The final contribution to the overall covariance matrix is a mixture of the Poisson and cosmic effect and hence depends on both the galaxy density and the survey volume. In the absence of a window function, the covariance matrices of the 2PCF (Meiksin & White 1999; Scoccimarro, Zaldarriaga & Hui 1999; Howlett & Percival 2017) and the LCF (Eggemeier & Smith 2017) are inversely proportional to the survey volume. This allows us to assume that the covariance matrices of the 2PCF and LCF both scale as the inverse of the survey volume, i.e. the covariance matrix  $\tilde{C}$  for a survey of volume  $V_{\text{eff}}$  is

$$\tilde{C} = \frac{V_{\text{sim}}}{V_{\text{eff}}} \hat{C}, \quad (11)$$

where  $V_{\text{sim}}$  is the simulation volume in which the raw covariance matrix  $\hat{C}$  was evaluated.

To gauge the effect of shot noise in our analysis, we randomly draw a subset of  $N_{\text{parts}}$  particles from the simulation box, such that the mean expected particle density  $\bar{n} = N_{\text{parts}}/V_{\text{sim}}$  equals  $2 \times 10^{-3} h^3 \text{ Mpc}^{-3}$ , approximately corresponding to the cumulative space density of haloes more massive than  $10^{12} M_{\odot}$  (Murray, Power & Robotham 2013) – the dynamical mass of typical  $M^*$ -galaxies. Strictly speaking, this subsampling process mimics shot noise in the matter field rather than in the galaxy field – a difference, which we neglect in this work, similarly to others (Eggemeier & Smith 2017). By subsampling the simulation box in this way, we find that the variance of the 2PCF estimator increases while that of the LCF decreases. This is because the expectation of the 2PCF is invariant to random subsampling of the particle field, while that of the LCF decreases significantly, a behaviour explored by Eggemeier & Smith (2017). Note that in computing the FIM in the presence of shot noise, we rescale the derivatives of the full LCF (without shot noise) by a factor  $f(r)$ , defined as the ratio between the expectation of the LCF of the subsampled boxes and the full LCF. This approach neglects the derivative of  $f(r)$  with respect to the cosmological parameters. Numerically, this simplification has a relatively small ( $\lesssim 20$  per cent) effect on the overall FI in the LCF. (An exact computation would nonetheless require more simulation boxes than available for this analysis.)

Modern galaxy redshift surveys probe cosmic volumes of a few  $h^{-3} \text{ Gpc}^3$ . For instance, the CMASS galaxy sample of the BOSS (Eisenstein et al. 2011) survey covers an effective volume of  $3 h^{-3} \text{ Gpc}^3$  (Ntelis et al. 2017) while future surveys might reach up to  $20 h^{-3} \text{ Gpc}^3$  (Duffy 2014). In this work all results are presented for a reference volume of  $V_{\text{eff}} = 1 h^{-3} \text{ Gpc}^3$ .

While the covariance matrix computed via equation (10) is an unbiased estimator, its inverse required for the FIM is generally not. To obtain an unbiased estimator, we apply the correction of

Hartlap, Simon & Schneider (2007)

$$C^{-1} = \frac{N_{\text{sim}} - N_{\text{dim}} - 2}{N_{\text{sim}} - 1} \tilde{C}^{-1}, \quad (12)$$

where  $N_{\text{dim}}$  is the order of  $\tilde{C}$ .

Fig. 2 shows the covariance matrix  $C$  (upper triangle), as well as its correlation matrix  $R$  (lower triangle), defined as

$$R = D^{-1} C D^{-1}, \quad (13)$$

where  $D$  is the square root of the diagonal matrix of  $C$ .

Fig. 2 reveals that the cross-correlation between the 2PCF and the LCF is small, as expected from the fact that the former estimator measures the Fourier amplitudes while the latter measures the Fourier phases. Only at scales  $\lesssim 20 h^{-1} \text{ Mpc}$  does the cross-correlation become significant, indicating a strong coupling between the Fourier phases and amplitudes due to non-linear growth. Note that the rescaling of equations 11 and 12 has no effect on the correlation matrix.

## 2.5 Incorporating bias uncertainty

Most cosmological surveys use galaxies as tracers of the matter density field  $\delta(\mathbf{r})$ . The mapping between the galaxy density  $\delta_g(\mathbf{r})$  and  $\delta(\mathbf{r})$  is complicated in detail. However, at large scales, the asymptotic effect is a uniform rescaling

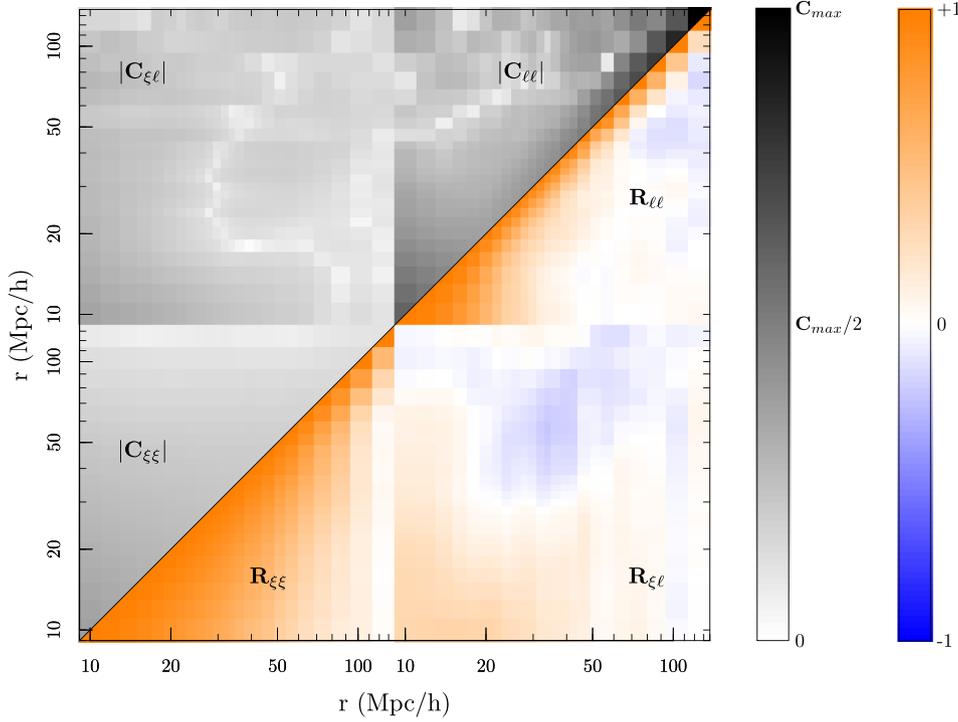
$$\delta_g(\mathbf{r}) = b \delta(\mathbf{r}), \quad (14)$$

where the scaling factor  $b$  is known as linear bias. This bias only affects the amplitudes, not the phases of the Fourier modes, i.e.  $\delta_g(\mathbf{k}) = b\delta(\mathbf{k})$ . Therefore, only the 2PCF is affected by linear bias, not the LCF. Explicitly, the galaxy 2PCF at large scales ( $\gtrsim 50 h^{-1} \text{ Mpc}$ ) is

$$\xi_{g,i} = b^2 \xi_i, \quad (15)$$

where  $\xi_i = \xi(r_i)$  is the estimated dark matter 2PCF at separation scale  $r_i$ . On smaller scales, the bias becomes non-linear. The effects of this non-linearity on the LCF are not yet well understood and will be neglected in this work, although they might be worth considering in the future.

