Cylinders out of a top hat: counts-in-cells for projected densities

Cora Uhlemann[1,2,3], Christophe Pichon[4,5], Sandrine Codis[4,6], Benjamin L'Huillier[7], Juhan Kim[5], Francis Bernardeau[4,8], Changbom Park[5], Simon Prunet[4,9]

[1] *Centre for Theoretical Cosmology, DAMTP, University of Cambridge, CB3 0WA, United Kingdom*
[2] *Fitzwilliam College, University of Cambridge, CB3 0DG, United Kingdom*
[3] *Institute for Theoretical Physics, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands*
[4] *CNRS & UPMC, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[5] *Korea Institute for Advanced Study (KIAS), 85 Hoegiro, Dongdaemun-gu, Seoul, 02455, Republic of Korea*
[6] *Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada*
[7] *Korea Astronomy and Space Science Institute (KASI), 776 Daedeokdae-ro, Yuseong-gu, Daejeon 34055, Republic of Korea*
[8] *CNRS & CEA, UMR 3681, Institut de Physique Théorique, F-91191 Gif-sur-Yvette, France*
[9] *Canada France Hawaii Telescope Corporation 65-1238 Mamalahoa Hwy Kamuela, Hawaii 96743 USA*

November 16, 2017

ABSTRACT

Large deviation statistics is implemented to predict the statistics of cosmic densities in cylinders applicable to photometric surveys. It yields few percent accurate analytical predictions for the one-point probability distribution function (PDF) of densities in concentric or compensated cylinders; and also captures the density-dependence of their angular clustering (cylinder bias). All predictions are found to be in excellent agreement with the cosmological simulation Horizon Run 4 in the quasi-linear regime where standard perturbation theory normally breaks down. These results are combined with a simple local bias model that relates dark matter and tracer densities in cylinders and validated on simulated halo catalogues. This formalism can be used to probe cosmology with existing and upcoming photometric surveys like DES, Euclid or WFIRST containing billions of galaxies.

Key words: cosmology: theory — large-scale structure of Universe — methods: analytical, numerical

1 INTRODUCTION

Understanding the nature of dark energy is the main challenge cosmology is facing today, as it accounts for ∼70 per cent of the energy budget of our Universe. Multiplexed fiber or slit-less instruments allow astronomers to collect thousands of spectroscopic redshifts (either ongoing GAMA (Driver et al. 2016), VIPERS (Guzzo et al. 2014), or upcoming DESI (DESI Collaboration et al. 2016) PFS (Takada et al. 2014), MSE (McConnachie et al. 2016)). Such surveys provide a direct mean to probe accurately the cosmic evolution of the large-scale structure. Yet photometric redshift remains the most efficient method to map large volumes of the sky to position billions of galaxies. Obtaining galaxy statistics such as the one-point density PDF and the angular correlation function in thick redshift slices does not require modelling the cosmology-dependent distance-redshift relation and allows us to accumulate enough statistics for populations of rare objects. This has motivated international collaborations such as DES (Rykoff et al. 2016), Euclid (Laureijs et al. 2011), WFIRST (Spergel et al. 2013), LSST (LSST Dark Energy Science Collaboration 2012), KiDs (de Jong et al. 2013), to carry out extensive large scale structure surveys relying primarily on photometric redshifts to study the details of

structure formation at different epochs and therefore offer insight into the engine of cosmic acceleration.

The prime estimators which will be implemented on these surveys are angular clustering, weak lensing and supernovae. The former will rely on photometric redshifts to estimate the distance of the lensed galaxies and improve the derived constraint on the equation of state of dark energy while reconstructing the mass distribution of lenses on dozens of distinct planes simultaneously (implementing so called weak lensing tomography). The extracted catalogues can at no extra cost be used to carry out count-in-cell statistics (Bouchet et al. 1993; Efstathiou 1995; Yang & Saslaw 2011; Wolk et al. 2013; Bel & Marinoni 2014; Bel et al. 2016; Friedrich et al. 2017; Gruen et al. 2017). Count-in-cell has been identified as an ancillary probe for these space missions, as it accesses the non-linear regime of clustering analytically up to cosmic variance of the order of one. Indeed Uhlemann et al. (2016, 2017c) derived the theory of one- and two-point counts-in-cells statistics for spherical symmetry in three dimensions using a fully analytical saddle-point approximation as a substitute for the numerical integration required in Bernardeau et al. (2014a); Codis et al. (2016b). This analytic estimator can quantify more modes of the underlying clustering, hence has improved statistical power (Codis et al. 2016a). It

has also recently been extended to account for tracer bias so as to operate directly on galaxy catalogues (Uhlemann et al. 2017a).

