Beyond Kaiser bias: mildly non-linear two-point statistics of densities in distant spheres

C. Uhlemann,[1] S. Codis,[2]★ J. Kim,[3] C. Pichon,[3,4] F. Bernardeau,[4,5] D. Pogosyan,[4,6] C. Park[3] and B. L'Huillier[3,7]

[1]*Institute for Theoretical Physics, Utrecht University, Princetonplein 5, NL-3584 CC Utrecht, the Netherlands*
[2]*Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada*
[3]*Korea Institute for Advanced Study (KIAS), 85 Hoegiro, Dongdaemun-gu, Seoul 02455, Republic of Korea*
[4]*CNRS and UPMC, UMR 7095, Institut d'Astrophysique de Paris, F-75014 Paris, France*
[5]*CNRS and CEA, UMR 3681, Institut de Physique Théorique, F-91191 Gif-sur-Yvette, France*
[6]*Department of Physics, University of Alberta, 412 Avadh Bhatia Physics Laboratory, Edmonton, Alberta T6G 2J1, Canada*
[7]*Korea Astronomy and Space Science Institute, 776 Daedeok-Daero, Yuseong-gu, Daejeon 34055, Republic of Korea*

ABSTRACT

We present simple parameter-free analytic bias functions for the two-point correlation of densities in spheres at large separation. These bias functions generalize the so-called Kaiser bias to the mildly non-linear regime for arbitrary density contrasts and grow as $b(\rho) - b(1) \propto (1 - \rho^{-13/21})\rho^{1 + n/3}$ with $b(1) = -4/21 - n/3$ for a power-law initial spectrum with index n. We carry out the derivation in the context of large-deviation statistics while relying on the spherical collapse model. We use a logarithmic transformation that provides a saddle-point approximation that is valid for the whole range of densities and show its accuracy against the 30 Gpc cube state-of-the-art Horizon Run 4 simulation. Special configurations of two concentric spheres that allow us to identify peaks are employed to obtain the conditional bias and a proxy for the BBKS extremum correlation functions. These analytic bias functions should be used jointly with extended perturbation theory to predict two-point clustering statistics as they capture the non-linear regime of structure formation at the per cent level down to scales of about 10 Mpc h^{-1} at redshift 0. Conversely, the joint statistics also provide us with optimal dark matter two-point correlation estimates that can be applied either universally to all spheres or to a restricted set of biased (over- or underdense) pairs. Based on a simple fiducial survey, we show that the variance of this estimator is reduced by five times relative to the traditional sample estimator for the two-point function. Extracting more information from correlations of different types of objects should prove essential in the context of upcoming surveys like *Euclid*, DESI and WFIRST.

Key words: methods: analytical – methods: numerical – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

The large-scale structure of the Universe puts very tight constraints on cosmological models. Deep spectroscopic surveys, like *Euclid* (Laureijs et al. 2011), DESI (Levi et al. 2013), PFS (Takada, Ellis & Chiba 2014) or WFIRST (Spergel et al. 2013), will allow astronomers to study the details of structure formation at different epochs, hence to probe cosmic acceleration. Yet, in order to reach the expected precision on the equation of state of dark energy, astronomers must address the following challenges:

non-linear gravitational evolution (Bernardeau et al. 2002), redshift-space distortions (Kaiser 1987; Taruya, Nishimichi & Saito 2010), bias (Kaiser 1984; Dekel & Rees 1987), intrinsic alignments (see e.g. Chisari et al. 2015; Joachimi et al. 2015; Kiessling et al. 2015; Kirk et al. 2015; Tenneti et al. 2015; Troxel & Ishak 2015; Velliscig et al. 2015, for recent reviews and numerical works) and baryonic physics (Zhan & Knox 2004; Guillet, Teyssier & Colombi 2010; van Daalen et al. 2014; Schneider & Teyssier 2015). In this context, two-point clustering has generated a lot of interest (e.g. Cooray & Sheth 2002, and references therein), as it allows one to investigate how the densest regions of space – where dark haloes usually reside – are clustered, which in turn sheds light on the so-called biasing between dark matter and haloes: as haloes correspond to peaks of

★ E-mail: codis@cita.utoronto.ca

the density field, they are not a fair tracer of that field. Kaiser (1984) showed that in the high-density-contrast ($\delta/\sigma \gg 1$, where δ is the overdensity with respect to the mean density and σ^2 is the variance of the density field), large-separation limit, the correlation function, $\xi_{>\delta/\sigma}(r)$, of peaks lying above this threshold reads

$$\xi_{>\delta/\sigma}(r) \approx \frac{1}{\sigma^2}\left(\frac{\delta}{\sigma}\right)^2 \xi(r), \tag{1}$$

so that the correlation function of high-density regions decreases more slowly than the density field correlation function, $\xi(r)$, with an amplification factor or *bias* that is proportional to the threshold squared. This analysis can also be restricted to the peaks of the density field above a given threshold following the seminal papers by Bardeen et al. (1986, hereafter BBKS) and Regos & Szalay (1995). For two point functions, the non-linear regime increases the number of modes used to better constrain cosmological parameters. Since only a Gaussian field is characterized solely by its two-point correlation, while the non-linearity of gravitational interaction induces the cosmic density field to become strongly non-Gaussian over time, there is significant cosmological information beyond the correlation function (or equivalently the power spectrum). Accessing this information encoded in the strongly correlated density field requires more than just considering a hierarchy of correlation functions as shown in Carron & Neyrinck (2012) that stimulated the study of so-called sufficient statistics for large-scale structure in Carron & Szapudi (2013) that are closely related to considering power spectra of non-linearly transformed fields such as the log-density in Neyrinck, Szapudi & Szalay (2009). Of particular (partially theoretically unexplored) interest is the possibility of computing conditional two-point correlations, e.g. two-point correlation between regions that have specific densities, so as to provide more robust estimates of the large-distance two-point correlation. Related ideas have been partially pursued by considering so-called marked or sliced correlation functions, see for example early works of Sheth & Tormen (2004), White & Padmanabhan (2009) as well as Neyrinck et al. (2016) and White (2016) for more recent work that has been published after the submission of our work.

It has been argued (Bernardeau, Pichon & Codis 2014; Bernardeau, Codis & Pichon 2015; Codis et al. 2016a; Uhlemann et al. 2016) that the statistics of cosmic densities in concentric spheres can leverage cosmic parameters competitively, as the corresponding spherical symmetry allows for analytical predictions in the mildly non-linear regime, beyond what is commonly achievable via other statistics in the context of perturbation theory. Indeed, the zero variance limit of the cumulant generating functions yields estimates of the joint density probability distribution function (PDF hereafter) that seems to match simulations in the regime of variances of order unity (Balian & Schaeffer 1989; Bernardeau 1992; Juszkiewicz, Bouchet & Colombi 1993; Valageas 2002; Bernardeau et al. 2014, 2015). This success was shown to originate from a regime of large deviations at play in the mildly non-linear evolution of the large-scale structure (Bernardeau & Reimberg 2016).

The aim of this paper is to show that the spherically symmetric framework that led to surprisingly accurate predictions for one-point statistics also accommodates, in the large-separation limit, analytic estimates of the two-point statistics and in particular of the bias factor associated with imposed constraints within concentric cosmic densities. Recently, Codis, Bernardeau & Pichon (2016b) studied the two-point statistics of the density within concentric spheres, whose redshift evolution was shown to be accurately predicted by large-deviation theory in the mildly non-linear regime, but relied on numerical integration of highly oscillating complex functions and was therefore subject to possibly significant numerical errors, in particular for large densities. Since Uhlemann et al. (2016) showed that very accurate analytic approximations could be found for one-point statistics by using a logarithmic transform of the density field and performing a saddle-point approximation, we propose in this paper to extend the use of the logarithmic transform to two-point statistics. Analytic formulas for the bias induced in high- or low-density regions are desirable because they allow for a simple assessment of the sensitivity to changes in the initial conditions (such as primordial non-Gaussianity) and the subsequent evolution (such as dark energy). This is advantageous compared to pure simulation-based approaches for which those investigations are computationally expensive, especially since the determination of large-scale correlations requires a large volume. Note that applying the theory presented here to actual galaxy surveys should allow us to account for the so-called assembly bias (Dalal et al. 2008), a residual variation in galaxy clustering with different galaxy properties (e.g. galaxy luminosity) at fixed halo mass, which is thought to originate from their assembly history. Indeed, by considering the joint bias measured at two different scales, we can quantify how much the clustering amplitude depends on the density on larger scale that is the essence of the idea behind assembly bias or peak-background split (BBKS).

It was shown in Codis et al. (2016a) that the one-point PDF can be fully predicted, modulo one parameter, the variance of the density field, which is the driving parameter of the theory, leading to two options: (i) higher order perturbation theory can be used to predict the value of this variance as a function of scale and redshift in order to recover the full PDF or (ii) this one-parameter model can be used to build optimal likelihood estimators for the variance based on the measurement of densities in spheres. Conversely, in the present paper, modulo the unknown underlying two-point correlation function of the dark matter density field, we will show that the same large-deviation formalism provides us with the full statistics of the two-point PDF of the density within concentric spheres separated by a distance r_e. Once again, one can (i) rely on perturbation theory to predict the underlying dark matter correlation function (e.g. Taruya et al. 2012), or (ii) build, from the present theory, optimal estimators for the dark matter correlation function to be applied to measured density in separated spheres.

In this paper, following Bernardeau (1996), the focus will be on predicting analytically the density two-point statistics for configurations shown in Fig. 1 and specifically the corresponding bias functions (the aforementioned density-dependent scalings of the two-point correlation). We will in particular consider the density field smoothed at two different scales in two concentric spheres that can be turned into an inner density and a slope (difference of density between the two spheres). This will allow us to focus on the conditional density-given-slope bias as a quasi-linear proxy for the BBKS peak correlation function. These bias functions generalize the so-called Kaiser linear bias in the mildly non-linear regime for large separations and arbitrary density contrasts. Hence, they provide alternative ways of using gravitational clustering to probe our cosmological model, in particular using specific regions of space (underdense/overdense, small/big slope, etc.). Leveraging conditionals on the value of the density at the legs of the correlations will allow for a more robust estimate of the two-point correlation function. We will illustrate in a fiducial experiment how the present formalism can be used to estimate optimally the underlying top-hat filtered correlation function.
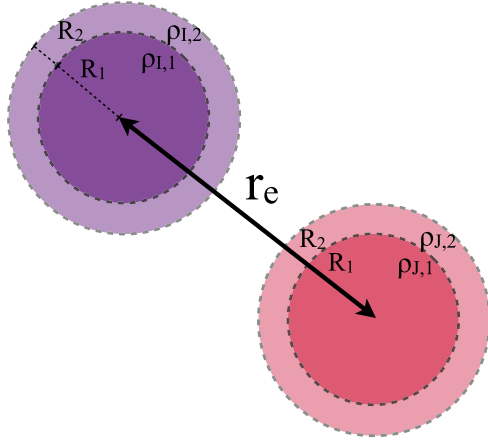
Figure 1. The two-point configuration with two concentric cells of radii R_1 and R_2 in one location (purple) and two other concentric cells of radii R_1 and R_2 in another location (red) separated by a distance $r_e \gg R_2$.

This paper is organized as follows. Section 2 presents briefly the implementation of a large-deviation principle on the joint statistics of concentric cells based on a saddle-point approximation. Section 3 compares these analytic predictions to the state-of-the-art dark matter simulation Horizon Run 4 (HR4). Section 4 demonstrates how to measure optimally the dark matter correlation function on a given survey. Finally, Section 5 wraps up. Appendix A shortly describes the accompanying package LSSFAST for the evaluation of the PDF and the bias function of a density in a sphere. Appendix B reviews the formalism of large deviations relevant to obtain the density PDF for concentric spheres and the joint PDF at large separations. Appendix C provides a description of bias functions in the Gaussian and weakly non-Gaussian regime based on perturbation theory. Appendix D provides a validation of HR4 at redshift $z = 4$ together with extended results for redshift $z = 0$.

2 LARGE DEVIATIONS AND SEPARATION

Appendix B presents rapidly the general formalism for deriving the PDF and the bias functions from the large-deviation principle. Let us in the main text focus on presenting directly the corresponding fully analytical predictions, relying on the so-called saddle-point approximation. For that purpose, we shall see that all that is needed is the so-called decay-rate function (controlling the exponential decay of the probability distribution with the density at leading order), the key quantity that connects the (Gaussian) initial to the non-Gaussian final distribution brought about by non-linear gravitational clustering.

In this section, we will determine the PDF of densities in *distant spheres* based on large-deviation statistics and the saddle-point approximation. This approximation has proven successful already to predict the one- and two-cell PDF for densities in *concentric spheres* in Uhlemann et al. (2016). After introducing the two-point PDF and the bias functions in Section 2.1 and the spherical collapse model as the mapping for large-deviation statistics in Section 2.2, we shortly review results for the one-point PDF in Section 2.3 and finally extend them to the two-point PDF of densities at large separations, hence the bias functions in Section 2.4.[1] The joint knowledge of the

PDF and the bias functions enables us to compute constrained bias functions that are of particular interest for the clustering of over- and underdense regions.

2.1 Definition of the bias functions

Let us consider two sets of N concentric spheres of radii $\{R_k\}$, separated by a distance r_e, and define the corresponding densities $\{\rho_k\}_{1 \le k \le N}$ and $\{\rho'_k\}_{1 \le k \le N}$. Following Bernardeau (1996) and Codis et al. (2016b), the joint PDF of these densities, $\mathcal{P}(\{\rho_k\}, \{\rho'_k\}; r_e)$, can be predicted from the one-point PDFs $\mathcal{P}(\{\rho_k\})$, $\mathcal{P}(\{\rho'_k\})$ and a correlation function $\xi_\circ$ of densities in spheres at finite separation

$$\mathcal{P}(\{\rho_k\}, \{\rho'_k\}; r_e) = \mathcal{P}(\{\rho_k\})\mathcal{P}(\{\rho'_k\}) \left[1 + \xi_\circ(\{\rho_k\}, \{\rho'_k\}; r_e)\right] . \quad (2a)$$

At large separations of the spheres $r_e \gg R_k$, which will at least require the distant spheres to not overlap, the sphere correlation function, $\xi_\circ(\{\rho_k\}, \{\rho'_k\}; r_e)$, can be related to the underlying (unbiased) top-hat smoothed dark matter two-point correlation function $\xi(r_e)$ via effective bias functions $b(\{\rho_k\})$ and $b(\{\rho'_k\})$:

$$\xi_\circ(\{\rho_k\}, \{\rho'_k\}; r_e) = \xi(r_e)b(\{\rho_k\})b(\{\rho'_k\}) . \quad (2b)$$

This is the count-in-cell analogue of the so-called peak-background split and defines the bias factor $b(\{\hat{\rho}_k\})$. Physically, this bias encodes the mean density in a sphere given that the densities in spheres of radii $\{R_k\}$ at large separation r_e are $\{\hat{\rho}_k\}$

$$1 + b(\{\hat{\rho}_k\})\xi(r_e) = \langle \rho' | \{\hat{\rho}_k\} ; r_e \rangle ,$$
$$= \frac{\int_0^\infty d\rho' \mathcal{P}(\{\hat{\rho}_k\}, \rho'; r_e)\rho'}{\mathcal{P}(\{\hat{\rho}_k\})} , \quad (3)$$

which follows from equation (2a) by integration and normalization of the PDFs. In the following, we will describe how this bias function, and hence the joint PDF of densities in spheres at large separation, can be predicted analytically.

2.2 Large-deviation statistics with spherical collapse

When considering a highly symmetric observable such as densities in spheres, one can argue that the most likely dynamics (amongst all possible mappings between the initial and final density field) is the one respecting the symmetry (Valageas 2002).[2] For spherical symmetry, one can then take advantage of the fact that non-linear solutions to the gravitational dynamics are known explicitly in terms of the spherical collapse model.