The FI of the 2PCF does not depend on the absolute value of  $b$ , since the covariance matrix  $C$  scales as  $b^4$ , while each of the two derivative terms scales as  $b^2$ , hence cancelling  $b$  in equation (6). Only the *uncertainty* of  $b$  affects the constraints on the cosmological parameters. We can account for this effect of linear bias uncertainty in two ways that correspond to slightly different concepts. The first approach consists of inferring the cosmological parameters from the *galaxy field*, which depends both on these parameters and  $b$ . One then computes the combined FIM of the model parameters and  $b$ , hence constraining them simultaneously. The subtle caveat of this approach is that a linear bias  $b$  only applies to the largest scales and hence should not be fitted to smaller scales. Ignoring this would result in significantly inflated constraints on  $b$ , although this problem can be somewhat alleviated by including a non-linear bias term (Eggemeier & Smith 2017). The alternative approach, used in this paper, is to assume that we infer the cosmological parameters from the *dark matter field*, which is itself inferred by applying an uncertain bias to the galaxy field. This translates into an uncertain dark matter 2PCF. In this case, the FIM is computed only for the cosmological parameters, but the bias uncertainty must be included in the estimator covariance matrix  $C$ . This approach has the advantage that we can impose a realistic uncertainty for the linear bias, accounting for the complexity of real surveys. The



**Figure 2.** The absolute value of covariance matrix (upper triangle) estimated from 500 L-PICOLA simulations and its associated correlation matrix (lower triangle) computed by inverting equation (13) as a function of scale length. As we go to larger scales the variance of the LCF increases while that of 2PCF decreases. The correlation matrix, on the other hand, increases at small scales for all combination of the parameters 2PCF-2PCF (bottom-left), 2PCF-LCF (bottom-right), and LCF-LCF (top-right). The increase in 2PCF-LCF correlation at  $r \lesssim 20 h^{-1}$  Mpc indicates strong coupling between Fourier phases and amplitudes. This property is absent in the early Universe and results from non-linear growth of structure due to gravity.

downside is that this approach doesn't guarantee that the assumed bias uncertainty is consistent with the best estimate of the bias that one might achieve from combining the 2PCF and LCF at the largest scales. This consistency therefore needs to be checked in a separate step, see Section 3.5.

Since the FIM does not depend on the absolute value of the linear bias, we can assume, without loss of generality, that  $\langle b^2 \rangle = 1$  and write the variance of  $b^2$  as  $\sigma_{b^2}^2$ . Using linear error propagation, the variance of the 2PCF with bias uncertainty then depends on  $\sigma_{b^2}$  via

$$\sigma_{\xi_i}^2(\sigma_{b^2}) = \sigma_{\xi_i}^2 + \sigma_{b^2}^2 (\xi_i^2 + \sigma_{\xi_i}^2), \quad (16)$$

where  $\sigma_{\xi_i}^2$  are the diagonal elements of  $C_{\xi\xi}$  without bias uncertainty. Note that this definition of  $\sigma_{\xi_i}^2(\sigma_{b^2})$  satisfies the condition  $\sigma_{\xi_i}^2(0) = \sigma_{\xi_i}^2$ .

Equation (16) allows us to correct the diagonal elements of the covariance matrix  $C$  in order to account for bias uncertainty. The remaining question is how to correct the off-diagonal elements. We found that linear error propagation as in equation (16) is not the right approach. This is because the bias uncertainty introduces cross-correlation between different scales  $r_i$  and  $r_j$ , which is not present in the estimations of the derivatives, when based only on *one* random realization of the universe (which necessarily can only have one bias value). To bypass this issue, we instead assume that the off-diagonal elements of the covariance matrix remain unchanged under the effect of bias uncertainties. Formally, the bias-corrected covariance matrix becomes

$$C(\sigma_{b^2}) = D(\sigma_{b^2})RD(\sigma_{b^2}), \quad (17)$$

where  $R$  is the correlation matrix of  $C$  given in equation (13) and  $D(\sigma_{b^2})$  is a diagonal matrix defined as

$$D(\sigma_{b^2}) = \begin{bmatrix} D_{\xi\xi}(\sigma_{b^2}) & \mathbf{0} \\ \mathbf{0} & D_{\ell} \end{bmatrix} \quad (18)$$

with the diagonal sub-matrices

$$\begin{aligned} (D_{\xi\xi})_{ii}(\sigma_{b^2}) &= \sigma_{\xi_{g,i}}(\sigma_{b^2}) \\ (D_{\ell})_{ii} &= \sigma_{\ell_i}. \end{aligned} \quad (19)$$

Note that this definition of the covariance matrix  $C(\sigma_{b^2})$  satisfies the condition  $C(0) = C$ , i.e. the original covariance matrix is recovered if the bias uncertainty vanishes. A consequence of this definition is the dependence of bias uncertainty on the galaxy survey volume as  $1/\sqrt{V}$ . We will use a fiducial survey volume throughout Section 3 and in Section 3.5 quantify the change in combined information of the estimators due to the bias uncertainty parameter.

### 3 APPLICATION TO SIMULATIONS

We will now evaluate the FIM of the 2PCF and LCF in different cosmological models using  $N$ -body simulations. All FIM computations are carried out at redshift  $z = 0$  and assume a fiducial survey volume of  $V_{\text{eff}} = 1 h^{-3} \text{Gpc}^3$  and a linear bias uncertainty of 13 per cent (standard deviation), that is  $\sigma_{b^2} \approx 0.26$ . The results can be rescaled to other volumes and bias uncertainties using the covariance scaling relations given in Section 2.5. The smallest length scale accounted for in the statistical analyses is set at  $r = 10 h^{-1}$  Mpc. This is a conservative estimate that avoids most effects of (not modelled) non-gravitational baryonic physics, e.g. hydrodynamics and radiation.

### 3.1 Standard $\Lambda$ CDM cosmology

The bulk of the cosmological  $N$ -body simulations used to compute the estimator derivatives use  $256^3$  particles in a periodic simulation box of side length  $L = 512 h^{-1}$  Mpc. The particles are initially placed on a regular cartesian grid, then displaced to obey the power spectrum generated by CAMB (Lewis & Bridle 2002) and evolved using second-order Lagrangian perturbation theory (Crocco & Scoccimarro 2006) to redshift  $z = 49$ . We then employ the GADGET-2 (Springel 2005)  $N$ -body solver to evolve the particles to redshift  $z = 0$ .