This paper theoretically models the one- and two-point statistics of densities-in-concentric-cylinders that can be obtained from measured 2D maps of the large scale structure in redshift bins. It focuses on extending and combining Uhlemann et al. (2016, 2017a) to projected densities extracted from photometric – quasi-cylindrical – galaxy surveys with tracer bias. In a highly symmetric configuration such as cylindrical symmetry, one can take advantage of the fact that a non-linear solution to the gravitational dynamics (so-called cylindrical collapse) is known exactly and can be used to accurately probe the non-linear regime. In particular, this statistics allows us to study the density dependence of gravitational clustering in terms of the conditional one-point PDF of the density within underdense (resp. overdense) regions and the modulation of angular clustering therein.

It also provides the basis for an application to cosmic shear experiments sensitive to the matter distribution itself, such as the weak lensing signal around shear peaks (Kacprzak 2016), galaxy troughs (Gruen et al. 2016) or more general density split statistics (Friedrich et al. 2017; Gruen et al. 2017). Indeed, density-split statistics from counts and lensing in cells can yield cosmological constraints competitive with the two-point function measurements if the stochasticity of the tracer-halo connection can be controlled (see fig. 10 in Gruen et al. 2017). In this respect, theoretical progress has been made in deriving the one-point PDF for weak lensing shear and local aperture mass in the mildly non-linear regime (Bernardeau & Valageas 2000; Munshi & Jain 2000, 2001; Munshi et al. 2004, 2014; Reimberg & Bernardeau 2017). This is particularly interesting because, despite an approximate log-normality of 3D matter and tracers densities (Coles & Jones 1991; Kayo et al. 2001; Bel et al. 2016; Hurtado-Gil et al. 2017; Agrawal et al. 2017) and weak lensing statistics (Taruya et al. 2002; Hilbert et al. 2011; Clerkin et al. 2017), lognormal models have fundamental limitations regarding joint modelling of the two (Xavier et al. 2016).

The outline of the paper is the following. Section 2 presents how large deviation statistics can be used to obtain the PDF of dark matter densities in cylinders using cylindrical collapse dynamics and perturbative arguments. Section 3 compares these theoretical predictions for the one-point PDF and the two-point clustering of dark matter densities in cylinders to the Horizon Run 4 simulation. Such statistics allow us to probe differentially the slope of the density field in regions of low or high density. Section 4 demonstrates how tracer densities can be related to dark matter densities using a simple local bias relation and leads to accurate predictions for the statistics of tracers. Section 5 concludes and provides an outlook for future use of the results presented here. Appendix A contains further details on the treatment in the main text, in particular regarding the cylinder depth and leading order cumulants from perturbation theory, while Appendix B shortly addresses the accuracy of lognormal reconstructions.

2   LARGE DEVIATION FOR PROJECTED DENSITIES

The general formalism of large deviation statistics for cosmic densities has been presented in Bernardeau et al. (2014b); Bernardeau & Reimberg (2016); Uhlemann et al. (2016) for 3D spherical cells. It is presented here in the context of the density field smoothed in cylinders. In that case the general framework is left unchanged provided that

(i) The cylindrical instead of spherical filter is used to compute the initial correlation matrix of projected densities.

(ii) The cylindrical instead of spherical collapse is used to map initial to final densities.

(iii) The dimension of the cylinders instead of spheres is properly rescaled from the initial to the final conditions according to mass conservation.

(iv) The finite character of the redshift slices is accounted for.

After addressing those four points in the following subsection, Section 2.2 describes how to compute the density PDF from large deviation statistics.

2.1   Large deviation statistics with symmetric collapse models

When considering a highly symmetric observable such as the density in spheres or infinitely long cylinders, one can argue that the most likely dynamics (amongst all possible mappings between the initial and final density fields) is the one respecting the symmetry (Valageas 2002)[1]. For spherical and cylindrical 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 and cylindrical collapse model, respectively.