Let us denote $\zeta_{SC}(\tau)$ the non-linear transform of an initial fluctuation with linear density contrast, τ, in a sphere of radius r, to the final density ρ (in units of the average density) in a sphere of radius R according to the spherical collapse model

$$\rho = \zeta_{SC}(\tau), \quad \text{with} \quad \rho R^3 = r^3, \quad (4)$$

where the initial and final radii are connected through mass conservation. An explicit possible fit for $\rho_{SC}(\tau)$ is given by

$$\rho_{SC}(\tau) = (1 - \tau/\nu)^{-\nu}, \quad (5)$$

density within the inner sphere and a second density within a given outer shell, one could build correlators for arbitrary separation. This will be the topic of upcoming investigations.

[2] This is a result of the so-called contraction principle in the context of large-deviation theory as explained in Bernardeau & Reimberg (2016), which formalizes the idea that amongst all unlikely fates (in the tail of the PDF) the least unlikely one (spherical collapse) dominates.

---

[1] Note that, by constructing PDFs for densities in a larger number of concentric spheres in special configurations, such as three spheres, with a central

where the collapse parameter ν can be adjusted to the actual values of the cosmological parameters ($\nu = 21/13$ provides a good description of the spherical dynamics for an Einstein–de Sitter background for the range of τ values of interest).

Thanks to this analytic spherical collapse model, the PDF of densities in concentric spheres and the corresponding bias functions, brought about by non-linear gravitational evolution, can be predicted explicitly from the given (Gaussian) initial conditions.

2.3 The density PDF in the large-deviation regime

2.3.1 PDF and decay-rate function for an initial Gaussian field

The principles of large-deviation statistics yield a formula for the probability of finding a certain density in a sphere given the initial conditions. The decay-rate function Ψ encodes the exponential decay of the PDF with increasing deviation from the mean. For Gaussian initial conditions, which we assume here, the PDF can be written in the following form:

$$\mathcal{P}_{\{r_k\}}^{\mathrm{ini}}(\{\tau_k\}) = \sqrt{\det\left[\frac{\partial^2 \Psi_{\{r_k\}}^{\mathrm{ini}}}{\partial \tau_i \partial \tau_j}\right] \frac{\exp\left[-\Psi_{\{r_k\}}^{\mathrm{ini}}(\{\tau_k\})\right]}{(2\pi)^{N/2}}}, \qquad (6)$$

the initial decay-rate function Ψ^{ini} is given by the usual quadratic form in the initial density contrasts τ_k within concentric spheres of radii r_k

$$\Psi_{\{r_k\}}^{\mathrm{ini}}(\{\tau_k\}) = \frac{1}{2} \sum_{i,j} \Xi_{ij}(\{r_k\})\, \tau_i \tau_j\,, \qquad (7)$$

where Ξ_{ij} is the inverse of the initial covariance matrix of densities in spheres of radii r_k, $\sigma_{ij}^2 = \sigma^2(r_i, r_j)$, encoding all dependence with respect to the initial power spectrum P^{lin} according to

$$\sigma_{ij}^2 = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} P^{\mathrm{lin}}(k) W_{\mathrm{3D}}(kr_i) W_{\mathrm{3D}}(kr_j)\,, \qquad (8)$$

where W_{3D} is the Fourier transform of the top-hat filter

$$W_{\mathrm{3D}}(k) = \frac{3}{k^2}(\sin(k)/k - \cos(k))\,. \qquad (9)$$

Note that equation (6) is an unusual rewrite of a Gaussian distribution, emphasizing the central role of the rate function (7). This rate function has a straightforward explicit expression in terms of the underlying covariances, hence the initial power spectrum.

2.3.2 Saddle-point PDF for an evolved non-Gaussian field

The final decay-rate function Ψ is obtained from re-expressing the initial decay-rate function Ψ^{ini} in terms of the final densities ρ_k within concentric spheres of radii R_k

$$\Psi_{\{R_k\}}(\{\rho_k\}) = \frac{1}{2} \sum_{i,j} \Xi_{ij}(\{R_k \rho_k^{1/3}\})\, \tau_i(\rho_i) \tau_j(\rho_j)\,, \qquad (10)$$

using the spherical collapse mapping characterized by equation (5) and mass conservation from equation (4). The previously known result (obtained from a saddle-point approximation) for the PDF of densities in concentric spheres at one point is given by

$$\mathcal{P}_{\{R_k\}}(\{\rho_k\}) = \sqrt{\det\left[\frac{\partial^2 \Psi_{\{R_k\}}}{\partial \rho_i \partial \rho_j}\right] \frac{\exp\left[-\Psi_{\{R_k\}}(\{\rho_k\})\right]}{(2\pi)^{N/2}}}, \qquad (11)$$

where the decay-rate function Ψ is given by equation (10). The saddle-point approximation provides a very good approximation

to the exact result from large-deviation statistics, as discussed in Appendix B, if the final decay-rate function is convex, i.e.

$$\det\left[\frac{\partial^2 \Psi_{\{R_k\}}}{\partial \rho_i \partial \rho_j}\right] > 0\,,$$

which simplifies to $\Psi_R''(\rho) > 0$ for the one-cell case. However, as has been shown in Bernardeau et al. (2014), this condition is only fulfilled below a critical density value ρ_{c}, where $\Psi_R''(\rho_{\mathrm{c}}) = 0$ in the one-cell case, and similarly inside an $N - 1$ manifold for the N-cell case. The main point of Uhlemann et al. (2016) is that this difficulty can be alleviated with an adequate change of variables such as a logarithmic density transform. The procedure is then to apply the saddle-point approximation to predict the PDF of the (logarithmically) mapped densities μ_k as

$$\mathcal{P}_{\mu,\{R_k\}}(\{\mu_k\}) = \sqrt{\det\left[\frac{\partial^2 \Psi_{\{R_k\}}}{\partial \mu_i \partial \mu_j}\right] \frac{\exp\left[-\Psi_{\{R_k\}}\right]}{(2\pi)^{N/2}}}, \qquad (12a)$$

where the transformation $\{\rho_k\} \to \{\mu_k\}$ has to be chosen to ensure the convexity of the decay-rate function. This result can then easily be translated into the PDF of the density via a change of variables,

$$\mathcal{P}_{\{R_k\}}(\{\rho_k\}) = \mathcal{P}_{\mu,\{R_k\}}[\{\mu_k(\{\rho_i\})\}]\left|\det\left[\frac{\partial \mu_i}{\partial \rho_j}\right]\right|, \qquad (12b)$$

where the Hessian of the decay-rate function $\Psi_{\{R_k\}}$ after a change of variables $\{\rho_k\} \to \{\mu_k\}$ is trivially given by

$$\frac{\partial^2 \Psi_{\{R_k\}}}{\partial \mu_i \partial \mu_j} = \frac{\partial \rho_k}{\partial \mu_i} \cdot \frac{\partial^2 \Psi_{\{R_k\}}}{\partial \rho_k \partial \rho_l} \cdot \frac{\partial \rho_l}{\partial \mu_j} + \frac{\partial^2 \rho_k}{\partial \mu_i \partial \mu_j} \cdot \frac{\partial \Psi_{\{R_k\}}}{\partial \rho_k}\,.$$

2.3.3 Ensuring normalization

Equation (12) assumes that the means of the log-transformed densities μ_j do not depend on the variance and vanishes. For a generic non-linear mapping, it will translate into a mean density that can deviate from one as the variance σ grows. In order to account for this effect, one has to consider the shifted PDF

$$\hat{\mathcal{P}}_{\mu,\{R_k\}}(\{\mu_k\}) = \mathcal{P}_{\mu,\{R_k\}}(\{\tilde{\mu}_k = \mu_k - \langle \mu_k \rangle\})\,, \qquad (12c)$$

with the shifts $\langle \mu_k \rangle$ chosen such that the resulting mean densities are one $\langle \rho_i \rangle = 1\ \forall i = 1, \ldots, n$. Furthermore, since the saddle-point method yields only an approximation to the exact PDF, the PDF obtained from equation (12) is not necessarily properly normalized. In practice, this can be accounted for by considering

$$\hat{\mathcal{P}}_R(\{\rho_k\}) = \mathcal{P}_R(\{\rho_k\})/\langle 1 \rangle\,, \qquad (12d)$$

with the shorthand notation $\langle 1 \rangle = \prod_k \int_0^\infty \mathrm{d}\rho_k\, \mathcal{P}_R(\{\rho_k\})$.

2.4 The bias functions at large separations

The saddle-point approximation applies also to the joint PDF of densities in spheres at large separation and hence the bias function defined in equation (3). Since initially, the field is Gaussian, the initial bias function is exactly given by the so-called Kaiser linear bias as described in Section 2.4.1. The subsequent quasi-linear evolution can then be predicted by the large-deviation principle as will be shown in Section 2.4.2.

2.4.1 Kaiser bias for an initial Gaussian field

Let us consider a set of density contrasts $\{\tau_k\}$ in concentric spheres of radii r_k and the contrast τ_1' in a sphere of radius r_1 at a distance r_e

away from the centre of the concentric spheres. If the density field is Gaussian, the covariance matrix C of $(\{\tau_k\}, \tau_1')$ simply reads

$$C = \begin{pmatrix} \sigma_{ij}^2 & \xi_{i1} \\ \xi_{1j} & \sigma_{11}^2 \end{pmatrix}, \qquad (13)$$

where we use the shorthand notation $\sigma_{ij}^2 = \sigma^2(r_i, r_j)$ for the covariance and $\xi_{ij} = \xi(r_i, r_j; r_e)$ for the cross-correlation with

$$\xi_{ij} = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} P^{\mathrm{lin}}(k) W_{3D}(kr_i) W_{3D}(kr_j) \exp(\iota k r_e \cos\theta).$$

In the Gaussian case, the bias function b^{G} can be computed analytically, for example by diagonalizing the covariance matrix by a change of variables as shown in Appendix C. The result for the one-cell density bias is

$$b^{\mathrm{G}}(\tau) = \frac{\langle \tau_1' | \tau \rangle}{\xi_{11}} = \frac{\tau}{\sigma_{11}^2}, \qquad (14)$$

which is proportional to the initial overdensity τ as expected from Kaiser (1984). The N-cell density bias follows as

$$b^{\mathrm{ini}}(\{\tau_k\}) = \sum_{i,j=1}^{N} \Xi_{ij}(r_i, r_j)\tau_i = \sum_{j=1}^{N} \frac{\partial \Psi_{\{r_k\}}^{\mathrm{ini}}(\{\tau_k\})}{\partial \tau_j}, \qquad (15)$$

if we assume that for large separations $r_e \gg r_k$ the cross-correlations are all approximately identical $\xi_{1i} \approx \xi_{11} \, \forall i$. In general, the Kaiser bias function is given in terms of the derivative of the decay-rate function of the initial PDF $\Psi_{\{r_k\}}(\{\tau_k\}) = -\log \mathcal{P}_{\{r_k\}}(\{\tau_k\})$ and hence the rate of decay of the PDF. This encodes the idea that unlikely configurations, corresponding to strongly positive or negative values of the initial density contrast τ, are more biased.

2.4.2 Saddle-point bias for an evolved non-Gaussian field

The saddle-point approximation of the bias function amounts to mapping the initial Kaiser bias function, equation (15), using the inverse spherical collapse dynamics from equation (5) together with mass conservation from equation (4),

$$b(\{\rho_k\}) = \sum_{i,j=1}^{n} \Xi_{ij}(R_i \rho_i^{1/3}, R_j \rho_j^{1/3})\tau_i(\rho_i). \qquad (16)$$

The spherical collapse can be shown to be the leading-order contribution for the statistics of densities in distant spheres, as was done for the PDF of densities in concentric spheres.[3] This saddle-point approximation is valid in the large-separation regime and as long as the PDF of the density can be obtained via a saddle-point approximation. In as much as the logarithmic transform significantly increases the region of applicability of the saddle-point approximation for the PDF, it also yields analytical bias functions. Using this saddle-point approximation, we therefore extend the results of Codis et al. (2016a) and present analytical predictions that do not require a numerical integration in the complex plane. This allows us to also provide predictions for the joint and constrained biases based on analytical predictions of two-cell quantities that were not easily accessible before.

---

[3] More precisely, the analytical asymptote of the bias function can be derived using a steepest descent method, see Appendix B.

2.4.3 Ensuring normalization

Because of the normalization of the density and joint density PDFs and the definition of the sphere correlation function $\xi_\circ$, the bias function b must obey the following two relations

$$\langle b(\{\rho_k\}) \rangle = \prod_k \int_0^\infty \mathrm{d}\rho_k \, \hat{\mathcal{P}}(\{\rho_k\}) b(\{\rho_k\}) = 0,$$

$$\langle \rho_i b(\{\rho_k\}) \rangle = \prod_k \int_0^\infty \mathrm{d}\rho_k \, \hat{\mathcal{P}}(\{\rho_k\}) b(\{\rho_k\})\rho_i = 1, \, \forall i = 1, \ldots, N.$$

$$(17)$$

Because of these properties, the bias function $b(\{\rho_k\})$ obtained from equation (16) still has to be normalized according to

$$\hat{b}(\{\rho_k\}) = \frac{b(\{\rho_k\}) - \langle b(\{\rho_k\}) \rangle}{\frac{1}{N} \sum_{i=1}^{N} (\langle \rho_i b(\{\rho_k\}) \rangle - \langle b(\{\rho_k\}) \rangle)}. \qquad (18)$$

This normalization procedure is necessary and can be understood easily: while the conditions (18) are trivially fulfilled for a purely Gaussian initial field with small variance and Kaiser bias (15), the bias will pick up corrections from non-Gaussianity via gravitational collapse. Those corrections modify the value of the bias at average density (and hence the mean bias and mean density-weighted bias) and become manifest already in the very mildly non-Gaussian (and hence perturbative) regime as shown in Appendix C. The non-Gaussian corrections that affect the mean are not accounted for in the saddle-point approximation used in this work. To correct for this effect, we choose to shift the non-perturbative result from spherical collapse dynamics at the end according to the non-zero mean bias and will show that it leads to accurate predictions that are robust to variances of order one. Furthermore, extrapolating the saddle-point approximation to finite variances requires to adjust the absolute normalization connected to the mean density-weighted bias.