To estimate the derivatives of the  $\Lambda$ CDM model, we use equation (7) with a background cosmology given by Planck Collaboration (2016) following which we rescale the LCF derivatives. We take the central parameter values of  $\theta_0 = \{\Omega_b, \Omega_M, \sigma_8, n_s, H_0(\text{km s}^{-1} \text{Mpc}^{-1})\} = \{0.048, 0.31, 0.82, 0.97, 68\}$  with a step spacing of  $\Delta\theta = \{0.10, 0.62, 0.86, 0.40, 46\} \times 10^{-2}$ . We do not include the reionization depth parameter in our analysis since the linear matter power spectrum is independent of this parameter and thus it cannot be constrained by our methodology.

We then evaluate the FIM using the covariance matrix from the 500 L-PICOLA simulations (Sections 2.4 and 2.5). The covariance matrix of the model parameters, inferred from the statistical measurement, is then given by

$$\Sigma = F^{-1}. \quad (20)$$

The order of the matrix  $\Sigma$  is equal to the number of model parameters. The FIM can be computed using the 2PCF, the LCF, or both of them simultaneously. To distinguish between the different parameter covariances resulting from these three cases, we use the symbols  $\Sigma_\xi$ ,  $\Sigma_\ell$ , and  $\Sigma_{\xi\ell}$ .

Fig. 3 shows the parameter uncertainties implied by this covariance matrix in the Gaussian approximation, i.e. for a probability distribution  $p(\theta) \propto \exp[-(\theta - \theta_0)^\dagger \Sigma^{-1}(\theta - \theta_0)/2]$ . The different line colours, respectively, show the constraints from the 2PCF, LCF, and their combination. Interestingly, the uncertainty ellipses of the 2PCF and the LCF often have significantly different orientations. For instance in the case of the  $\sigma_8$ - $\Omega_M$  pair, the LCF helps breaking the classic degeneracy.

Generally, a parameter  $\theta_i$  is better constrained by the 2PCF+LCF than by the 2PCF alone. We quantify this *gain* in constraining power as  $g_i = \sqrt{\Sigma_{\xi,ii}/\Sigma_{\xi\ell,ii}}$ . For  $\Lambda$ CDM, we find that subsampling particles (i.e. mimicking shot noise) within the simulation box increases the contribution of the LCF to the combined constraints,  $\mathbf{g} = \{1.17, 1.19, 1.87, 1.23, 1.26\}$  (for parameters  $\Omega_b, \Omega_M, \sigma_8, n_s, H_0$ ), compared to using the complete set of particles (i.e. no shot noise) in the analysis  $\mathbf{g}_{\text{complete}} = \{1.04, 1.06, 1.52, 1.13, 1.13\}$  for the same bias uncertainty. The same is not case when we neglect bias in the analysis since higher order functions are more susceptible to Poisson noise. This emphasises the importance of bias uncertainty in our analysis, and we verify our choice of method in Section 3.5.

Comparing our results with a previous study by Eggemeier & Smith (2017), we find a good match between the gain values of  $\sigma_8$  and total matter density,  $\Omega_M$  (with shot noise), i.e.  $\mathbf{g} = \{1.25, 1.90\}$  (for parameters  $\Omega_M, \sigma_8$ ). However, their gains on other parameters  $\mathbf{g} = \{1.02, 1.07, 1.07\}$  (for  $\Omega_b, n_s, H_0$ ) are significantly smaller than those found by our full analysis. This disagreement might, primarily, be due to their study using CMB priors which put sharp constraints on  $\Omega_b, n_s$ , and  $H_0$  parameters as compared to  $\Omega_M$  and  $\sigma_8$ . Furthermore, they incorporate the linear and non-linear bias into

the FIM analysis and use  $N$ -body simulations at multiple redshifts to determine the overall parameter constraints.

Fig. 4 (left) depicts the increase in the parameter uncertainty (decrease in information) as we increase the smallest length scale  $r$  used in the FIM computation. One finding from this representation is that the relative constraining power from the LCF slowly increases with decreasing  $r$  and for certain parameters contributes significantly to the combined constraints. For instance, the LCF provides similar constraints to the 2PCF for the  $\sigma_8$  parameter while contributing at most half as much to the combined constraints of the other standard parameters. For all parameters, the increase in constraining power of the LCF at small scales is much steeper than that of the 2PCF due to the onset of non-linear regime.

### 3.2 WDM cosmology

As a first alternative cosmological model, we consider the WDM model, often evoked as a possible solution to potential sub-structure problems in CDM (Klypin et al. 1999).

Our approach to modelling a WDM universe consists in reducing the power of the initial CDM density field on small scales to mimic the free-streaming of the WDM. This is achieved by truncating the input linear matter power at large  $k$  values (small scales), following Bode, Ostriker & Turok (2001),

$$P(k)_{\text{WDM}} = [1 + (\alpha k)^{2\nu}]^{-5\nu} P(k)_{\text{CDM}}, \quad (21)$$

where  $P(k)$  is the linear CDM power spectrum,  $\nu$  is a numerical constant, and  $\alpha$  is a non-linear function of the dark matter particle mass such that  $\lim_{m_{\text{WDM}} \rightarrow \infty} \alpha = 0$  and we recover the standard cosmology. We use the numerically fitted value of  $\nu = 0.5$  by Bode et al. (2001). In the FI analysis, we consider  $m_{\text{WDM}}^{-1}$  as the free additional cosmological parameter to be constrained. To compute the estimator derivatives, we use a spacing of  $\Delta m_{\text{WDM}}^{-1} = (0.2 \text{ keV})^{-1}$  centred at the a cosmology with  $m_{\text{WDM}}^{-1} = (0.2 \text{ keV})^{-1}$  and the same  $\Lambda$ CDM parameters as in the previous section.