2.1.1   Filters for cylinders and spheres

To relate the average density in a region of a given shape to the underlying density field, the appropriate top-hat filters are used. For a finite cylinder with radius R and depth d, a top-hat filter is used in cylindrical coordinates $W(r_\perp, r_\parallel) \propto \Theta(R - r_\perp)\Theta(d/2 - r_\parallel)$ where $(r_\perp, r_\parallel)$ is the distance perpendicular and parallel to the cylinder axis, while Θ is the Heaviside step function. The Fourier transform of the cylindrical filter is a product of a two- and one-dimensional top-hat filter

$$\tilde{W}_{\text{cyl}}(Rk_\perp, dk_\parallel) = \frac{2J_1(Rk_\perp)}{Rk_\perp}\text{sinc}\left(\frac{dk_\parallel}{2}\right), \quad (1)$$

$$= \tilde{W}_{\text{2D}}(Rk_\perp)\tilde{W}_{\text{1D}}\left(\frac{dk_\parallel}{2}\right),$$

where J_1 is the Bessel function of the first kind of order one. For an infinitely long cylinder $d \to \infty$, the cylindrical filter (1) reduces to a disk top-hat filter $W_{\text{disk}}(r_{\text{2D}}) \propto \Theta(R - r_{\text{2D}})$ in 2 dimensions. In comparison, for a spherical top-hat filter $W_{\text{sph}}(r_{\text{3D}}) \propto \Theta(R - r_{\text{3D}})$ in 3 dimensions one has

$$\tilde{W}_{\text{sph}}(kR) = \frac{3(kR\cos(kR) - \sin(kR))}{(kR)^3}. \quad (2)$$

For a galaxy redshift survey, one would eventually consider truncated conical filters $W_{\text{cone}}(r_\parallel, \hat{\gamma})$, that select objects depending on their radial distance $r_\parallel$ and solid angle $\hat{\gamma}$ on the sky. For conical shapes, one typically puts a selection function $F(r_\parallel)$ along the radial direction (pointing along the line of sight $\hat{n}_{\text{los}}$) that encodes the mean density of observable objects at given distance (and hence redshift). Then one considers all solid angles $\hat{\gamma}$ around the line of sight $\hat{n}_{\text{los}}$ with an angle $\theta = \angle(\hat{n}_{\text{los}}, \hat{\gamma})$ smaller than a given aperture θ_0 such that $W_{\text{cone}}(r_\parallel, \hat{\gamma}) \propto F(r_\parallel)\Theta(\theta_0 - \theta)$. If the redshift

---

[1] 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 (symmetric collapse) dominates.

slices are thin compared to the radial distance of the redshift bin, a truncated cone should be virtually identical to a cylinder centred at the corresponding mean redshift.

Using equations (1) and (2), the filtered linear covariance for concentric volumes is then obtained from the linear power spectrum according to

$$\sigma_{s,L}^2(\mathcal{R}_{s,1}, \mathcal{R}_{s,2}) = \int \frac{\mathrm{d}^{d_s}\mathbf{k}}{(2\pi)^d} P_L(k) \tilde{W}_s(\mathbf{k}, \mathcal{R}_{s,1}) \tilde{W}_s(\mathbf{k}, \mathcal{R}_{s,2}), \quad (3)$$

where d_s encodes the dimension of integration and $\tilde{W}_s$ is the Fourier transform of an appropriate filter enclosing a volume of a given shape s (e.g cylinders, spheres) and characteristic dimensions $\mathcal{R}_s$. For a cylinder, $\mathcal{R}_{cyl} = \{R, d\}$ and for a sphere $\mathcal{R}_{sph} = R$. Fig. A3 shows a comparison of the scale-dependence of the variance for cylinders and spheres and Appendix A3 describes an efficient power-law parametrisation for the covariance.

2.1.2 From spherical to cylindrical collapse

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

$$\rho = \zeta_{SC}(\tau), \quad \text{with} \quad \rho R^{d_s} = R_{ini}^{d_s}, \quad (4)$$

where the initial and final radii are connected through mass conservation. Bernardeau (1992, 1995) showed that in 2D and 3D, the spherical collapse obeys the following differential equation

$$-\zeta_{SC}\tau^2\zeta_{SC}'' + c(\tau\zeta_{SC}')^2 - \frac{3}{2}\zeta_{SC}\tau\zeta_{SC}' + \frac{3}{2}\zeta_{SC}^2(\zeta_{SC}-1) = 0, \quad (5)$$

with $c_{2D} = 3/2$ and $c_{3D} = 4/3$. While equation (5) is strictly valid only for an Einstein-de Sitter background, it is also a good approximation for a general cosmological background. In both cases, one has $\zeta(\tau) \simeq 1 + \tau + \mathcal{O}(\tau^2)$ for small values of the linear density contrast τ as expected in the linear regime. The second and third orders then allows to predict the skewness from the spherical collapse dynamics as is described in Appendix A1. Equation (5) has a parametric solution for an Einstein-de Sitter background in three dimensions and can be solved numerically in other cases. An explicit possible fit for the relation between the initial density contrast τ and the final density ρ is given by

$$\rho_{SC}(\tau) = (1 - \tau/\nu)^{-\nu} \Leftrightarrow \tau_{SC}(\rho) = \nu(1 - \rho^{-1/\nu}), \quad (6)$$

where ν depends on the dimension and the symmetry of collapse ($d_s = 3$ for spherical and $d_s = 2$ for cylindrical/disk-like) and can be adjusted to the actual values of the cosmological parameters. The Zeldovich approximation, which is exact in 1D, corresponds to setting $\nu_{ZA} = 1$[2].