The reason why we expect this procedure to work is the following. If the bias function is not obtained from a saddle-point approximation but from the exact integral of the inverse Laplace transform that is presented in equation (B9), computed using the cumulant generating function (obtained in the limit $\sigma \to 0$), one obtains bias functions that are automatically normalized correctly. This has been demonstrated in Codis et al. (2016b) and the same is true for obtaining the PDF from computing the inverse Laplace transform from equation (B4) exactly as demonstrated in Bernardeau et al. (2014). This property, as already noted in Valageas (2002), is ensured by the fact that the cumulant generating functions at leading order in the variance (the quadratic part in the expansion of $\varphi(\lambda)$ for $\lambda \to 0$) are enough to obtain the exact first two moments. Hence, the role of the normalization procedure is to improve the saddle point approximation to more closely resemble the full result obtained from the numerical computation of the inverse Laplace transform

3 VALIDATION WITH THE HR4 SIMULATION

Let us now evaluate the simple analytical predictions for the density PDF in concentric spheres, equation (11), together with the bias functions at large separations, equation (16), and compare them to measurements in the HR4 simulation that is presented in Section 3.1. The estimators for the measurements of the bias functions are described in Section 3.2 while the parametrization for the covariance matrix is given in Section 3.3. For brevity, we will focus our comparison in the main text to redshift $z = 0.7$, which is in the redshift range that is most interesting to current surveys, while
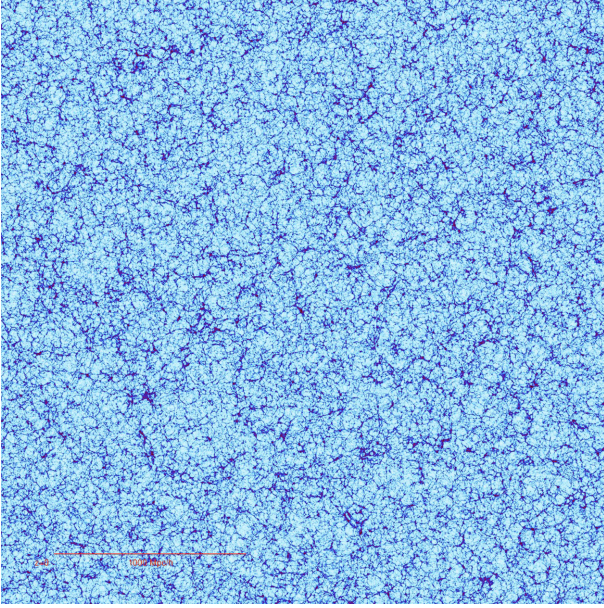
Figure 2. A thin $(1.5\,\mathrm{Mpc}\,h^{-1})$ slice through the HR4 simulation at redshift 0. The simulation is $3150\,\mathrm{Mpc}\,h^{-1}$ across sampled by 6300^3 dark matter particles.

results for $z = 0$ and a validation at high redshift $z = 4$ are shown in Appendix D.

We will validate our analytical results in Section 3.4 against the HR4 measurements, then present conditional bias function in Section 3.5 and discuss the modulation of the matter correlation function induced by biasing in Section 3.6.

3.1 The HR4 simulation

The HR4 simulation (Kim et al. 2015) is a state-of-the-art dark matter simulation modelling gravitational clustering in a Hubble-like volume. It was run on the Tachyon-2 system at Korea Institute of Science and Technology Information to study the non-linear matter evolution using $N_p = 6300^3$ particles in a cubic box with a side length of $3150\,h^{-1}$ Mpc. The particle mass is about $m_p \simeq 9.0 \times 10^9\,h^{-1}\,\mathrm{M}_\odot$. The adopted cosmology is comparable with the *WMAP* 5 year Λ cold dark matter model with $\sigma_8 = 0.794$ and matter, baryonic matter and dark energy density parameters of $\Omega_{m,0} = 0.26$, $\Omega_{b,0} = 0.044$ and $\Omega_{\Lambda,0} = 0.74$, respectively. The initial conditions were generated at redshift $z_i = 100$ according to the second-order Lagrangian perturbation theory (Jenkins 2010). The gravitational force on each particle was calculated using the PM-Tree method by the GOTPM (Dubinski et al. 2004) down to redshift $z = 0$ in 2000 global time-steps, see Fig. 2. The force resolution is $0.05\,h^{-1}$ Mpc, i.e. 1/10 of the mean particle separation.

To enhance the positional accuracy of particles, we adopt the shift vector rather than the position vector for particle information. In such a huge simulation like HR4, the 32-bit floating-point accuracy has a round-off error around sub-per cent level in terms of the mean particle separation, and a large time-step evolution may accumulate such errors and may affect the final matter distributions. Without requesting any more memory space, we devised a trick using the particle index as the Lagrangian position and the shift vector to calculate the particle position. For more details about the simulation and methods adopted by the GOTPM, see Kim et al. (2015).

3.2 Measuring the bias functions

To measure the bias functions b, we will make use of equation (3) applied to a specific number of concentric spheres. In practice, we determine the values of the PDFs in bins of a certain width around the given value by counting the number of densities in a given bin using stepwise functions. The HR4 simulation density is estimated via dark matter count in cell in $N_t = 252^3$ cells separated by $12.5\,\mathrm{Mpc}\,h^{-1}$. We use spheres of radii $R = 10, 11, 12, 13, 14, 15\,\mathrm{Mpc}\,h^{-1}$ with a separation $r_e = 37.5\,\mathrm{Mpc}\,h^{-1}$ that is big enough to ensure that we are in the large-separation regime, see Codis et al. (2016b).

Correlation function. We measure the sphere correlation function $\hat{\xi}$ at distance r_e according to

$$\hat{\xi}(r_e) = \frac{\sum_{I=1}^{N_t} \sum_{j=1}^{6} \rho_I \rho_{\alpha_{I,j}}}{6N_t} - 1\,, \tag{19}$$

where ρ_I is the density in the sphere at grid point I and $\alpha_{I,j}$ labels the six neighbouring grid points at separation r_e. In practice, we count all pairs of spheres only once when computing $\hat{\xi}(r_e)$ by only considering three neighbours for each sphere.

Density bias. The density bias $b(\rho)$ is estimated using the cross-correlations of densities ρ in spheres with radius R defined in equation (3) for N concentric spheres that reduces to equation (26) for the case of single spheres. More precisely, we compute a density-weighted sum including the neighbours $\alpha_{I,j}$ of each sphere I with density ρ_I close to $\hat{\rho}$,

$$\hat{b}(\hat{\rho}) = \frac{1}{\hat{\xi}} \left[\frac{\sum_I \sum_{j=1}^{6} \mathcal{B}(|\rho_I - \hat{\rho}| \le \Delta\rho/2)\rho_{\alpha_{I,j}}}{6\sum_I \mathcal{B}(|\rho_I - \hat{\rho}| \le \Delta\rho/2)} - 1 \right], \tag{20}$$

where $\mathcal{B}$ is a Boolean function that evaluates to one if the density is in a bin centred on $\hat{\rho}$ with width $\Delta\rho = 0.1$. Note that the density bias can also be measured using auto-correlations that has been shown to give results consistent with cross-correlations, see Codis et al. (2016b).

Joint density–slope bias. In order to measure the joint density–slope bias in the simulation, where the slope $s = \rho_2 - \rho_1$ is the density difference between two spheres, we consider two spheres with similar radii $R_{1/2}$ such that $R_2 - R_1 = 1\,\mathrm{Mpc}\,h^{-1}$. By reducing equation (3) to two concentric spheres, explicitly written in equation (28), we compute again a density-weighted sum over all grid points I with density and slope close to the desired values including the neighbouring grid points $\alpha_{I,j}$ at distance r_e,

$$\hat{b}(\hat{\rho}, \hat{s}) = \frac{1}{\hat{\xi}} \left[\frac{\sum_{I,j} \mathcal{B}(|\rho_I - \hat{\rho}| \le \Delta\rho/2, |s_I - \hat{s}| \le \Delta s/2)\rho_{\alpha_{I,j}}}{6\sum_I \mathcal{B}(|\rho_I - \hat{\rho}| \le \Delta\rho/2, |s_I - \hat{s}| \le \Delta s/2)} - 1 \right],$$

where $\mathcal{B}$ is a Boolean function that evaluates to one if the density is in a bin centred on $\hat{\rho}$ with width $\Delta\rho = 0.2$ and the slope is in a bin centred on $\hat{s}$ with width $\Delta s = 0.02R_1$.

Constrained density bias given environment. From the joint bias of the density and slope, one can also determine a constrained density bias given an environment either specified by a positive or negative slope s,

$$\hat{b}(\hat{\rho}|s \lessgtr 0) = \frac{1}{\hat{\xi}} \left[\frac{\sum_{I,j} \mathcal{B}(|\rho_I - \hat{\rho}| \le \Delta\rho/2, s_I \lessgtr 0)\rho_{\alpha_{I,j}}}{6\sum_I \mathcal{B}(|\rho_I - \hat{\rho}| \le \Delta\rho/2, s_I \lessgtr 0)} - 1 \right],$$

or an over- or underdense shell $\rho_{12} = (R_2^3\rho_2 - R_1^3\rho_1)/(R_2^3 - R_1^3)$

$$\hat{b}(\hat{\rho}|\rho_{12} \lessgtr 1) = \frac{1}{\hat{\xi}} \left[\frac{\sum_{I,j} \mathcal{B}(|\rho_I - \hat{\rho}| \le \Delta\rho/2, \rho_{12,I} \lessgtr 1)\rho_{\alpha_{I,j}}}{6\sum_I \mathcal{B}(|\rho_I - \hat{\rho}| \le \Delta\rho/2, \rho_{12,I} \lessgtr 1)} - 1 \right],$$

which are both measured with bins of width $\Delta\rho = 0.1$.

Table 1. Variances of the density ρ and the log-density $\mu = \log \rho$ for different radii R and redshifts z as measured from the HR4 simulation.

z	R (Mpc h^{-1})	10	11	12	13	14	15
4.0	$\hat{\sigma}_\rho \simeq \hat{\sigma}_\mu$	0.18	0.17	0.16	0.15	0.14	0.13
0.7	$\hat{\sigma}_\rho$	0.51	0.47	0.44	0.41	0.39	0.37
0.7	$\hat{\sigma}_\mu$	0.46	0.43	0.41	0.39	0.37	0.35
0.0	$\hat{\sigma}_\rho$	0.74	0.68	0.63	0.59	0.55	0.52
0.0	$\hat{\sigma}_\mu$	0.61	0.58	0.55	0.52	0.50	0.48

3.3 Parametrizing the covariance matrix

In order to determine the decay-rate function, the joint PDFs and hence also the bias functions, one needs to compute the covariance matrix between initial densities in spheres of radii R_i and R_j as defined in equation (8). For the sake of simplicity, we choose here to parametrize this covariance matrix in analogy with a power-law initial spectrum with spectral index $n = n(R_p)$ at a pivot scale R_p by

$$\sigma^2(R_i, R_i) = \sigma^2(R_p) \left(\frac{R_i}{R_p} \right)^{-n(R_p)-3}, \qquad (21a)$$

$$\sigma^2(R_i, R_{j>i}) = \sigma^2(R_p) \, \mathcal{G}\left(\frac{R_i}{R_p}, \frac{R_j}{R_p}, n(R_p) \right), \qquad (21b)$$

where

$$\mathcal{G}(x, y, n) = \frac{\int d^3k \, k^n W_{3D}(kx) W_{3D}(ky)}{\int d^3k \, k^n W_{3D}(kR_p) W_{3D}(kR_p)}$$

$$= \frac{(x+y)^\alpha \left(x^2 + y^2 - \alpha xy \right) - (y-x)^\alpha \left(x^2 + y^2 + \alpha xy \right)}{2^\alpha (n+1) x^3 y^3},$$

with $\alpha = 1 - n$. The key parameter in the prediction of the PDF is the value of the variance at the pivot scale R_p that we measure in the simulation and use as an input to our theoretical model. We report the results for the measured variance of both the density ρ and the log-density $\mu = \log \rho$ in Table 1.

3.4 The density PDF in concentric spheres

For the predictions of the one- and two-cell PDF of the density in spheres, we use equation (12) specialized to an appropriate logarithmic mapping that provides a wide range of applicability for the saddle-point approximation. Note that the functional forms were presented in Uhlemann et al. (2016) and compared to measurements from a 500 Mpc h^{-1} GADGET2 simulation (Springel 2005) sampled with 1024^3 particles. Here, we confront them with the significantly more accurate measurements from the HR4 simulation at redshift $z = 0.7$; the results for $z = 0$ are shown in Appendix D.

3.4.1 One-cell density PDF

For the one-cell PDF, the appropriate mapping leading to an accurate density PDF is simply the logarithm of the density,

$$\mu = \log \rho . \qquad (22)$$

The density PDF is then obtained as

$$\mathcal{P}_R(\rho) = \sqrt{\frac{\Psi_R''[\rho] + \Psi_R'[\rho]/\rho}{2\pi}} \exp\left(-\Psi_R[\rho]\right). \qquad (23a)$$

The normalized PDF with the corrected mean is obtained from equation (23a) according to

$$\hat{\mathcal{P}}_R(\rho) = \mathcal{P}_R\left(\rho \cdot \frac{\langle\rho\rangle}{\langle 1\rangle} \right) \cdot \frac{\langle\rho\rangle}{\langle 1\rangle^2} . \qquad (23b)$$

In Fig. 3, we compare the saddle-point approximations of the PDF obtained from equation (23) evaluated with the help of LSSFAST (Codis et al. 2016a, see also Appendix A) to the measurements from the HR4 measurement for six different radii $R = 10, 11, \ldots,$ 15 Mpc h^{-1} at redshift $z = 0.7$. The agreement is spectacular over a wide range of densities.

3.4.2 Two-cell joint density–slope PDF

A suitable and physically motivated change of variables for the two-cell case is given by the logarithmic transform of the sum and difference of mass,

$$\mu_1 = \log\left(r^3 \rho_2 + \rho_1 \right), \quad \mu_2 = \log\left(r^3 \rho_2 - \rho_1 \right), \qquad (24)$$
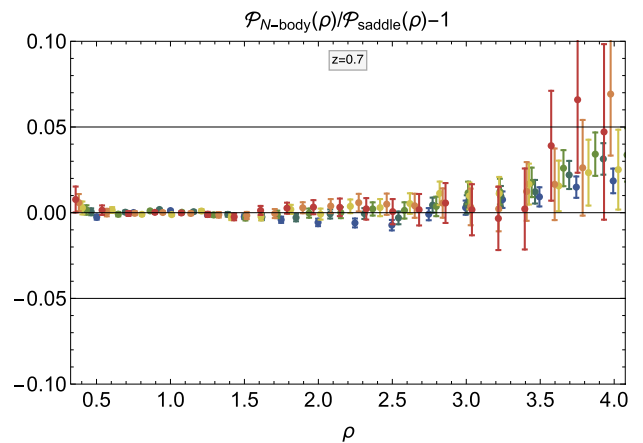

Figure 3. The density PDF $\mathcal{P}(\rho)$ (left-hand panel) predicted from the saddle-point approximation equation (23) using LSSFAST (solid lines) for redshift $z = 0.7$ and radii $R = 10, 11, \ldots, 15$ Mpc h^{-1} (from blue to red) with variances as given in Table 1, in comparison to the measurement from HR4 (data points with error bars) and the corresponding residuals (right-hand panel). See also Fig. D3 for lower redshift PDFs. For this figure, we used $\nu = 1.59$ instead of $\nu = 21/13 \simeq 1.61$ because it leads to smaller residuals (see Appendix D).
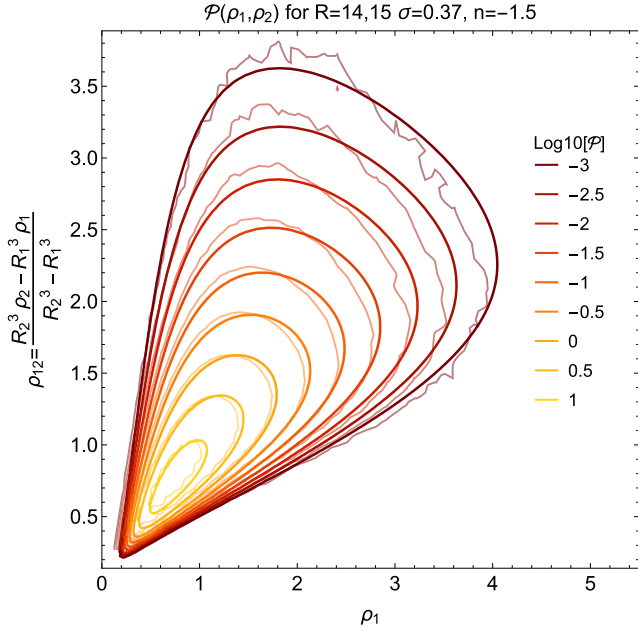
Figure 4. The joint density–slope PDF $\mathcal{P}(\rho_1, \rho_2)$ predicted from the saddle-point approximation equation (25) (thick lines) as a function of the central density ρ_1 and the shell density ρ_{12} for radii $R_{1,2} = 14$, 15 Mpc h^{-1} and redshift $z = 0.7$ where $\sigma_\mu = 0.37$ in comparison to the measurement from HR4 (thin wiggly lines). The agreement is quite good and demonstrates the wide dynamical range of this simulation. See also Fig. D3 for the same PDF at redshift $z = 0$.