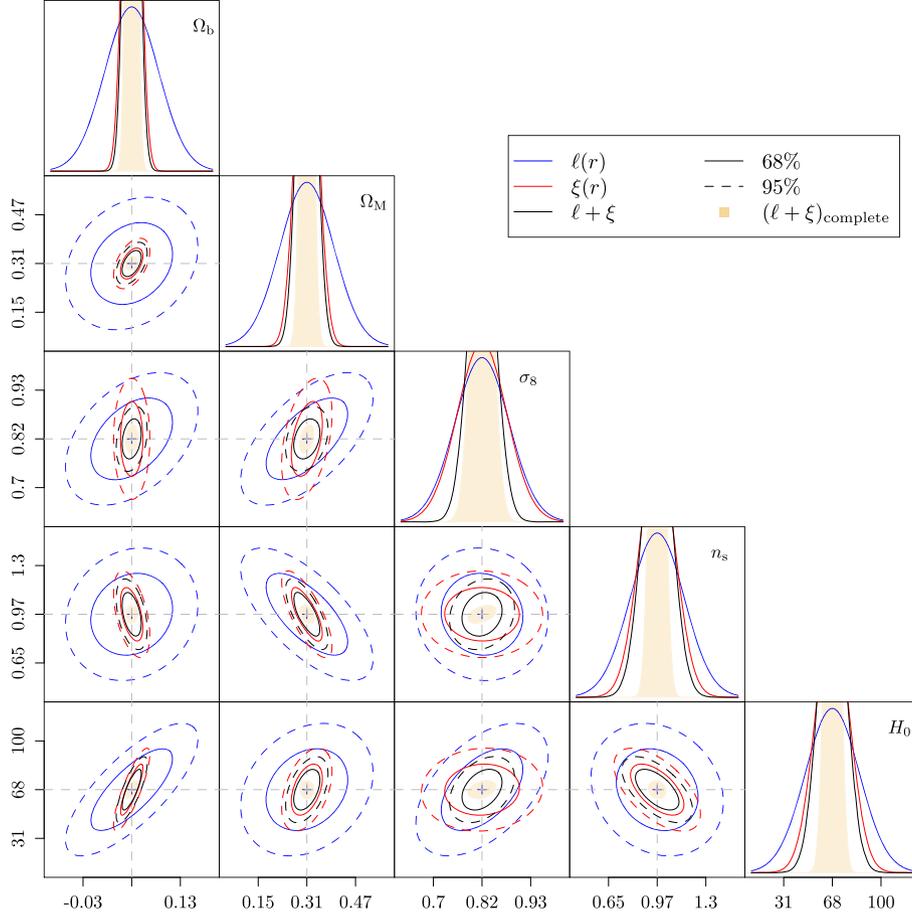
Fig. 5 (top) shows the diagonal and off-diagonal components of  $\Sigma$  determined by inverting the FIM. We find a *gain* of 3.28 on the  $m_{\text{WDM}}$  parameter with a non-noticeable change in constraints on the  $\Lambda$ CDM parameters. This is likely the result of the elliptical orientations of the 2PCF and the LCF being vastly different for  $m_{\text{WDM}}^{-1}$  and  $\Lambda$ CDM parameter pairs. Furthermore, the relative constraining power of the LCF for the  $m_{\text{WDM}}^{-1}$  increases faster when compared to that of standard parameters as shown in Fig. 4 (right). Since the LCF is less sensitive to the linear growth, its variation is mostly dictated by the local gravitational interactions and hence the properties of the underlying dark matter field. This attribute of the LCF allows it to contain significantly more information about the  $m_{\text{WDM}}^{-1}$  for the fiducial survey we assume in our analysis.

### 3.3 $f(R)$ Cosmology

The first modified gravity (MG) analysed in this paper is a specific realization of the Hu-Sawicki  $f(R)$  model (Hu & Sawicki 2007). The original model was defined as a modification to the Ricci scalar,  $R$ , of the Einstein-Hilbert term, i.e.  $R \rightarrow R + f(R)$ , with a free function  $f$  such that the action,  $S$ , for this model is given by

$$S = \int \sqrt{-g} \left[ \frac{R + f(R)}{16\pi G} + \mathcal{L}_m \right] d^4x, \quad (22)$$

where  $\mathcal{L}_m$  and  $g$  are the matter Lagrangian and the Einstein frame metric, respectively. In this prescription, the free function  $f$  is defined



**Figure 3.** The constraints provided by the estimated 2PCF (red), LCF (blue), and their combination (black) for the standard cosmology parameters is shown on the left. The diagonal elements are shown as a 1-D Gaussian while the off-diagonal elements are depicted by their 68 per cent (solid) and 95 per cent (dashed) contours. The shaded yellow regions indicate combined constraints of the two estimators when all particles are used in the analysis (i.e. no shot noise) and as expected provide better constraints than the combined information from smaller particle density. A significant portion of the LCF ellipses have different orientations to those of the 2PCF allowing better constraints on the  $\Lambda$ CDM parameters. The highest gain in constraining power is placed on  $\sigma_8$  with a *gain* of 1.87. For this analysis we assume a fiducial cosmological survey with  $\sigma_b = 0.13$ ,  $\bar{n} = 2 \times 10^{-3} h^3 \text{ Gpc}^{-3}$ , and  $V_{\text{eff}} = 1 h^{-3} \text{ Gpc}^3$ .

as

$$f(R) = -m^2 \frac{c_1(R/m^2)^n}{c_2(R/m^2)^n + 1}, \quad (23)$$

where  $m^2 = H_0^2 \Omega_M$ . The variables  $c_1$ ,  $c_2$ , and  $n$  are the free parameters for this MG model.

The ratio  $c_1/c_2$  describes the expansion of the universe, and enforcing a  $\Lambda$ CDM-like expansion history reduces the number of free parameters to two, i.e.  $n$  and  $c_1$  (or  $c_2$ ). In this case, the derivative  $f_R = df(R)/dR$ , at present day ( $z = 0$ ), becomes

$$f_{R_0} = -n \frac{c_1}{c_2} \left( \frac{\Omega_M}{3(\Omega_M + 4\Omega_\Lambda)} \right)^{n+1}. \quad (24)$$

The geodesic equation of this model takes the form

$$\ddot{\mathbf{x}} + 2H\dot{\mathbf{x}} + \frac{\nabla\Phi}{a^2} - \frac{1}{2} \frac{\nabla f_R}{a^2} = 0, \quad (25)$$

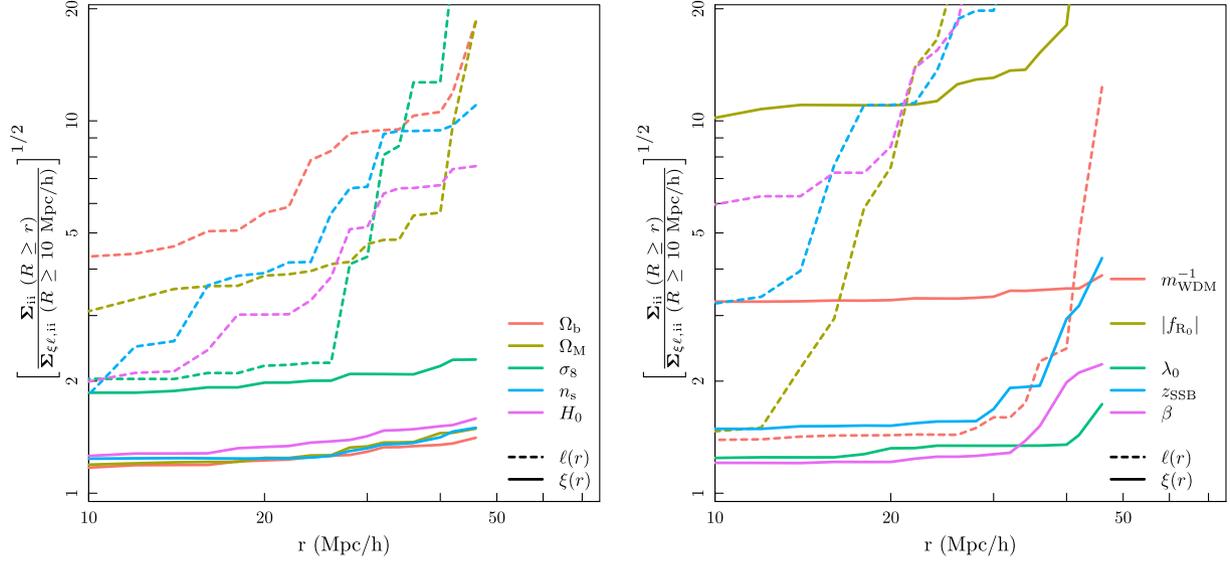
where  $a$  and  $\Phi$  are the scale factor and the scalar perturbation (or the gravitational potential in the classical sense), respectively. The final term on left-hand side is the additional ‘fifth’ force beyond the standard gravity. The range of this force depends on the two free