where the relative shell thickness is $r = R_2/R_1$ and mass conservation ensures μ_2 to be real. The PDF $\mathcal{P}(\rho_1, \rho_2)$ can then be approximated via equation (12), which can explicitly be rewritten as

$$\mathcal{P}_{R_1, R_2}(\rho_1, \rho_2) = \frac{\exp\left[-\Psi_{R_1, R_2}\right]}{2\pi} \sqrt{p_{R_1, R_2}(\rho_1, \rho_2)}, \qquad (25a)$$

with

$$
\begin{aligned}
p_{R_1, R_2}(\rho_1, \rho_2) &= \det\left[\frac{\partial^2 \Psi_{R_1, R_2}}{\partial \mu_i \partial \mu_j}\right] \left(\det\left[\frac{\partial \mu_i}{\partial \rho_j}\right]\right)^2 \\
&= \left(\frac{1}{2r^3}\Psi_{,22} + \Psi_{,12} + \frac{r^3}{2}\Psi_{,11} + \frac{\Psi_{,2} + r^3\Psi_{,1}}{r^3\rho_2 + \rho_1}\right) \\
&\quad \times \left(\frac{1}{2r^3}\Psi_{,22} - \Psi_{,12} + \frac{r^3}{2}\Psi_{,11} + \frac{\Psi_{,2} - r^3\Psi_{,1}}{r^3\rho_2 - \rho_1}\right) \\
&\quad - \left(\frac{\Psi_{,22}}{2r^3} - \frac{r^3\Psi_{,11}}{2}\right)^2, \qquad (25b)
\end{aligned}
$$

with $\Psi_{,1}$ and $\Psi_{,2}$ denoting partial derivatives of the decay-rate function Ψ with regard to ρ_1 and ρ_2, respectively. Analogously to the one-cell case, one still has to enforce the mean and normalization for the saddle-point PDF obtained from equation (25) following the procedure described in equations (12c) and (12d). In practice, it is often useful to express the joint density PDF $\mathcal{P}(\rho_1, \rho_2)$ not in terms of the two densities but rather as a function of the inner density $\rho = \rho_1$ and the slope $s = (\rho_2 - \rho_1)/(R_2/R_1 - 1)$ or the density in the outer shell $\rho_{12} = (R_2^3\rho_2 - R_1^3\rho_1)/(R_2^3 - R_1^3)$. The result is displayed in Fig. 4 and compared to the measurements. Overall, we observe a very good agreement; it is however not as good as for the one-cell case as we did not compute it with the exact linear power spectrum but assumed a power-law power spectrum through the parametrization (21). It was shown in Bernardeau et al. (2014)

that taking into account the running of the spectral index can further improve the result.

3.5 The bias functions at large separation

To predict the bias functions for densities in spheres, we will use equation (16) with the variance of the log-density from Table 1 as input together with the normalization procedure described in equation (18). While the functional form of the density bias was introduced already in Bernardeau (1996) for the saddle-point approximation applied to the density, we provide improved analytical predictions based on the log-density mapping together with a normalization scheme that are as good as results from a numerical integration presented in Codis et al. (2016b). Furthermore, we provide new predictions for the joint density bias $b(\rho_1, \rho_2)$ measured for densities ρ_1 and ρ_2 in two concentric spheres and the derived constrained biases $b(\rho_1|\rho_2)$ given the density environment.

3.5.1 One-cell density bias

The density bias describes the cross-correlation of spherical cells given one spherically averaged density at separation r_e,

$$1 + b(\rho)\xi(r_e) = \langle \rho' | \rho ; r_e \rangle = \frac{\int_0^\infty \mathrm{d}\rho' \mathcal{P}(\rho, \rho' ; r_e)\rho'}{\mathcal{P}(\rho)} . \qquad (26)$$

Physically, the density bias describes the mean of the density found in a sphere of radius R given that the density in a sphere of the same radius at separation r_e is ρ.

The density bias can be straightforwardly obtained from the saddle-point approximation for the log-density $\mu = \log\rho$ by evaluating equation (16) for a given ρ in a cell of radius R as

$$b_R(\rho) = \frac{\nu(1 - \rho^{-1/\nu})}{\sigma_\mu^2(R\rho^{1/3})} \simeq \frac{\nu(1 - \rho^{-1/\nu})\rho^{\frac{3+n(R)}{3}}}{\sigma_\mu^2(R)} , \qquad (27a)$$

and then normalizing according to

$$\hat{b}_R(\rho) = \frac{b_R(\rho) - \langle b_R(\rho) \rangle}{\langle \rho b_R(\rho) \rangle - \langle b_R(\rho) \rangle} , \qquad (27b)$$

where the averages denoted by $\langle \cdot \rangle$ are computed as integrals with the one-cell saddle-point PDF equation (23). The result is plotted in Fig. 5 for redshift $z = 0.7$ and radii $R = 10$, 12, 14 Mpc h^{-1} with variances as indicated in the legend. In Fig. D6 in Appendix D, we also show redshift $z = 0$. One can see that unbiased results are obtained for densities close to the mean density, but note that the bias at the mean density is not exactly zero, as explained in Appendix C on the basis of perturbation theory. The amplitude of the bias grows with the deviation from the background density and is, as expected, positive for overdense and negative for underdense regions. We can clearly see that the saddle-point bias functions predict a deviation from the linear growth of Kaiser bias from equation (14) and are in excellent agreement with the measurements; the normalization procedure (18) correctly captures the finite value of the bias at mean density. In Fig. 6, we show that the prediction of equation (27), which relies on large-deviation statistics and spherical collapse, extends the classical Kaiser bias result from equation (14) valid in the Gaussian and hence linear regime towards the non-Gaussian, mildly non-linear regime. Hence, we obtain corrections to the linear growth of the bias with density contrast: we observe that the bias for underdense regions is significantly enhanced because in the non-Gaussian regime the sharp
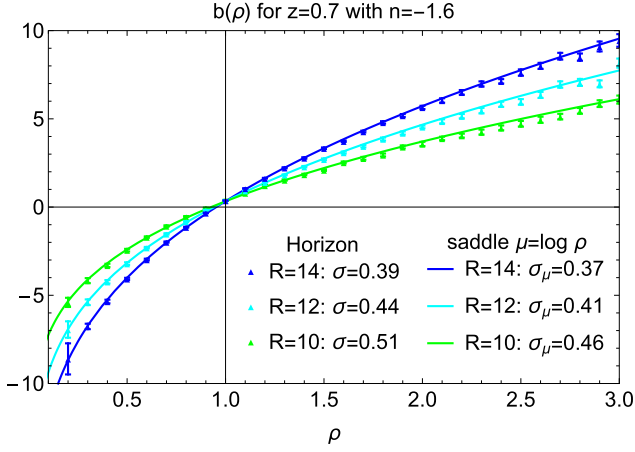
Figure 5. The normalized density bias function $\hat{b}_R(\rho)$ predicted from the saddle-point approximation given by equation (27) for the log-density $\mu = \log \rho$ for $n_s = -1.6$ and different values of the variance. This prediction is compared to the HR4 measurements at redshift $z = 0.7$ for different radii R in Mpc h^{-1} and hence variances as indicated in the legend. See also Fig. D5 for redshift $z = 0$.
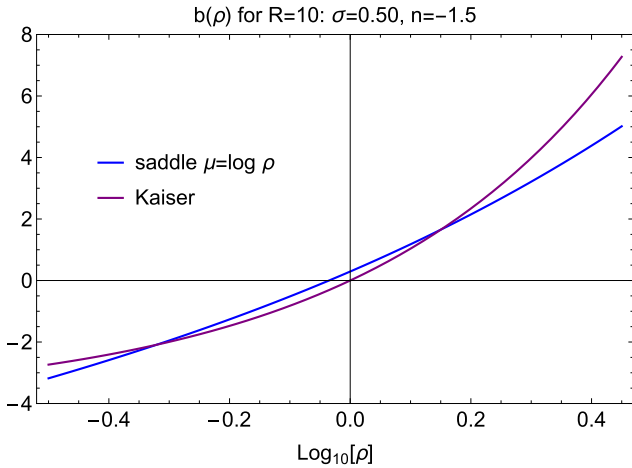

Figure 6. Density bias $b(\rho)$ from equation (27) (blue line) predicted from the saddle-point approximation for $\sigma = 0.50$ and spectral index $n_s = -1.5$ and compared with the Kaiser bias from equation (15) (purple line) valid only in the Gaussian regime but extrapolated here into the non-Gaussian regime.

cutoff at zero density becomes apparent that suppresses the rare-event tail further and disfavours very small densities. In turn, the bias for overdense regions is suppressed because the gravitational evolution bends the rare-event tail upwards and therefore favours large overdensities.

3.5.2 Joint density–slope bias

The two-cell bias is given by

$$1 + b(\rho_1, \rho_2)\xi(r_e) = \langle \rho' | (\rho_1, \rho_2); r_e \rangle$$
$$= \frac{\int_0^\infty d\rho' \, \mathcal{P}(\{\rho_1, \rho_2\}, \rho'; r_e)\rho'}{\mathcal{P}(\rho_1, \rho_2)}, \qquad (28)$$
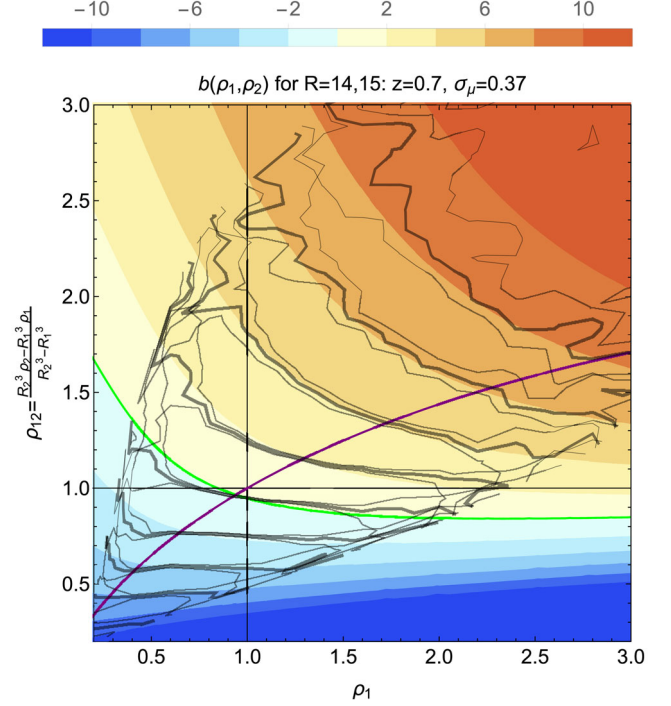

Figure 7. A contour plot of the joint density and slope bias function $b(\rho_1, \rho_2)$ predicted from the saddle-point approximation equation (29) with normalization from equation (18) as a function of the central density ρ_1 and the shell density ρ_{12} for $R_1 = 14$ Mpc h^{-1} and $R_2 = 15$ Mpc h^{-1} at redshift $z = 0.7$ where $\sigma_\mu = 0.37$ (corresponding to $\sigma_\rho = 0.39$) in comparison to the measurements from HR4 (mean as thick black lines, and mean $\pm$ error on the mean as thin black lines). Shown is also the stationary line $\rho_2(\rho_1)$ (purple line) along which the joint bias function has to be evaluated to obtain the density bias $b(\rho_1) = b(\rho_1, \rho_2(\rho_1))$. The green line corresponds to zero bias.

where $\mathcal{P}(\{\rho_1, \rho_2\}, \rho'; r_e)$ is a marginal of the two-cell PDF

$$\mathcal{P}(\{\rho_1, \rho_2\}, \rho'; r_e) = \int_0^\infty d\rho'_2 \, \mathcal{P}(\{\rho_1, \rho_2\}, \{\rho', \rho'_2\}; r_e).$$

The joint density and slope bias describes the mean of the density found in a sphere of radius $R = R_1$ given that at a distance r_e, the densities in spheres of radii R_1 and R_2 are respectively $(\hat{\rho}_1, \hat{\rho}_2)$. The result, before normalization, is straightforwardly obtained from the general formula (16),

$$b_{R_1, R_2}(\rho_1, \rho_2) = \nu \sum_{i,j=1}^{2} \Xi_{ij} \left(R_i \rho_i^{1/3}, R_j \rho_j^{1/3} \right) (1 - \rho_i^{-1/\nu}), \qquad (29)$$

where the covariance matrix is given by equation (8) and can be parametrized by equation (21) for a power-law initial spectrum. We can again use the variables describing the inner density ρ_1 together with the slope $s = (\rho_2 - \rho_1)/(R_2/R_1 - 1)$ or the density in the outer shell $\rho_{12} = (R_2^3 \rho_2 - R_1^3 \rho_1)/(R_2^3 - R_1^3)$. The result, normalized according to equation (18), is shown for radii $R_{1,2} = 14, 15$ Mpc h^{-1} at redshift $z = 0.7$ in Fig. 7 (and at redshift $z = 0$ in Fig. D6). Besides the general trend that the bias increases with increasing over- or underdensity, one can see that unbiased results are obtained along the green line for which either both densities are close to the background density or the over- or underdensity of the central density is roughly counterbalanced by an under- or overdense shell, respectively. Again we find a deviation from the linear growth predicted

by the linear bias described in equation (15) and good agreement with the measurements.

Consistency check. Note that by decimation of variables we can obtain the one-cell density bias $b(\rho_1)$ from the joint two-cell density bias $b(\rho_1, \rho_2)$ by evaluating it along the stationary line $\rho_2^{\text{stat}}(\rho_1)$ of the decay-rate function (shown as the purple line in Fig. 7) as

$$0 = \frac{\partial \Psi}{\partial \rho_2}\Big|_{\rho_2 = \rho_2^{\text{stat}}(\rho_1)} \Rightarrow b(\rho_1) = b(\rho_1, \rho_2^{\text{stat}}(\rho_1)), \quad (30)$$

which indeed gives back the density bias shown in Fig. 5.

Constrained bias given environment. Given that we have the full knowledge of the two-cell bias function, we can now determine two constrained quantities: the density bias in a positive or negative slope environment $b(\rho_1|\rho_2 \gtrless \rho_1)$ as well as the bias in an over- or underdense shell $b(\rho_1|\rho_{12} \gtrless 1)$.

3.5.3 Density bias in a positive or negative slope environment

The constrained bias for the density ρ given a positive or negative slope s can be obtained from

$$(b \cdot \mathcal{P})(\rho_1|\rho_2 \gtrless \rho_1) = \int_0^\infty \mathrm{d}\rho_2 \, (b \cdot \mathcal{P})(\rho_1, \rho_2)\theta(\pm(\rho_2 - \rho_1)), \quad (31a)$$

where θ is the Heaviside step function. Note that apart from normalization this equation resembles the definition of the constrained PDF when $b = 1$ and we have that

$$b(\rho_1) = \frac{(b \cdot \mathcal{P})(\rho_1|\rho_2 < \rho_1)}{\mathcal{P}(\rho_1)} + \frac{(b \cdot \mathcal{P})(\rho_1|\rho_2 > \rho_1)}{\mathcal{P}(\rho_1)}, \quad (31b)$$

where we define the constrained bias as

$$b(\rho_1|\rho_2 \gtrless \rho_1) = \frac{(b \cdot \mathcal{P})(\rho_1|\rho_2 \gtrless \rho_1)}{\mathcal{P}(\rho_1|\rho_2 \gtrless \rho_1)}. \quad (31c)$$

Hence, we can easily compare the constrained bias function to its unconstrained analogue. This is done in Fig. 8. A positive slope increases the bias for all densities with a strength growing with the central density, while a negative slope has the opposite effect. Because a mean central density with a negative or positive slope will appear as overall under- or overdense, respectively, the value of the bias at mean density and the point of vanishing bias appear shifted. An intuition about this shift can be gained from a peak-background split argument by computing the mean density given positive or negative slope that yields $\langle \rho_1|\rho_2 > \rho_1 \rangle = 1.15$ and $\langle \rho_1|\rho_2 < \rho_1 \rangle = 0.85$, respectively. The constrained density bias given slope provides us with a proxy for peaks in the spirit of BBKS, which correspond to overdensities with negative slope (peaks) and underdensities with positive slope (voids), respectively. Those configurations also give the asymptotes of the density bias for extreme densities, because large strongly under- or overdense regions will mostly have positive or negative slopes, respectively, which causes one contribution from equation (31b) to dominate in the regime of extreme densities. Note that, while BBKS determine peaks in the initial Gaussian field and then apply collapse criteria based on spherical collapse, we here use spherical collapse to predict the final non-Gaussian statistics of densities in spheres and use special configurations to get a proxy for peaks in the final field. It is worth nothing that our formalism describes large peaks of the density field in a similar (but not equivalent) fashion as peak bias studies do following the definition introduced in BBKS (see e.g Desjacques et al. 2010). Those two different approaches give a complementary insight on
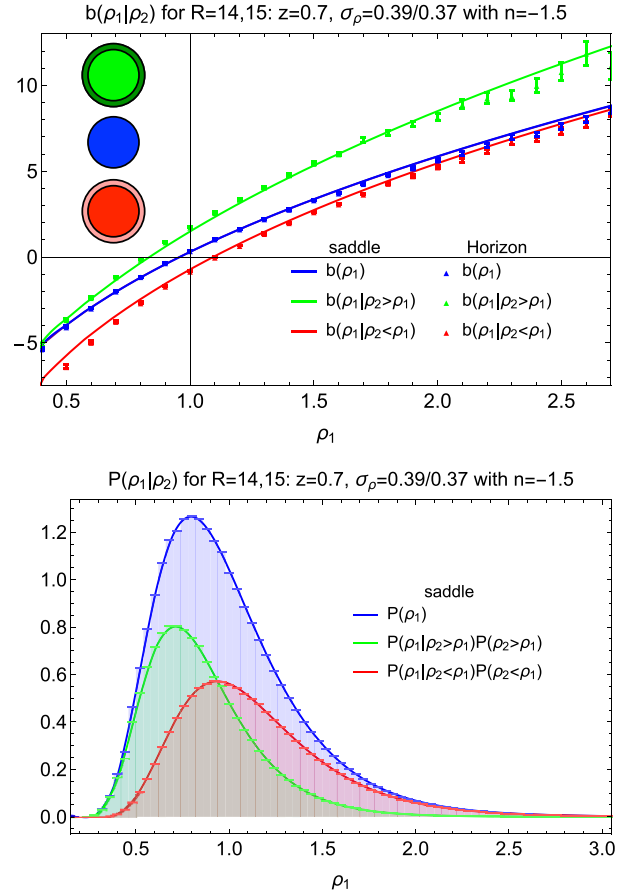

Figure 8. The constrained density bias function for negative slopes $b(\rho_1|\rho_2 < \rho_1)$ (red line) and positive slopes $b(\rho_1|\rho_2 > \rho_1)$ (green line) compared to the unconstrained density bias $b(\rho)$ (blue line) that agrees with the density bias from the one-cell saddle-point approximation $b_1(\rho)$ (purple line) shown in Fig. 5. The coloured spheres, where darker colour indicates higher density, sketched in the inset illustrate the different cases considered for the correspondingly coloured lines in the plot. All results are obtained from the saddle-point approximation for the log-density for radii $R_1 = 14$ Mpc h^{-1} and $R_2 = 15$ Mpc h^{-1} at redshift $z = 0.7$ with variance $\sigma_\mu = 0.37$. Once again the agreement is excellent.

halo biasing, either by focusing on overdense regions with negative slopes in the mildly non-linear regime (this work) or on the maxima of the Gaussian Lagrangian density field (BBKS).

3.5.4 Density bias in an over- or underdense shell

The constrained bias for the central density ρ given an over- or underdense shell with density $\rho_{12} = (R_2^3\rho_2 - R_1^3\rho_1)/(R_2^3 - R_1^3)$ can be obtained from

$$(b \cdot \mathcal{P})(\rho_1|\rho_{12} \gtrless 1) = \int_0^\infty \mathrm{d}\rho_{12} \, (b \cdot \mathcal{P})(\rho_1, \rho_{12})\theta(\pm(\rho_{12} - 1)), \quad (32a)$$

such that

$$b(\rho_1) = \frac{(b \cdot \mathcal{P})(\rho_1|\rho_{12} < 1)}{\mathcal{P}(\rho_1)} + \frac{(b \cdot \mathcal{P})(\rho_1|\rho_{12} > 1)}{\mathcal{P}(\rho_1)}, \quad (32b)$$

where we define the effective constrained bias to be

$$b(\rho_1|\rho_{12} \gtrless 1) = \frac{(b \cdot \mathcal{P})(\rho_1|\rho_{12} \gtrless 1)}{\mathcal{P}(\rho_1|\rho_{12} \gtrless 1)}. \quad (32c)$$
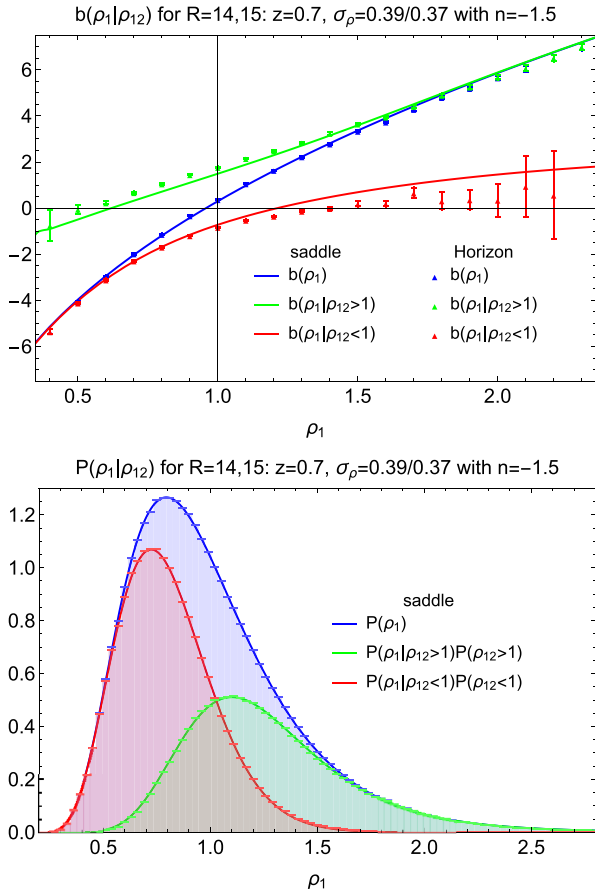
Figure 9. The constrained effective density bias function in underdense shells $b(\rho_1|\rho_{12} < 1)$ (red line) and overdense shells $b(\rho_1|\rho_{12} > 1)$ (green line) compared to the unconstrained density bias $b(\rho_1)$ (blue line). All results are obtained from the saddle-point approximation for log-density with radii $R_1 = 14$ Mpc h^{-1} and $R_2 = 15$ Mpc h^{-1} at redshift $z = 0.7$ with the variance $\sigma_\mu = 0.37$. The agreement is quite good so long as the PDF is significantly non-zero.

The results are shown in Fig. 9 where we see that the constrained density bias given an underdense- or overdense environment gives the asymptote of the density bias for small and large densities, respectively. This is due to the fact that large, strongly under- or overdense regions will mostly have under- or overdense environment, respectively, which causes one contribution from equation (32b) to dominate in the regime of extreme densities. The interesting regime is where the constrained biases deviate from the averaged bias that is when we find an overdensity residing inside an underdensity (red line in the half where $\rho_1 > 1$) or vice versa an underdensity residing inside an overdensity (green line in the half where $\rho_1 < 1$). In this specific configuration, over- and underdensities are not only peaks, but even isolated, such that we can think of them as voids surrounded by walls or clusters surrounded by voids. Because a mean central density with a surrounding under- or overdense shell will appear as overall under- or overdense, respectively, the value of the bias at mean density and the point of vanishing bias appear shifted. An intuition about this shift can be gained once again from a peak-background split argument by computing the mean density given an over- or underdense shell that yields $\langle \rho_1|\rho_{12} > 1 \rangle = 1.3$ and $\langle \rho_1|\rho_{12} < 1 \rangle = 0.8$, respectively. Besides that, the bias of overdensities is reduced for isolated overdensities because of the

'screening' effect of the surrounding underdense shell and vice versa for isolated underdensities.

3.6 The two-point correlation function of densities in spheres

The two-point correlation function of densities in spheres was introduced in equation (2b) to relate the joint PDF of densities at large separation to individual density PDFs. After having obtained analytical predictions for the density bias and constrained density bias, we can use them to predict the modulation for the correlation function $\xi(r_e)$ that is introduced by those bias functions. In Fig. 10, we show the modulation function $b(\rho)b(\rho') = \xi_\bigcirc(\rho, \rho'; r_e)/\xi(r_e)$ computed from the unconstrained density bias and constrained density bias given slope as was shown in Fig. 8. As mentioned before, the constrained density bias given a negative or positive slope can be viewed as giving a proxy for peaks when we have a negative slope around an overdensity (positive peak) or positive slope around an underdensity (negative peak), respectively. Hence, the different bias modulations we show are the ones for the auto- and cross-correlations of masses (unconstrained densities) and peaks (densities with slopes), in the spirit of Regos & Szalay (1995) and Baldauf et al. (2016), but with the added value of capturing the quasi-linear regime of structure formation.

The upper panel shows the auto-correlation for positive peaks (negative slopes), mass and negative peaks (positive slopes). The mass correlation function in the upper-middle panel shows that over- and underdensities among themselves are positively correlated and more strongly clustered than spheres of average density (lower-left and upper-right part of the plot) while over- and underdense spheres are negatively correlated with each other (lower-right and upper-left part). The point of zero bias, when compared to the unconstrained case, is shifted into the quadrant that corresponds to over- or underdensities if positive or negative peaks are involved that also points to the interesting region of the plots where the peak correlation differs from the average correlation. The lower panel shows the cross-correlations between positive peaks (negative slopes), mass and negative peaks (positive slopes).

Overall, the ab initio analytic bias functions for the two-point correlation of density in spheres, equations (27) and (29), have been shown to be in very good agreement with the HR4 simulation. This is remarkable, given that these functions are very simple explicit algebraic functions of the underlying linear power spectrum via equation (8) or (21). It also demonstrates that modern simulations capture very accurately the one- and two-point statistics of non-linear gravitational clustering.

4 DARK MATTER CORRELATION ML ESTIMATOR

Equation (2a) allows us to *analytically* model the statistics of the cosmic density field in two locations of space. This model only depends on two parameters: the variance of the density field measured at present time, $\sigma^2(R)$, and the value of the two-point dark matter correlation function, $\xi(r_e)$, at the separation. Therefore, following the ideas developed in Codis et al. (2016a) for the estimation of the variance $\sigma(R)$, one can build a maximum likelihood (ML) estimator for the two-point correlation $\xi(r_e)$ that should perform better than the sample estimator as time grows and non-Gaussianities arise.

Let us focus here on the two-point density statistics at one scale only for which equation (2a) becomes

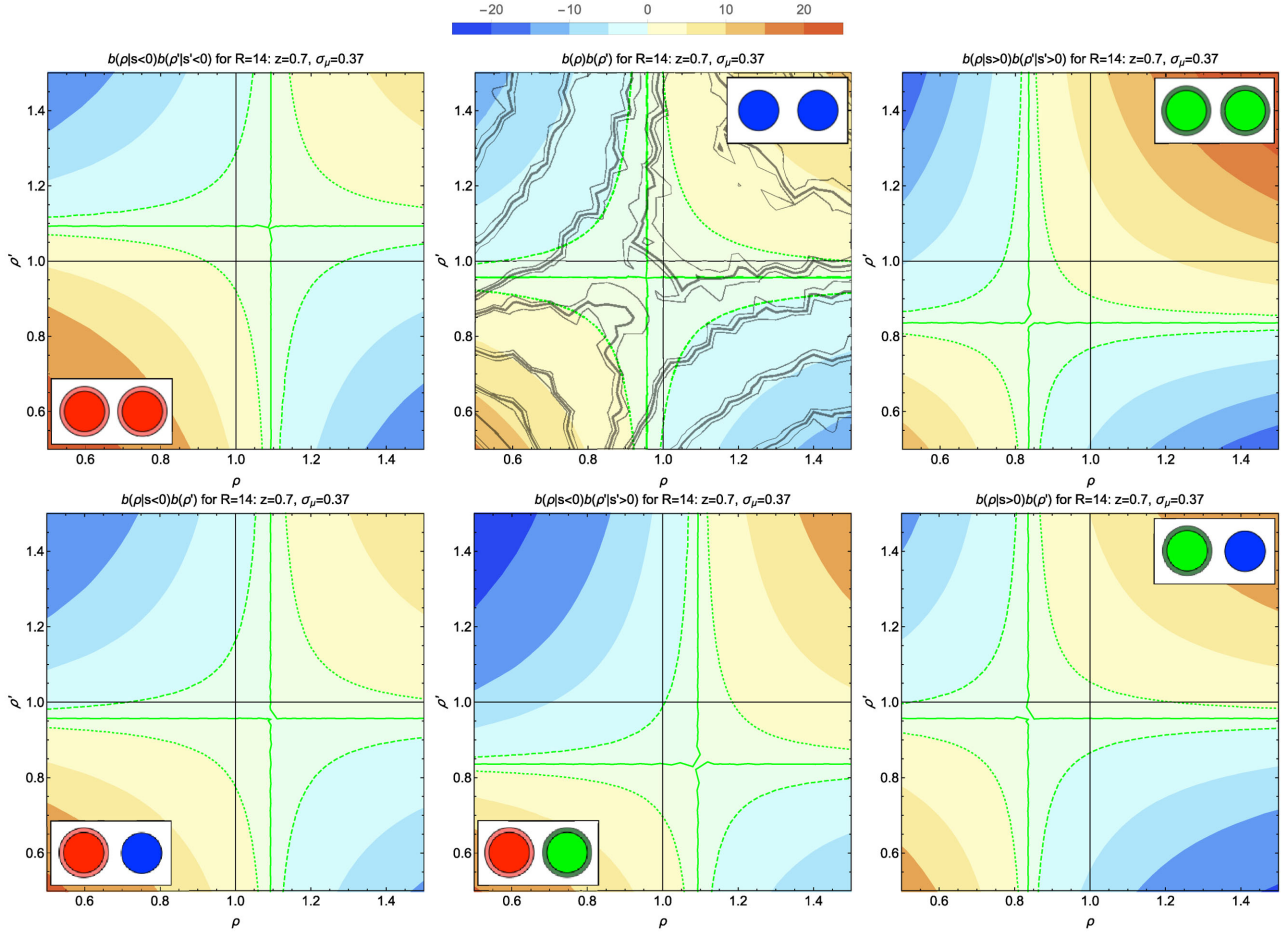$$\mathcal{P}(\rho, \rho') = \mathcal{P}(\rho)\mathcal{P}(\rho')(1 + \xi(r_e)b(\rho)b(\rho')), \qquad (33)$$

Figure 10. The bias modulation of the two-point sphere correlation function $b(\rho)b(\rho') = \xi_\bigcirc(\rho, \rho'; r_e)/\xi(r_e)$ for radii $R = R' = 14$ Mpc h^{-1} at redshift $z = 0.7$ with the variance $\sigma_\mu = 0.37$ computed from the bias and constrained bias in an over- or underdense mass shell shown in Fig. 8. Besides equi-bias contours, we show the line of zero bias $b = 0$ (green line) and unity bias $b = \pm 1$ (green dotted and dashed line) and sketch the configurations in the insets. Upper panels: auto-correlations for densities with negative slope (positive peaks; left), unconstrained densities in comparison to HR4 measurements (mean as thick black lines, and mean $\pm$ error on the mean as thin black lines; middle) and densities with positive slope (negative peaks; right). Lower panels: cross-correlations between densities with negative slope and unconstrained densities (left), densities with negative slope and densities with positive slope (middle) or densities with positive slope and unconstrained densities (right). This modulation captures the expected bias clustering of peaks and voids beyond the linear regime.

where ρ' is the density at a distance r_e from ρ. In equation (33), the one-point PDFs only depend on the variance and are computed using the public code LSSFAST described in Appendix A, and the bias $b(\rho)$ is predicted via equation (16). This two-cell PDF is shown in the left-hand panel of Fig. 11 where the effect of the spatial clustering (dashed line) is compared to the case with no spatial correlation (solid line).