parameters and at  $z = 0$  can be quantified as

$$\lambda_\phi^0 = 3 \sqrt{\frac{n+1}{\Omega_M + 4\Omega_\Lambda}} \sqrt{\frac{|f_{R_0}|}{10^{-6}}} h^{-1} \text{ Mpc}, \quad (26)$$

which is usually expressed in units of  $h^{-1} \text{ Mpc}$ . As mentioned earlier we only look at a special case of the Hu-Sawicki model where  $n = 1$ . This allows us to fully define the  $f(R)$  model using a single parameter,  $|f_{R_0}|$ . See Llinares et al. (2014) for further details about this model and the reason for choosing this particular parametrization.

The numerical simulations analysed in this paper are taken from Llinares et al. (2014), which uses *ISIS*, a derivative of the adaptive mesh refinement (AMR) code *RAMSES* (Teyssier 2002), to fully evolve density fields sampled from an initial power spectrum generated by *LINGER* (Bode & Bertschinger 1995). We evaluate the derivatives using the *FORGE* and  $\Lambda$ CDM runs under the assumption that the central cosmology is given by  $|f_{R_0}| = 10^{-6}$  with the same spacing, i.e.  $\Delta|f_{R_0}| = 10^{-6}$ . Since the standard  $\Lambda$ CDM and the  $f(R)$  runs are both computed by the same codes with the same random seeds, the deviation between them purely results from the fifth force, allowing a robust calculation of the estimator derivative with respect to  $|f_{R_0}|$ . Note that the  $\Lambda$ CDM parameters of our previous runs (Sections 3.1 and 3.2) slightly differ from those in (Llinares



**Figure 4.** The relative constraints placed on each standard parameter (left) and alternate cosmological parameter (right) as smaller scales are included in the FIM computation. The largest scales ( $\geq 50 h^3 \text{ Mpc}^{-3}$ ) suffer from statistical noise due to lack of modes and, hence, have been omitted from the figure. We find a sharp increase in constraining power of the LCF when compared to the 2PCF across all parameters. For this analysis we assume a fiducial cosmological survey with  $\sigma_b = 0.13$ ,  $\bar{n} = 2 \times 10^{-3} h^3 \text{ Gpc}^{-3}$ , and  $V_{\text{eff}} = 1 h^{-3} \text{ Gpc}^3$ .

et al. 2014), but we assume that this does not impact the derivatives with respect to  $|f_{R_0}|$ . Another difference is that the  $\Lambda\text{CDM}$  and  $f(R)$  simulations of Llinares et al. (2014) use a higher spatial resolution ( $512^3$  particles in a  $L = 256 h^{-1} \text{ Mpc}$  box). This higher resolution is required to properly take into account non-linear effects associated with the screening mechanism. Despite this increase in resolution, we restrict the FI computations to correlation scales  $r \geq 10 h^{-1} \text{ Mpc}$  to be consistent with the previous cosmological models.

For the fiducial survey volume of  $V_{\text{eff}} = 1 h^{-3} \text{ Gpc}^3$  and a linear bias uncertainty of 13 percent (standard deviation), the *gain* on  $|f_{R_0}|$  constraints from including the LCF is  $g = 9.99$  as shown in Fig. 5 (middle). The orientations of the 2PCF and the LCF ellipses are different for all  $|f_{R_0}|$  and  $\Lambda\text{CDM}$  parameter pairs which leads to the extra constraining power of the combined estimators. For this MG model, we find a sharp increase in the information content of the LCF at small scales as compared to that of the 2PCF as shown in Fig. 4 (right). Hence, the LCF is more susceptible to non-linear growth in  $f(R)$  model and requires us to measure scale down to  $10 h^{-1} \text{ Mpc}$ .

### 3.4 Symmetron

The final cosmological model analysed in this paper is the Symmetron model. This MG model, which was originally explored by Hinterbichler & Khoury (2010), uses a scalar field  $\phi$  governed by a potential and a conformal factor equation

$$V(\phi) = -\frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4, \quad (27)$$

$$A(\phi) = 1 + \frac{1}{2}\frac{\phi^2}{M^2}, \quad (28)$$

where  $\mu$  and  $M$  are mass scales with  $\lambda$  being a positive dimensionless constant. These three parameters define the simplest Symmetron model.

In this model, the fluctuations of the scalar field couple to the matter in regions of low density  $\rho \lesssim \bar{\rho}$ , where  $\bar{\rho}$  is the average

matter density of the universe. These correspond to cosmic voids. However, in overdense regions, the coupling becomes negligible – an effect known as ‘screening’. The vacuum expectation value of the scalar  $\phi$  is

$$\phi_0 = \frac{\mu}{\sqrt{\lambda}} \quad (29)$$

and it determines the coupling strength between the scalar field and matter. For numerical convenience we rewrite the equations in terms of a dimensionless scalar field  $\chi = \phi/\phi_0$  using the following parameters