4.1 Fiducial experiment from HR4

Let us carry out the following experiment: consider the 252^3 spheres of radius $R = 15$ Mpc h^{-1} of the HR4 simulation at $z = 0.7$ equally spaced on a grid of resolution $\Delta R = 12.5$ Mpc h^{-1} and let us estimate the corresponding dark matter correlation function at different separations keeping the variance fixed. To do so, for each separation $r_i = 4\Delta R, \ldots, 20\Delta R$, we look for all pairs of spheres separated by r_i and compute the log-likelihood $\mathcal{L}$ for different models of the joint PDF $\mathcal{P}(\rho, \rho'; r_i)$ described by the correlation function $\xi(r_i)$

$$\mathcal{L}(\xi(r_i)) = \sum_{(p,q)} \log \mathcal{P}(\rho_p, \rho_q' | \xi(r_i)), \qquad (34)$$

where the indices (p, q) describe all pairs of spheres separated by r_i. The maximum of the likelihood can then be found together with the α sigma contours where $\mathcal{L} = \max \mathcal{L}(\xi(r_i)) - 1/2\alpha^2$. In practice, we only consider separations above $4\Delta R = 50$ Mpc h^{-1} to avoid the small region where the modelled two-point PDF is not yet in the large-separation regime (but the bias functions, which describe the *mean* density at given separation, are in such regime long before that). However, it is expected that one could restrict the analysis around the maximum of the PDF and therefore get an estimate of the dark matter correlation function even for smaller separations.

The right-hand panel of Fig. 11 shows the corresponding ML estimate for $\xi(r_e)$ as a function of the separation compared to a sample estimator with no prior on the underlying PDF

$$\hat{\xi}_A = \langle \rho_p \rho_q \rangle - 1. \qquad (35)$$

The agreement between the sample and likelihood estimators is remarkable, highlighting that the model presented here for the two-point density PDF is very good and could be used to measure more accurately the dark matter correlation function.
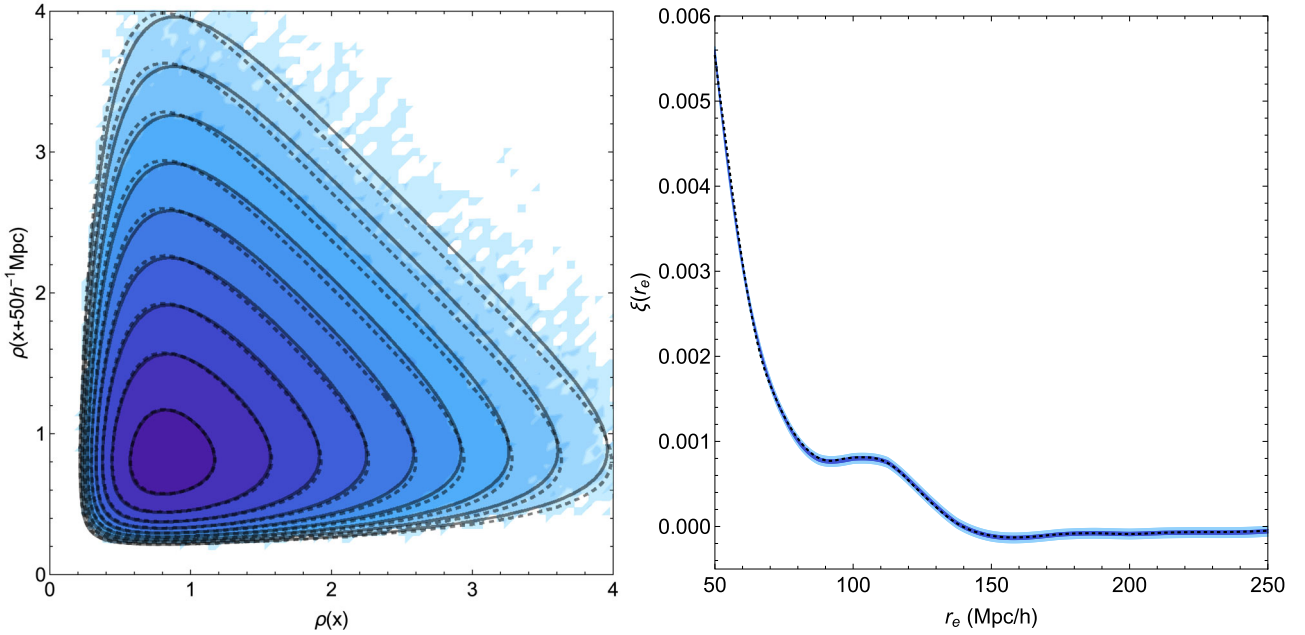
Figure 11. Left-hand panel: PDF of the densities separated by $r_e = 50$ Mpc h^{-1} in the HR4 simulation (from dark to light blue) and predicted for $\xi(r_e) = 0$ (dashed) and 0.005 58 (dotted). Contours are displayed for $\log \mathcal{P} = 1, 0.5, 0, \ldots, -4$. Right-hand panel: one- (dark blue) and three-sigma (light blue) contours for the maximum likelihood estimate of the dark matter correlation function, $\hat{\xi}_{\mathrm{ML}}$, compared with the arithmetic estimate, $\hat{\xi}_{\mathrm{A}}$ (dashed line).

4.2 Qualifying the estimator

This procedure is expected to perform better than usual sample estimators when the field becomes mildly non-linear. Indeed, in analogy with the analysis presented in Codis et al. (2016a), one can show that the scatter of the maximum likelihood estimator,

$$\hat{\xi}_{\mathrm{ML}} = \mathrm{argmax}\,\mathcal{L}(\xi), \qquad (36)$$

is much smaller than that of the sample estimator, $\hat{\xi}_{\mathrm{A}}$. To illustrate this point, for different separations r_e between 50 and 250 Mpc h^{-1}, we randomly divide the pairs of spheres of radius $R = 15$ Mpc h^{-1} into 216 subsets. For each subset, we estimate the dark matter correlation function via the sample estimator and the ML estimator. The mean and one standard deviation are shown in Fig. 12. Both are shown to be unbiased, as the mean is consistent with the correlation function measured from the full simulation, and do not depend on the separation. But, the ML estimator is shown to give a tighter measurement of the dark matter correlation function than the arithmetic estimate, the scatter being reduced by a factor of 5 in this case. The significant improvement is similar to what was described in Codis et al. (2016a) for the scatter of the variance of the density field estimated from one-point statistics of densities in spheres where a factor of 2 improvement compared to the arithmetic estimate was found. We describe in more details how it occurs in Appendix E. This method could therefore be applied successfully to real surveys provided one is able to model galaxy biasing (Feix et al., in preparation). Such likelihood estimators could also be generalized to subset regions of the top-hat filtered field where the density has a given value, which could be chosen so as to optimize the sought level of non-linearity.

Note finally that the underlying cosmological parameters enter the likelihood function $\mathcal{L}(z)$ at a given redshift, z, twice: via the bias function, $b(\rho, z)$, and via the generalized perturbation theory redshift-dependent two-point function, $\xi(r, z)$. Following Codis et al. (2016a), one could imagine in the long run building an
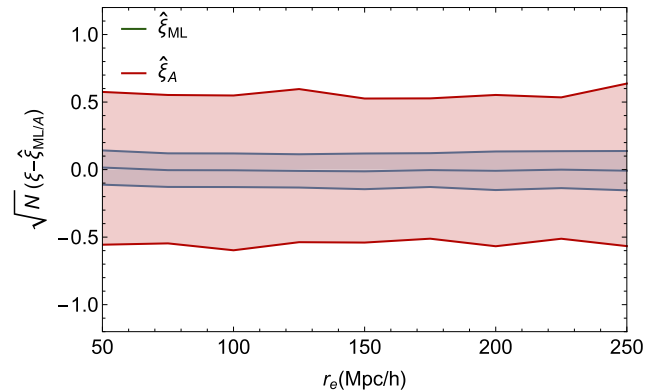

Figure 12. Estimate of the dark matter correlation function at $z = 0.7$ in 216 random subsamples of $N = 222\,264$ pairs of spheres of radius $R = 15$ Mpc h^{-1} separated by $r_e = 50, \ldots, 250$ Mpc h^{-1}. The mean and one standard deviation area are shown in blue for the maximum likelihood estimator $\hat{\xi}_{\mathrm{ML}}$ and in red for the usual sample estimator $\hat{\xi}_{\mathrm{A}}$. We display only the difference compared to the dark matter correlation function measured in the full simulation that is assumed to be the true underlying value of ξ. As expected, the ML estimator provides a much smaller scatter than the sample estimator.

optimal dark energy experiment that would measure the dark energy parameters while leveraging both the one- and two-point statistics

$$(\hat{w}_0, \hat{w}_a) = \mathop{\mathrm{argmax}}_{w_0, w_a} \sum_z \mathcal{L}(z|w_0, w_a),$$

where w_0 and w_a would parametrize the equation of state of dark energy (Glazebrook & Blake 2005).

5 CONCLUSIONS

This paper presented simple parameter-free analytic bias functions for the two-point correlation of density in spheres, in equations (27)

and (29). These bias functions generalize the so-called Kaiser bias in the mildly non-linear regime for (not so) large separation and for arbitrary contrast when considering the density smoothed with a top-hat filter (or equivalently measured in spheres). The derivation was carried out using a large-deviation principle, while relying on the spherical collapse model. A logarithmic transformation allowed for a saddle-point approximation, which was shown to be extremely accurate against the state-of-the-art HR4 *N*-body simulation throughout the range of measured densities, e.g. extending the match to the theory by a factor of 10 or more on joint PDFs, conditionals and marginals. This is both a success of the theory and an assessment of the quality of this simulation.

The conditional density-given-slope and density-given-mass biases were also presented as a quasi-linear proxy for the BBKS extremum correlation functions operating at lower redshifts. As an illustration, Fig. 10 presents the expected bias modulation of the sphere–sphere correlation function at redshift 0.7 in spheres of 14 Mpc h^{-1}.

Codis et al. (2016b) recently showed how such bias functions could be used as a mean of mitigating correlation errors when computing count-in-cell statistics on finite surveys. Conversely, based on the knowledge of the joint PDF of the density in spheres separated by r_e, we presented and implemented in Section 4 an ML estimator for the underlying top-hat smoothed dark matter density, which was shown to be unbiased and very accurate for separations above 50 Mpc h^{-1}. Its variance is up to five times smaller than that of the classical sample estimator. Hence, these analytic bias functions should be used jointly with analytic models for the two-point function from perturbation theory for cosmic parameter estimation, as they capture the biasing effect of non-linear regime of structure formation.

Let us stress in closing that the saddle-point PDFs presented in this work are not arbitrary fitting functions, but a clear prediction of the theory of gravitational clustering that allows for direct comparison with data at low redshift. These PDFs should also be compared favourably with fits to a lognormal PDF that provide a much worse match as illustrated in Fig. D1. The saddle-point approximation presented here gives, at very little extra cost, a few per cent accuracy over about four orders of magnitude in the values of the one- and two-cell PDFs and per cent accuracy on the bias functions for all densities probed by the simulation, with an explicit dependence on both cosmology, through the initial power spectrum, and the chosen theory of gravitation, through the spherical collapse model. In this paper, we ignored redshift-space distortion or galaxy biasing that will be investigated in Feix et al. (in preparation).

ACKNOWLEDGEMENTS

REFERENCES

Baldauf T., Codis S., Desjacques V., Pichon C., 2016, MNRAS, 456, 3985
Balian R., Schaeffer R., 1989, A&A, 220, 1
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15 (BBKS)
Bernardeau F., 1992, ApJ, 390, L61
Bernardeau F., 1996, A&A, 312, 11
Bernardeau F., Reimberg P., 2016, Phys. Rev. D, 94, 3520
Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1
Bernardeau F., Pichon C., Codis S., 2014, Phys. Rev. D, 90, 103519
Bernardeau F., Codis S., Pichon C., 2015, MNRAS, 449, L105
Carron J., Neyrinck M. C., 2012, ApJ, 750, 28
Carron J., Szapudi I., 2013, MNRAS, 434, 2961
Chisari N., Codis S., Laigle C., Dubois Y., Pichon C., Devriendt J., 2015, MNRAS, 454, 2736
Codis S., Pichon C., Bernardeau F., Uhlemann C., Prunet S., 2016a, MNRAS, 460, 1549
Codis S., Bernardeau F., Pichon C., 2016b, MNRAS, 460, 1598
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Dalal N., White M., Bond J. R., Shirokov A., 2008, ApJ, 687, 12
Dekel A., Rees M. J., 1987, Nature, 326, 455
Desjacques V., Crocce M., Scoccimarro R., Sheth R. K., 2010, Phys. Rev. D, 82, 103529
Dubinski J., Kim J., Park C., Humble R., 2004, New Astron., 9, 111
Gay C., Pichon C., Pogosyan D., 2012, Phys. Rev. D, 85, 023011
Glazebrook K., Blake C., 2005, ApJ, 631, 1
Guillet T., Teyssier R., Colombi S., 2010, MNRAS, 405, 525
Jenkins A., 2010, MNRAS, 403, 1859
Joachimi B. et al., 2015, Space Sci. Rev., 193, 1
Juszkiewicz R., Bouchet F. R., Colombi S., 1993, ApJ, 412, L9
Kaiser N., 1984, ApJ, 284, L9
Kaiser N., 1987, MNRAS, 227, 1
Kiessling A. et al., 2015, Space Sci. Rev., 193, 67
Kim J., Park C., L'Huillier B., Hong S. E., 2015, J. Korean Astron. Soc., 48, 213
Kirk D. et al., 2015, Space Sci. Rev., 193, 139
Laureijs R. et al., 2011, preprint (arXiv:e-prints)
Levi M. et al., 2013, preprint (arXiv:e-prints)
Neyrinck M. C., Szapudi I., Szalay A. S., 2009, ApJ, 698, L90
Neyrinck M. C., Szapudi I., McCullagh N., Szalay A., Falck B., Wang J., 2016, preprint (arXiv:e-prints)
Regos E., Szalay A. S., 1995, MNRAS, 272, 447
Schneider A., Teyssier R., 2015, J. Cosmol. Astropart. Phys., 12, 049
Sheth R. K., Tormen G., 2004, MNRAS, 350, 1385
Spergel D. et al., 2013, preprint (arXiv:1305.5422)
Springel V., 2005, MNRAS, 364, 1105
Stuart A., Ord K., 2009, Kendall's Advanced Theory of Statistics: Vol. 1: Distribution Theory. Wiley, New York
Takada M., Ellis R. S., Chiba M., 2014, PASJ, 66, R1
Taruya A., Nishimichi T., Saito S., 2010, Phys. Rev. D, 82, 063522
Taruya A., Bernardeau F., Nishimichi T., Codis S., 2012, Phys. Rev. D, 86, 103528
Tenneti A., Singh S., Mandelbaum R., Matteo T. D., Feng Y., Khandai N., 2015, MNRAS, 448, 3522
Troxel M. A., Ishak M., 2015, Phys. Rep., 558, 1
Uhlemann C., Codis S., Pichon C., Bernardeau F., Reimberg P., 2016, MNRAS, 460, 1529
Valageas P., 2002, A&A, 382, 412
van Daalen M. P., Schaye J., McCarthy I. G., Booth C. M., Dalla Vecchia C., 2014, MNRAS, 440, 2997
Velliscig M. et al., 2015, MNRAS, 454, 3328
White M., 2016, JCAP, 11, 057
White M., Padmanabhan N., 2009, MNRAS, 395, 2381
Zhan H., Knox L., 2004, ApJ, 616, L75

APPENDIX A: LSSFAST PACKAGE

The one-point density PDF and the bias functions for power-law and arbitrary power spectra are made available in the LSSFAST package distributed freely at http://cita.utoronto.ca/∼codis/LSSFast.html. Two versions of the code are presented. The simpler version, ρPDFns and ρbiasns, assumes a running index, meaning that the variance is given by

$$\sigma^2(R) = \frac{2\sigma^2(R_p)}{\left(R/R_p\right)^{3+n_s+\alpha} + \left(R/R_p\right)^{3+n_s-\alpha}}, \tag{A1}$$

where α can be non-zero to take into account the variation of the spectral index n_s. The density PDF is analytically computed from equation (23) and the bias from equation (27). This code is very efficient and runs in about 1 s on one processor for one evaluation. Note that the functions ρPDFns and ρbiasns take three arguments, ρ, σ and n_s, and have one option, α.