$$\lambda_0 = \frac{1}{\sqrt{2}\mu} \quad (30)$$

$$\beta = \frac{\phi_0 M_{\text{pl}}}{M^2}$$

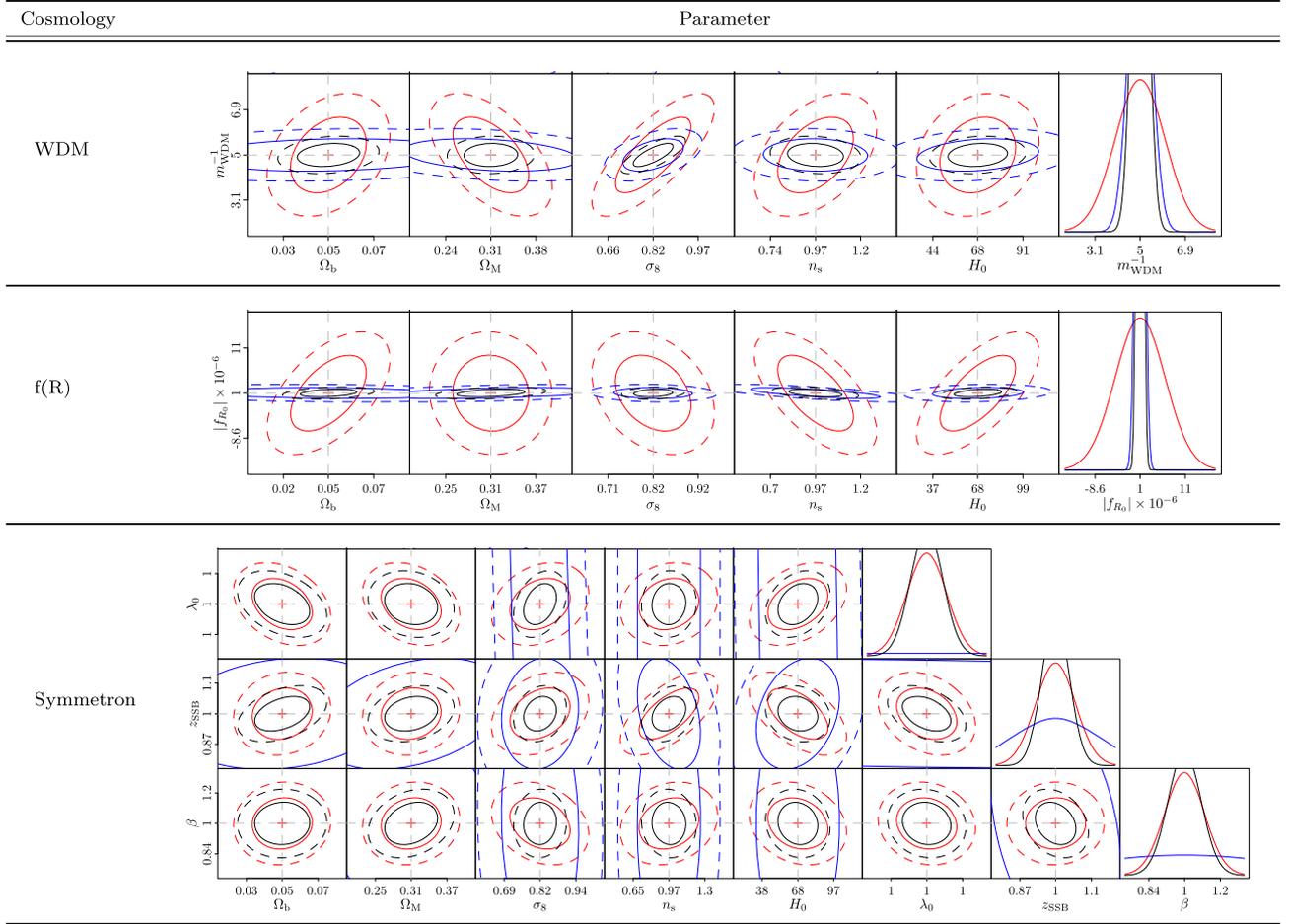
$$a_{\text{SSB}} = \frac{\rho_0}{\mu^2 M^2}$$

where  $M_{\text{pl}}$  and  $\rho_0$  are the Planckian mass and the background density at  $z = 0$ , respectively. The physical interpretation of the parameters  $\lambda_0$ ,  $\beta$ , and  $z_{\text{SSB}}$  are the range of the scalar field, its coupling strength and the redshift  $z_{\text{SSB}}$  (or its associated scale factor  $a_{\text{SSB}}$ ) at which the symmetry is broken (screening) for a given background cosmology, respectively. Using these parameters, the geodesic equation with the additional fifth force term becomes

$$\ddot{\mathbf{x}} + 2H\dot{\mathbf{x}} + \frac{\nabla\Phi}{a^2} - 6\frac{\Omega_M H_0^2}{a^2} \frac{\beta^2 \lambda_0^2}{a_{\text{SSB}}^3} \chi \nabla\chi = 0. \quad (31)$$

See Llinares et al. (2014) for full details of this model and the reason for choosing this parametrization.

We assume the Symmetron parameters are centred at  $\theta_0 = \{\lambda_0 (h^{-1} \text{ Mpc}), z_{\text{SSB}}, \beta\} = \{1, 1, 1\}$  and use the  $\Lambda\text{CDM}$  box in conjunction with `Symm A-C` runs from Llinares et al. (2014) to evaluate the derivatives. Fig. 5 (bottom) shows the parameter constraints obtained by employing the two estimators in our analysis. The gain of 2PCF+LCF, relative to the 2PCF alone, is about  $g = \{1.24, 1.48, 1.21\}$ . As with the previous MG model we find



**Figure 5.** The 2PCF (red), the LCF (blue), and the combined (black) constraints for each of the alternate cosmological parameters determined by our analysis. The colour scheme is same as that used in Fig. 3. The top, middle, and bottom rows show the allowable range of WDM,  $f(R)$ , and Symmetron parameters, respectively, along with their pairwise constraints with the standard parameters. The elliptical orientations of the standard-alternate parameter pairs are significantly different for the 2PCF and LCF, which allows a higher gain in constraining power for these models. The strongest constraints are comparable to that of  $\sigma_8$  which is the best determined standard parameter following the FI analysis. The analysis assumes a fiducial survey with 13 per cent bias uncertainty,  $2 \times 10^{-3} h^3 \text{ Gpc}^{-3}$  particle number density, and  $V_{\text{eff}} = 1 h^{-3} \text{ Gpc}^3$ .

the orientation of the Symmetron and  $\Lambda\text{CDM}$  pair ellipses having different orientations for the two estimators. This provides a boost in the combined constraints.

Fig. 4 (right) indicates the importance of incorporating smaller scales in our analysis. In the case of Symmetron model, the relative constraints from the LCF decrease faster than the 2PCF. Although, the individual constraining power of the 2PCF is better than that of the LCF, small-scale measurements still provide better combined constraint than standard parameters with the exception of  $\sigma_8$ .

### 3.5 Effect of the linear bias uncertainty

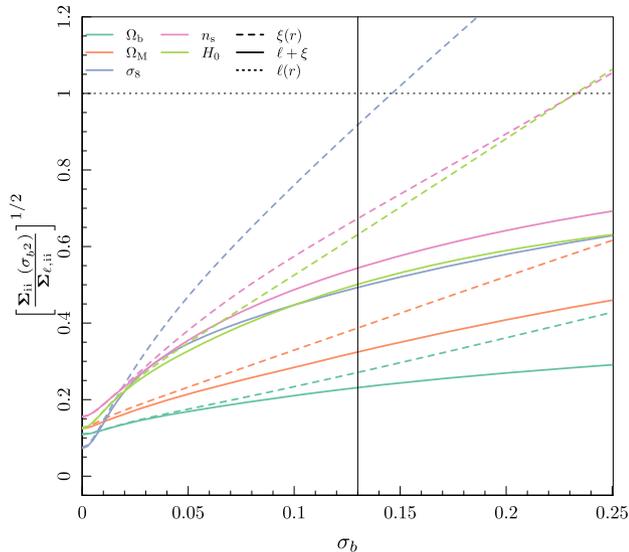
The statistical uncertainty of the galaxy (or halo) bias deteriorates the information extracted from cosmic large-scale structure. At linear order, this effect only applies to the 2PCF, not to the LCF, since the latter is insensitive to the linear bias  $b$ . Hence, the information of the LCF relative to the 2PCF increases with the uncertainty of  $b$ . So far, we have accounted for this effect assuming that  $b$  has a fixed normal uncertainty of  $\sigma_b = 13$  per cent. The variation of the relative FI with this uncertainty is shown in Fig. 6. Since the relative information in the 2PCF decreases with increasing  $\sigma_b$ , the combined information asymptotes to that of the LCF estimator as