The second version of the code, ρPDF and ρbias, can be applied to arbitrary power spectra. In this case, the function $\sigma^2(R)$ is tabulated using equation (8). Once this tabulation is done (typically 1 min on one processor), each evaluation of the PDF and the bias takes about the same time as for the power-law case (≈ 1 s).

APPENDIX B: PDF AND BIAS FUNCTION DERIVATION

Let us present shortly the idea behind the large-deviation principle that allows us to obtain the PDF of densities in concentric cells and the generalization to the joint PDF of densities at different positions. For more details, we refer to Uhlemann et al. (2016) and Codis et al. (2016b).

B1 The PDF of density in concentric spheres

Bernardeau et al. (2014) computed the joint PDF, $\mathcal{P}(\{\rho_k\})$, of densities in concentric spheres, a highly symmetric configuration that allows one to take advantage of the spherical collapse model for gravitational dynamics. To obtain the PDF, we use the cumulant generating function of densities in concentric cells, $\varphi(\{\lambda_k\})$, defined via a Laplace transform of the density PDF $\mathcal{P}(\{\rho_k\})$

$$\varphi(\{\lambda_k\}) = \log\left[\int \Pi_k d\rho_k \exp(\Sigma_k \lambda_k \rho_k) \mathcal{P}(\{\rho_k\})\right],$$
$$= \log[\langle \exp(\Sigma_k \lambda_k \rho_k)\rangle] = \sum_{p_i=0}^{\infty} \langle \Pi_i \, \rho_i^{p_i}\rangle_c \frac{\Pi_i \lambda_i^{p_i}}{\Pi_i p_i}. \tag{B1}$$

This relationship is useful because, in the limit of zero variance, the cumulant generating function is obtained analytically from the decay-rate function $\Psi(\{\rho_k\})$ via a Legendre transformation

$$\varphi(\{\lambda_k\}) = \sum_i \lambda_i \rho_i - \Psi(\{\rho_k\}), \quad \lambda_i = \frac{\partial}{\partial \rho_i}\Psi(\{\rho_k\}), \tag{B2}$$

where the conjugate variables $\{\lambda_k\}$ are functions of the densities $\{\rho_k\}$ via the stationary condition on the decay-rate function that in turn has been obtained from the initial decay-rate function by a simple remapping according to spherical collapse as described in equation (10) (as a result of the contraction principle). The PDF of the density is then given as an inverse Laplace transform of the cumulant generating function $\varphi(\{\lambda_k\})$

$$\mathcal{P}(\{\rho_k\}) = \int \prod_k \frac{d\lambda_k}{2\pi i} \exp\left[-\sum_k \lambda_k \rho_k + \varphi(\{\lambda_k\})\right]. \tag{B3}$$

Hence, PDF can be obtained from a numerical integration in the complex plane as done in Bernardeau et al. (2014, 2015) or evaluated using a saddle-point approximation for the log-density (which has a close-to-optimal range of validity) as described in Uhlemann et al. (2016).

B2 Two-point clustering of concentric spheres

Let us now consider two sets of concentric spheres separated by a distance r_e and define the corresponding densities $\{\rho_k\} \equiv \{\rho_{1,k}\}$ and $\{\rho'_k\} \equiv \{\rho_{2,k}\}$. We are interested in the joint density PDF $\mathcal{P}(\{\rho_k\}, \{\rho'_k\}; r_e)$ that, at large separations $r_e > R_k$, can be predicted from the individual PDFs $\mathcal{P}(\{\rho_{1/2,k}\})$ and some effective bias functions $b(\{\rho_{1/2,k}\})$ according to equation (2b). The effective bias functions encode the correlations of the densities in spheres that are hence related to the two-point correlation function ξ of the underlying dark matter distribution. In analogy with the density PDF at one point, also the joint density PDF at two different points can be obtained from the corresponding cumulant generating function $\varphi(\{\lambda_k\}, \{\lambda'_k\}; r_e)$ of the joint cumulants $\langle \rho_1^{p_1} \ldots \rho_n^{p_n} \rho'^{q_1}_1 \ldots \rho'^{q_m}_m\rangle_c$ as inverse Laplace transform

$$\mathcal{P}(\{\rho_k\}, \{\rho'_k\}; r_e)$$
$$= \int \prod_k \frac{d\lambda_k}{2\pi i}\frac{d\lambda'_k}{2\pi i} \exp\left[-\Sigma_k(\lambda_k\rho_k + \lambda'_k\rho'_k) + \varphi(\{\lambda_k\}, \{\lambda'_k\}; r_e)\right]. \tag{B4}$$

In Codis et al. (2016b), it has been shown that in the large-separation limit, where the separation distance r_e is much larger than all radii R_k at the individual points, the joint cumulant generating function $\varphi(\{\lambda_k\}, \{\lambda'_k\}; r_e)$ can be derived based on the following idea: for large enough separations, the joint cumulants can be shown to be well approximated by

$$\langle \rho_1^{p_1} \ldots \rho_n^{p_n} \rho'^{q_1}_1 \ldots \rho'^{q_m}_m\rangle_c$$
$$= \frac{1}{\xi(r_e)}\langle \rho_1^{p_1} \ldots \rho_n^{p_n}\rho'_1\rangle_c\langle \rho_1\rho'^{q_1}_1 \ldots \rho'^{q_m}_m\rangle_c. \tag{B5}$$

Hence, in this limit, it is enough to know the subset of cumulants of the type $\langle \rho_1^{p_1} \ldots \rho_n^{p_n}\rho'_1\rangle_c$ to determine the generating function of joint cumulants $\varphi(\{\lambda_k\}, \{\mu_k\}; r_e)$. The generating function of the joint cumulants of this special type can be shown to be

$$\varphi_b(\{\lambda_k\}, r_e) = 1 + \xi(r_e)b_\varphi(\{\lambda_k\}), \tag{B6}$$

with the bias cumulant generating function defined as

$$b_\varphi(\{\lambda_k\}) \equiv \sum_{i=1}^n \sum_{j=1}^n \Xi_{ij}\tau_j. \tag{B7}$$

Equations (B5) and (B6) can then be used to express the joint cumulant generating function $\varphi(\{\lambda_k\}, \{\lambda'_k\}; r_e)$ in terms of the already known generating function $\varphi(\{\lambda_k\})$ of cumulants at one point via

$$\varphi(\{\lambda_k\}, \{\lambda'_k\}; r_e)$$
$$= \varphi(\{\lambda_k\}) + \varphi(\{\lambda'_k\}) + \xi(r_e)\,b_\varphi(\{\lambda_k\})\,b_\varphi(\{\lambda'_k\}). \tag{B8}$$

The bias cumulant generating function, $b_\varphi(\{\lambda_k\})$, is therefore defined as the sum of the first partial derivatives of the initial decay-rate function and hence closely related to equation (10). The bias function is then obtained from the bias cumulant generating function via

$$b(\{\rho_k\})\mathcal{P}(\{\rho_k\})$$
$$= \int \prod_k \frac{d\lambda_k}{2\pi i} b_\varphi(\{\lambda_k\}) \exp\left(-\sum_k \lambda_k\rho_k + \varphi(\{\lambda_k\})\right). \tag{B9}$$

Evaluating the integral in equation (B9) using a saddle-point approximation then gives

$$b(\{\rho_k\}) \approx b_\varphi \left(\left\{ \lambda_k = \frac{\partial \Psi(\{\rho_i\})}{\partial \rho_k} \right\} \right) = \sum_{i,j=1}^{n} \Xi_{ij} \tau_i(\rho_i) \, . \qquad \text{(B10)}$$

APPENDIX C: WEAKLY NON-GAUSSIAN BIAS

Let us revisit here the origin of the normalization shift discussed in the main text by looking at the Gaussian and weakly non-Gaussian predictions for the mean density bias.

C1 Kaiser bias

To study one-cell and two-cell bias, we first diagonalize the covariance matrix from equation (13), by transforming from τ_1', τ_1, τ_2 set of correlated variable to the following set of independent variables

$$\nu_1 = \frac{\tau_1}{\sigma_{11}}, \quad \zeta = \frac{\sigma_{11}^2 \tau_1' - \xi_{11} \tau_1}{\sigma_{11}\sqrt{\sigma_{11}^4 - \xi_{11}^2}}, \qquad \text{(C1)}$$

$$\eta = \frac{\sigma_{11}}{\sqrt{\sigma_{22}^2\sigma_{11}^2 - \sigma_{12}^4 - \alpha^2}} \left(\tau_2 - \frac{\sigma_{12}^2}{\sigma_{11}^2} \tau_1 - \frac{\alpha}{\sigma_{11}} \zeta \right), \qquad \text{(C2)}$$

which are built to be decorrelated and normalized by their variance:

$$\langle \nu_1^2 \rangle = \langle \zeta^2 \rangle = \langle \eta^2 \rangle = 1 \, , \quad \langle \nu_1 \zeta \rangle = \langle \nu_1 \eta \rangle = \langle \zeta \eta \rangle = 0$$

once α is set to $\alpha = (\xi_{12}\sigma_{11}^2 - \xi_{11}\sigma_{12}^2)/\sqrt{\sigma_{11}^4 - \xi_{11}^2}$. Thanks to the diagonalization (ν_1, ζ, η) now follow a standard normal distribution, such that it is easy to check that the density bias reads

$$b^{\mathrm{G}}(\tau) = \frac{\langle \tau_1'(\zeta, \nu_1) | \nu_1 = \tau/\sigma_{11} \rangle}{\xi_{11}} = \frac{\tau}{\sigma_{11}^2} \, , \qquad \text{(C3)}$$

which is proportional to the initial overdensity τ as expected from Kaiser (1984). The two-cell density bias also follows as

$$b^{\mathrm{G}}(\tau_1, \tau_2) = \frac{\langle \tau_1'(\zeta, \nu_1) | (\tau_1, \tau_2) \rangle}{\xi_{11}} \approx \sum_{i,j=1}^{2} \Xi_{ij} \tau_j \, , \qquad \text{(C4)}$$

if we assume that for large separations r_e we have $\xi_{12} \approx \xi_{11}$.

C2 Expected offset of density bias

Fig. 5 shows that for the non-linear density field, the one-cell bias $b(\rho)$ is non-vanishing and positive at the mean density $\rho = 1$ whereas the Gaussian result from equation (C3) predicts zero bias. Here we compute this offset using perturbative methods in the weakly non-Gaussian regime that are expected to be accurate for variances of the order of $\sigma \lesssim 0.2$.

We use a moment expansion for the two-cell distribution function. The Gaussian limit provides the kernel to define the orthogonal polynomials of the expansion (Gay, Pichon & Pogosyan 2012). In the decorrelated variables ν_1 and ζ introduced in equation (C1), these are just products of Hermite polynomials, which for the first non-Gaussian correction to the two-point distribution function give

$$\begin{aligned} P(\nu_1, \zeta) \approx &\frac{1}{2\pi} \exp\left(-\frac{\nu_1^2}{2} - \frac{\zeta^2}{2}\right) \\ &\times \left[1 + \frac{1}{6}\langle \nu_1^3 \rangle H_3(\nu_1) + \frac{1}{6}\langle \zeta^3 \rangle H_3(\zeta) \right. \\ &\left. + \frac{1}{2}\langle \nu_1 \zeta^2 \rangle H_1(\nu_1) H_2(\zeta) + \frac{1}{2}\langle \nu_1^2 \zeta \rangle H_2(\nu_1) H_1(\zeta) \right]. \end{aligned}$$

$$\text{(C5)}$$

The conditional mean that determines the one-cell bias is

$$\langle \tau_1' | \tau_1 = \tau \rangle = \frac{\int_{-\infty}^{\infty} \mathrm{d}\nu_1 \int_{-\infty}^{\infty} \mathrm{d}\zeta \; \tau_1'(\nu_1, \zeta) P(\nu_1, \zeta)\delta_D(\tau_1(\nu_1, \zeta) - \tau)}{\int_{-\infty}^{\infty} \mathrm{d}\nu_1 \int_{-\infty}^{\infty} \mathrm{d}\zeta \; P(\nu_1, \zeta)\delta_D(\tau_1(\nu_1, \zeta) - \tau)}$$

where inverting equation (C1) gives

$$\tau_1 = \nu_1 \sigma_{11}, \quad \tau_1' = \frac{\nu_1 \xi_{11} + \zeta\sqrt{\sigma_{11}^4 - \xi_{11}^2}}{\sigma_{11}} \, . \qquad \text{(C6)}$$

After some algebra, and expressing the moments of ν_1, ζ back via the moments of τ_1, τ_1', the one-cell bias at $\rho = \bar{\rho} = 1$ in the leading non-Gaussian order becomes

$$b(\rho = 1) \sim \frac{1}{2}\left(\frac{\langle \tau_1^3 \rangle}{\sigma_{11}^4} - \frac{\langle \tau_1' \tau_1^2 \rangle}{\sigma_{11}^2 \xi_{11}(r)} \right). \qquad \text{(C7)}$$

As expected, it is zero for $r = 0$. In the large-separation limit, the cubic moments can be computed using perturbation theory (Bernardeau et al. 2002) and give

$$\frac{\langle \tau_1^3 \rangle}{\sigma_{11}^4} = S_3 \approx \frac{34}{7} - (n+3), \quad \frac{\langle \tau_1' \tau_1^2 \rangle}{\sigma_{11}^2 \xi_{11}} = C_{21} \approx \frac{68}{21} - \frac{1}{3}(n+3)$$

such that the bias offset evaluates to

$$b(1) \approx \frac{1}{2}(S_3 - C_{21}) = -\frac{4}{21} - \frac{n}{3} \overset{n=-1.6}{\approx} 0.34 \, , \qquad \text{(C8)}$$

a value fully consistent with the measured one.

APPENDIX D: REDSHIFT ZERO MATCH

While we focused our comparison between the theoretical predictions and the HR4 simulation in Section 3 to redshift $z = 0.7$, we here provide results from high redshift $z = 4$ until today $z = 0$ to outline the reach of our formalism.

D1 Saddle-point versus lognormal PDF

The one-cell saddle-point PDF presented in equation (23) obtained from a log-density mapping has to be contrasted to an ad hoc lognormal PDF

$$\mathcal{P}_{\mathrm{lognorm}}(\rho, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{\rho} \exp\left[-\frac{(\log\rho + \sigma^2/2)^2}{2\sigma^2} \right] \qquad \text{(D1)}$$

with a best fit for the variance σ that mismatches the simulated PDFs at the 10 per cent level or more in its tail, as shown in Fig. D1. This is to be compared with the excellent match seen in Fig. 3. Note that, while doing a joint fit of the mean (which is otherwise assumed to be $\sigma^2/2$) and variance does improve the fit around the mean density, it worsens the mismatch in the tail.

D2 The density PDF in concentric spheres

In Figs D2 and D3, we show the one-cell PDFs for redshifts $z = 0$ and 4 comparing the saddle-point approximation computed using LSSFAST with the measurements from the HR4 simulation. Furthermore, we show the two-cell PDF for redshift $z = 0$ in Fig. D4. Note that we have chosen to use $\nu = 1.59$ instead of $21/13 \approx 1.61$ at low redshift as the residuals were smaller in this case. This suggests that in order to get per cent precision on the PDF at low redshift, one probably has to account for next-to-leading-order correction to the skewness (that a slightly lower value of ν seems to reproduce). This is clearly seen at $z = 0$ when the numerical integration of the
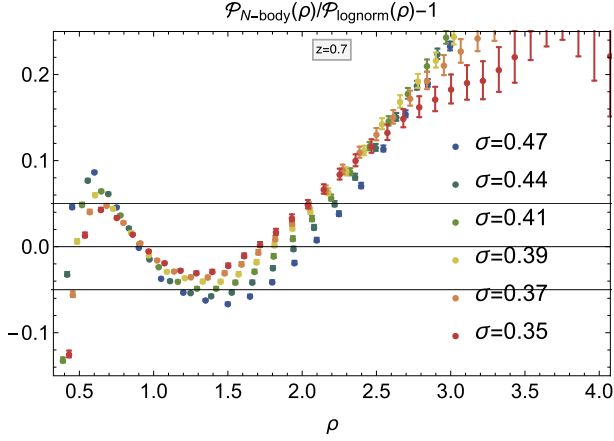
Figure D1. The residuals of the best-fitting lognormal PDF $\mathcal{P}(\rho)$ for redshift $z = 0.7$ and radii $R = 10, 11, \ldots, 15$ Mpc h^{-1} (from blue to red) in comparison to the measurement from HR4 (with error bars). This is to be contrasted to the quasi-perfect match of the saddle PDF presented in Fig. 3.
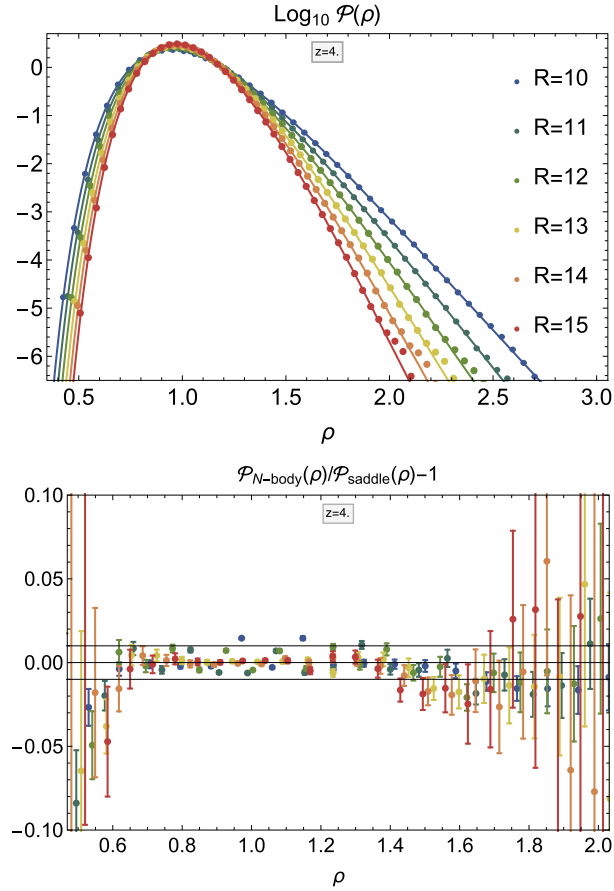

Figure D2. The density PDF $\mathcal{P}(\rho)$ predicted from equation (23) using LSSFAST (solid lines) for radii $R = 10, 11, \ldots, 15$ Mpc h^{-1} (from blue to red) at redshift $z = 4$ with $\nu = 21/13$ in comparison to the HR4 measurement (data points) and the corresponding residuals.

inverse Laplace transform is carried out and shows residuals proportional to the typical third-order Hermite polynomial, characteristic of the skewness. Adding higher order corrections to the skewness (by means of perturbation theory) is left for future work.
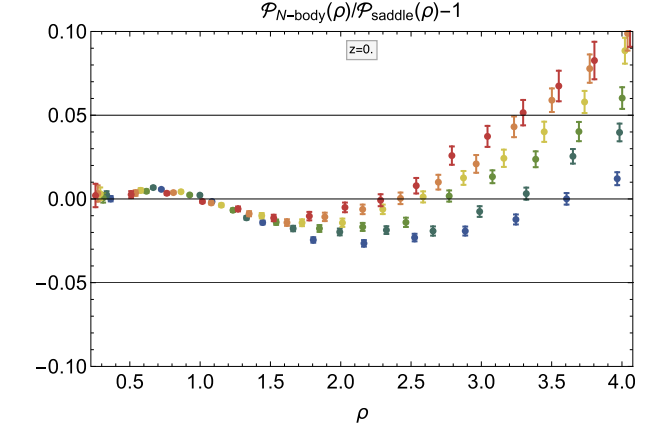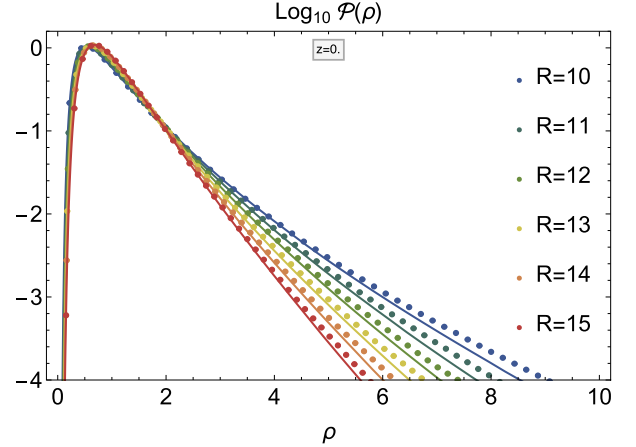

Figure D3. The density PDF $\mathcal{P}(\rho)$ predicted from equation (23) using LSSFAST (solid lines) for radii $R = 10, 11, \ldots, 15$ Mpc h^{-1} (from blue to red) at redshift $z = 0$ with a slightly adjusted $\nu = 1.59$ in comparison to the HR4 measurement (data points) and the corresponding residuals.
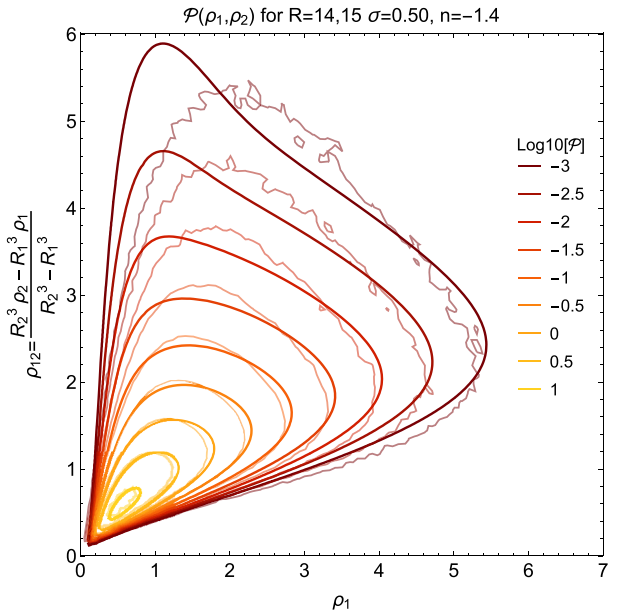

Figure D4. The joint density–slope PDF $\mathcal{P}(\rho_1, \rho_2)$ predicted from equation (25) for radii $R_{1,2} = 14, 15$ Mpc h^{-1} at redshift $z = 0$ with $\nu = 21/13$ in comparison to the measurement from HR4 (thin wiggly lines).
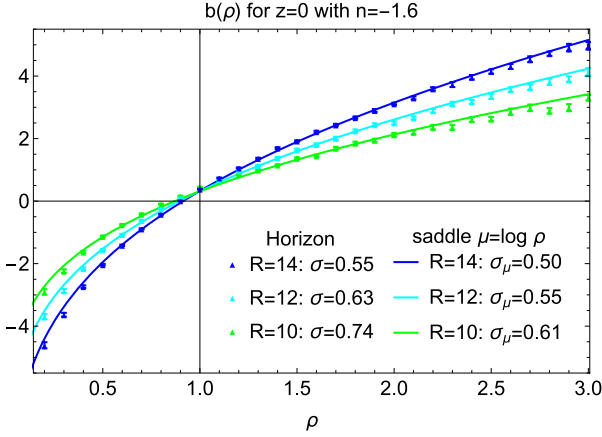
Figure D5. The density bias function $b(\rho)$ predicted from the saddle-point approximation for the log-density $\mu = \log \rho$ at redshift $z = 0$ for different radii and hence variances as indicated in the legend for $n_s = -1.6$.

D3 Bias functions

D3.1 Density bias

As a complement to Fig. 5 that shows the density bias for redshift $z = 0.7$, we show in Fig. D5 the corresponding result for redshift $z = 0$ finding again excellent agreement with the simulation results.

D3.2 Joint density–slope bias

As a complement to Fig. 7 that shows the joint density bias for redshift $z = 0.7$, we show in Fig. D6 the corresponding results for redshift $z = 0$ finding again good agreement with the simulation.

APPENDIX E: COMPARING ML AND SAMPLE ESTIMATORS FOR ξ

In the main text, we have compared the precision reached on measuring the two-point correlation function using either an ML estimator given by equation (36) or the usual sample estimator given by equation (35). Let us here try and compare analytically their respective variance [assuming that our model for the joint PDF – see equation (33) – is exact].

Following the same procedure as in Codis et al. (2016a), we first focus on the ML estimator and compute the Fisher information. This maximum likelihood estimator, $\hat{\xi}_{\rm ML}$, by definition (Stuart & Ord 2009) converges towards the true value ξ in a probabilistic sense (via the so-called relation of consistence), and it is said to be asymptotic normal which means that the asymptotic distribution of $\sqrt{N}(\hat{\xi}_{\rm ML} - \xi)$ is a Gaussian of zero mean and variance given by the inverse Fisher information

$$\Sigma_{\rm ML}^2 = -1/\left\langle \mathcal{L}''(\rho, \rho'|\xi) \right\rangle, \tag{E1}$$

where $\mathcal{L}''(\rho, \rho'|\xi) = \partial^2 \log \mathcal{P}(\rho|\xi)/\partial \xi^2$ and the total number of pairs for N spheres is given by $N = N(N-1)/2$. If the joint PDF is exactly given by equation (33), the Fisher information reads

$$-\left\langle \mathcal{L}''(\rho, \rho'|\xi) \right\rangle = \left\langle \frac{(b(\rho)b(\rho'))^2}{(1 + \xi b(\rho)b(\rho'))^2} \right\rangle \approx \left\langle b^2(\rho) \right\rangle^2. \tag{E2}$$

For the spheres of radius $R = 15$ Mpc h^{-1} of the HR4 simulation at $z = 0.7$, it gives $\Sigma_{\rm ML} = 0.13$ in perfect agreement with
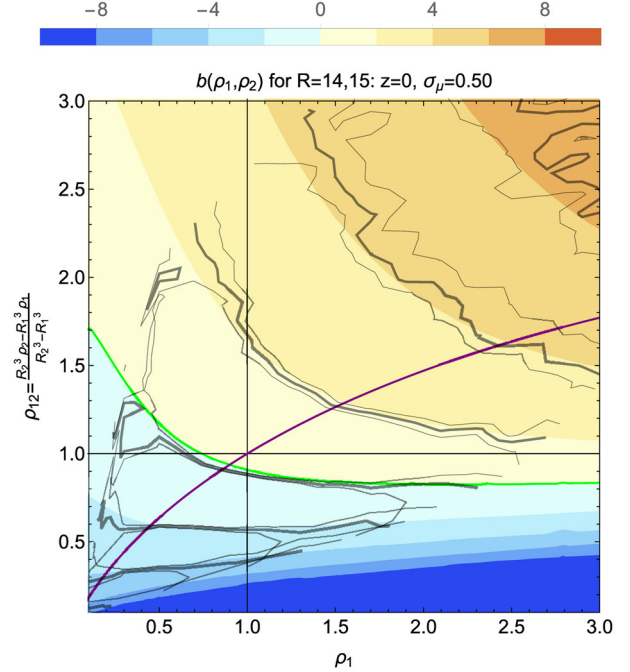

Figure D6. A contour plot of the joint density and slope bias function $b(\rho, s)$ predicted from the saddle-point approximation equation (29) with normalization from equation (18) for $R_1 = 14$ Mpc h^{-1} and $R_2 = 15$ Mpc h^{-1} at redshift $z = 0$ where $\sigma_\mu = 0.50$ (corresponding to $\sigma_\rho = 0.55$), in comparison to the measurements from HR4 (mean as thick black lines, and mean $\pm$ error on the mean as thin black lines). Shown is also the stationary line $\rho_2(\rho_1)$ (purple line) along which the joint bias function has to be evaluated to obtain the density bias $b(\rho_1) = b(\rho_1, \rho_2(\rho_1))$. The green line corresponds to zero bias.

the result obtained directly on the HR4 simulation as displayed in Fig. 12.

In contrast, the expected variance of the arithmetic estimator, $\sqrt{N}\hat{\xi}_{\rm A}$ is given by

$$\Sigma_{\rm A}^2 = \frac{\left\langle \left(\sum_{i,j} (\rho_i - 1)(\rho_j - 1) - N\xi \right)^2 \right\rangle}{N}. \tag{E3}$$

At leading order and after some algebra relying on Wick's theorem, we get

$$\Sigma_{\rm A,G}^2 \approx \sigma^4, \tag{E4}$$

which is exactly the Gaussian prediction for the maximum likelihood estimator

$$\Sigma_{\rm ML,G}^2 \approx \left\langle b^2(\rho) \right\rangle^{-2} = \left(\frac{\langle \tau^2 \rangle}{\sigma^4} \right)^{-2} = \sigma^4 \approx \Sigma_{\rm A,G}^2. \tag{E5}$$

As the field becomes non-Gaussian, the ML estimator performs increasingly better than the arithmetic estimator. The decrease of $\Sigma_{\rm A}/\Sigma_{\rm ML}$ can be studied numerically as was done in Fig. 12 that showed a factor of 5 increase in the variance of the arithmetic estimator compared to the ML approach.

This paper has been typeset from a T$_{\rm E}$X/L$^{\rm A}$T$_{\rm E}$X file prepared by the author.