$\sigma_b \rightarrow \infty$ . In all cosmological models the LCF adds significant extra constraints to the 2PCF, if  $\sigma_b = 0.13$  (as in Figs 3 and 5) or larger, whereas, in a scenario where bias is perfectly constrained, the 2PCF constrains each parameter by up to four times.

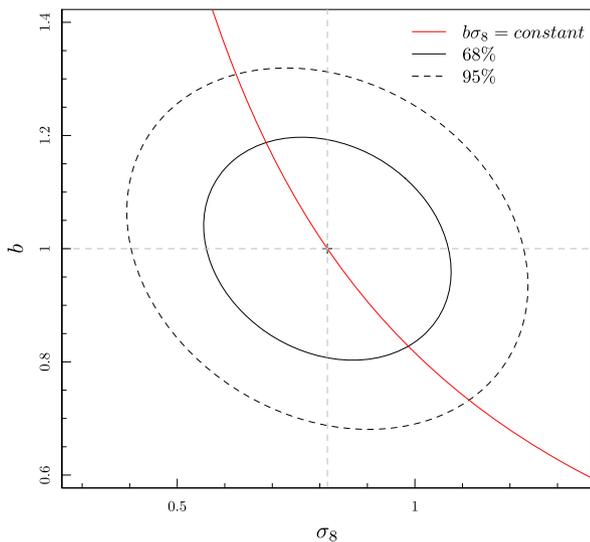
The discussion above highlights the power of combining the two-point statistics with a linear bias-independent estimator for parameter estimation. In turn, this property of the LCF also allows us to infer the bias parameter  $b$  itself. This can be achieved by considering  $b$  as an additional free parameter (see first approach discussed in Section 2.5) and computing the full FIM of all cosmological parameters and  $b$ . This requires the derivative of the galaxy 2PCF with respect to  $b$ , whose analytical expression is simply

$$\frac{\partial \xi_{\text{g},i}}{\partial b} = 2b\xi_i. \quad (32)$$

Since we now treat  $b$  as an additional model parameter, its uncertainty must not be included in the covariance via equation (18). Importantly, the linear bias  $b$  is only a good approximation of the full bias on large scales ( $r \geq 50 h^{-1} \text{ Mpc}$ ). Hence, we restrict this analysis to these large scales, and maintain the fiducial survey volume of  $V_{\text{eff}} = 1 h^{-3} \text{ Gpc}^3$ . The LCF at these scales is susceptible to statistical noise, hence, we use 50 L-PICOLA to better



**Figure 6.** The sensitivity of the  $\Lambda$ CDM constraints on the bias uncertainty,  $\sigma_b$ . The constraints placed by the LCF are independent of linear bias and, hence, have been used for normalization. The solid vertical line shows the 13 per cent uncertainty used in this paper. As this parameter increases the constraints asymptote to that of the LCF. The constraining power of the 2PCF pertaining to the  $\sigma_8$  parameter has the sharpest decrease since this parameter is highly correlated with linear bias. The analysis assumes a fiducial survey volume of  $V_{\text{eff}} = 1 h^{-3} \text{Gpc}^3$  and a space density of  $\bar{n} = 2 \times 10^{-3} h^3 \text{Gpc}^{-3}$ .



**Figure 7.** The bounds placed on the  $b$ - $\sigma_8$  using the combined information from the 2PCF and the LCF. The ellipses show the 68 per cent (solid) and 95 per cent (dotted) Gaussian contours inferred by only using length scales  $\geq 50 h^{-1} \text{Mpc}$ . For smaller scales the non-linear correction to bias becomes significant and hence the constraints become unreliable. If we were to carry out a full likelihood analysis we expect the constraints to be parallel to the  $b$ - $\sigma_8$  curve (solid red line). We assume a fiducial survey volume of  $V_{\text{eff}} = 1 h^{-3} \text{Gpc}^3$  to scale the covariance matrix and subsample particles to a final number density of  $\bar{n} = 2 \times 10^{-3} h^3 \text{Gpc}^{-3}$ .

estimate the large-scale derivatives of the 2PCF and the LCF for each standard parameter. The resulting combined constraints on  $\sigma_8$  and  $b$  are shown in Fig. 7. In the Laplace approximation (Gaussian likelihood), these constraints correspond to elliptical uncertainties.

However, from the collapse theory of haloes,  $\sigma_8 b$  is constant if the other cosmological parameters are fixed. This means that the degeneracy between  $\sigma_8$  and  $b$  corresponds to a hyperbola shown as solid curve in Fig. 7. As expected, the major axis of the ellipses roughly align with the direction of this hyperbola. Quantitatively, this analysis finds a linear bias uncertainty of  $\sigma_b \approx 13$  per cent, which we use throughout this study. We also compare the constraints found by the two methods for each of the  $\Lambda$ CDM parameters and find similar results with  $\sim 10$  per cent difference.

## 4 CONCLUSIONS

In this paper, we have combined two-point statistics and the line correlation function, three-point Fourier phase estimator, to infer cosmological parameters in standard and non-standard cosmological models. The Fisher information matrix was used to quantify the information in these estimators and derive parameter posteriors.

In the absence of linear bias uncertainties, the extra information provided by the LCF in addition to that already present in the 2PCF is marginal (about 5 per cent on average). However, for a fiducial linear bias uncertainty of 13 per cent and a survey volume of  $V_{\text{eff}} = 1 h^{-3} \text{Gpc}^3$ , we found that the addition of the LCF improves the 2PCF-based parameter constraints by a significant factor of about  $\sim 1.2$  (i.e. parameter uncertainties become 1.2-times smaller) in standard  $\Lambda$ CDM cosmology and by a factor up to  $\sim 1.2$ – $10$  in MG models and  $\sim 3.3$  in WDM cosmology. The relative information in the LCF increases with increasing bias uncertainty. However, to fully gauge the usefulness of the LCF we need to take into account non-linear bias terms in the computation. The constraints determined by the LCF in this study are likely an overestimate however our treatment simplifies the methodology and can be regarded as a first step towards more realistic constraints. To optimally benefit from the information in the LCF scales down to about  $r = 10 h^{-1} \text{Mpc}$  should be resolved. The combination of the 2PCF and the LCF can also be used to infer the linear bias, which is hard to measure otherwise and consequently to break the degeneracy between bias and  $\sigma_8$ .

Overall, these results advocate the use of Fourier phase statistics in addition to standard two-point statistics (2PCF of power spectrum) when inferring cosmological parameters from modern galaxy redshift surveys. The Fourier phase space is an excellent probe of the local interactions between the dark matter particles and even at larger scales before the full onset of non-linear gravitational growth the LCF provides more information about the properties of the underlying dark matter density field than the 2PCF. The gain from including the LCF is especially important in modified theories of gravity like in the  $f(R)$  and Symmetron models. This reflects the fact that modified gravity models possess a very rich phenomenology in the non-linear regime, or more precisely at the transition from the linear to the non-linear regime, where the screening mechanism takes place. As a consequence the LCF, which probes directly the emergence of non-linear correlations is very sensitive to these modifications of gravity and provides stronger constraints than the 2PCF.

The analysis of phase statistics comes with its own caveats. Apart from being computationally quite expensive, phase statistics such as the LCF are still difficult to measure in real surveys with irregularly shaped non-periodic volumes and complex selection functions. Clever methods for extracting such phase statistics, similarly to those used for the 2PCF (Landy & Szalay 1993), remain yet to be developed. Furthermore, redshift space distortions have their own

perturbing effects (Eggemeier et al. 2015), which, when harvested carefully, might be used for additional cosmological constraints.

## ACKNOWLEDGEMENTS

We thank our referee for helpful comments to improve the paper. This work was supported by the Research Collaboration Award 12105205 of the University of Western Australia. DO acknowledges support from Australian Research Council grants DP160102235. CLL acknowledges support from STFC consolidated grant ST/L00075X/1 & ST/P000541/1. This work used the DiRAC Data Centric system at Durham University. This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grants ST/H008519/1 and ST/K00087X/1, STFC DiRAC Operations grant ST/K003267/1 and Durham University. DiRAC is part of the National E-Infrastructure. Parts of this research were conducted by the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through project number CE110001020. Part of the simulations were done on the Sciama High Performance Compute (HPC) cluster which is supported by the ICG, SEPNet, and the University of Portsmouth. CB and FOF acknowledge support by the Swiss National Science Foundation.

## REFERENCES

- Alam S. et al., 2017, *MNRAS*, 470, 2617  
 Bode P., Bertschinger E., 1995, preprint ([arXiv:e-print](#))  
 Bode P., Ostriker J. P., Turok N., 2001, *ApJ*, 556, 93  
 Bonvin C., Hui L., Gaztanaga E., 2016, *J. Cosmol. Astropart. Phys.*, 8, 021  
 Byun J., Eggemeier A., Regan D., Seery D., Smith R. E., 2017, *MNRAS*, 471, 1581  
 Cai Y.-C., Padilla N., Li B., 2015, *MNRAS*, 451, 1036  
 Crocce M., Scoccimarro R., 2006, *Phys. Rev. D*, 73, 063519  
 da Cunha E. et al., 2017, *PASA*, 34, 47  
 DES Collaboration, 2017, preprint ([arXiv:1708.01530](#))  
 Dewdney P. E., Hall P. J., Schilizzi R. T., Lazio T. J. L. W., 2009, *IEEE Proc.*, 97, 1482  
 Duffy A. R., 2014, *Annalen der Physik*, 526, 283  
 Eggemeier A., Smith R. E., 2017, *MNRAS*, 466, 2496  
 Eggemeier A., Battefeld T., Smith R. E., Niemeyer J., 2015, *MNRAS*, 453, 797  
 Eisenstein D. J. et al., 2011, *AJ*, 142, 72  
 Falck B., Koyama K., Zhao G.-b., Cautun M., 2018, *MNRAS*, 475, 3262  
 Gil-Marín H., Schmidt F., Hu W., Jimenez R., Verde L., 2011, *J. Cosmol. Astropart. Phys.*, 11, 019  
 Gil-Marín H., Noreña J., Verde L., Percival W. J., Wagner C., Manera M., Schneider D. P., 2015, *MNRAS*, 451, 539  
 Hartlap J., Simon P., Schneider P., 2007, *A&A*, 464, 399  
 Hellwing W. A., Li B., Frenk C. S., Cole S., 2013, *MNRAS*, 435, 2806  
 Hinterbichler K., Khoury J., 2010, *Phys. Rev. Lett.*, 104, 231301  
 Howlett C., Percival W. J., 2017, *MNRAS*, 472, 4935  
 Howlett C., Manera M., Percival W. J., 2015, *Astron. Comput.*, 12, 109  
 Hu W., Sawicki I., 2007, *Phys. Rev. D*, 76, 064004  
 Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, *ApJ*, 522, 82  
 Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64  
 Laureijs R. et al., 2011, preprint ([arXiv:1110.3193](#))  
 Lewis A., Bridle S., 2002, *Phys. Rev. D*, 66, 103511  
 Llinares C., McCullagh N., 2017, *MNRAS*, 472L, 80L  
 Llinares C., Mota D. F., Winther H. A., 2014, *A&A*, 562, A78  
 Lombriser L., Simpson F., Mead A., 2015, *Phys. Rev. Lett.*, 114, 251101  
 Meiksin A., White M. J., 1999, *MNRAS*, 308, 1179  
 Murray S. G., Power C., Robotham A. S. G., 2013, *Astron. Comput.*, 3, 23  
 Ntelis P. et al., 2017, *J. Cosmol. Astropart. Phys.*, 6, 019  
 Obreschkow D., Power C., Bruderer M., Bonvin C., 2013, *ApJ*, 762, 115  
 Planck Collaboration. XIII. Cosmological parameters, 2016, *A&A*, 594, A13  
 Sabiu C. G., Mota D. F., Llinares C., Park C., 2016, *A&A*, 592, A38  
 Scoccimarro R., 1997, *ApJ*, 487, 1  
 Scoccimarro R., Zaldarriaga M., Hui L., 1999, *ApJ*, 527, 1  
 Sefusatti E., Vernizzi F., 2011, *J. Cosmol. Astropart. Phys.*, 3, 047  
 Springel V., 2005, *MNRAS*, 364, 1105  
 Tassev S., Zaldarriaga M., Eisenstein D. J., 2013, *J. Cosmol. Astropart. Phys.*, 6, 036  
 Tegmark M., Taylor A. N., Heavens A. F., 1997, *ApJ*, 480, 22  
 Teyssier R., 2002, *A&A*, 385, 337  
 The Dark Energy Survey Collaboration, 2005, preprint ([arXiv:e-print](#))  
 Voivodic R., Lima M., Llinares C., Mota D. F., 2017, *Phys. Rev. D*, 95, 024018  
 White M., 2016, *J. Cosmol. Astropart. Phys.*, 11, 057  
 Wolstenhulme R., Bonvin C., Obreschkow D., 2015, *ApJ*, 804, 132

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.