In 3D, $\nu_{3D} = 21/13 \simeq 1.6$ provides a good description of the spherical dynamics for an Einstein-de Sitter background for the range of τ values of interest. In particular, this number allows us to reproduce the tree-order reduced skewness S_3 which captures the behaviour for small and intermediate values of τ. From equation (5), Bernardeau (1995) found that for the 2D case asymptotically $\zeta(\tau) \propto \tau^{-(\sqrt{13}-1)/2}$ for large τ and therefore proposed to

use the parametrisation (6) with $\nu_{2D} = (\sqrt{13}-1)/2 \sim 1.30$. If instead, in analogy to the 3D case, one wants to reproduce the value of the tree-order unsmoothed skewness for cylindrical symmetry $S_3^{2D} = 36/7$ then one shall set $\nu_{2D} = 1.4$. Appendix A1 recaps the expression for the skewness at tree-order in perturbation theory and its relation to the spherical collapse parameter ν.

2.1.3 Relating initial to final volumes

Because of mass conservation, the dependence of the variance on the initial shape of the cells $\mathcal{R}_{ini}$ translates into a density dependence when the final dimensions of the cylinder $\mathcal{R} = \{R, d\}$ are fixed. Hence, a prescription of the form $\mathcal{R}_{ini} = \mathcal{R}_{ini}(\mathcal{R}, \rho)$ is needed to map the final to the initial dimensions of the cell given the enclosed final density. For a maximally symmetric case, like spherical symmetry in 3D, mass conservation completely fixes the relation between initial and final radii via $R_{sph}^{ini} = R_{sph}\rho_{sph}^{1/3}$. The same holds true for an infinitely long cylinder that corresponds to a 2D disk, for which mass conservation enforces to map the initial radius to the final one via

$$\text{long cylinders:} \quad R_{disk}^{ini} = R_{disk} \cdot \rho_{disk}^{1/2}, \quad (7)$$

where ρ_{disk} is the surface density in the cylinder. In these maximally symmetric cases, the only unknown is the initial radius and the sole constraint of mass conservation is sufficient to fix it.

Conversely, for a finite cylinder, two unknowns have to be determined (radius and depth) from one single constraint that relates the initial and final volume according to mass conservation $V_{cyl}^{ini} = V_{cyl}\rho_{cyl}$. One can then distribute the density ρ_{cyl} between the two scales of the problem: the radius R and depth d of the cylinder. For a cylinder which has sphere-like dimensions $d \simeq 2R_{2D}$, one can expect that the ratio of cylinder radius to depth remains constant since the cylinder should collapse similarly to an equivalent sphere. Hence, for a sphere-like cylinder one should perform an equal splitting

$$\text{sphere-like cylinders:} \quad (R_{cyl}^{ini}, d_{cyl}^{ini}) = (R_{cyl}, d_{cyl}) \cdot \rho_{cyl}^{1/3}, \quad (8)$$

as opposed to infinitely long cylinders for which all the density factor is attributed to the radius.

In the remainder of the text, the focus will be on long cylinders with axis ratios $2R/d < 1/10$ for which the approximation of infinitely long cylinder should hold.[3] Indeed, for those axis ratios, previous studies for prolate spheroids (elongated symmetric ellipsoids that should closely resemble cylinders) in Yoshisato et al. (1998, 2006) suggest that effective two-dimensional spherical collapse is accurate and the evolution of the axis corresponds to the one of infinitely long cylinders (7). This argument is seconded in Appendix A1 by comparing the skewness of the measured density field to tree-order perturbation theory.

2.2 The one-point density PDF in the large deviation regime

Let us now compute the one-point PDF of the density field smoothed in a cylinder from large deviation statistics.

---

[2] For an analytical treatment of the PDF arising from 1D gravitational collapse and non-perturbative effects therein, we refer to Pajer & van der Woude (2017) which also addresses the relation of the PDF in Lagrangian and Eulerian space.

[3] It was checked that as far as the one-point PDF is concerned, sphere-like cylinders with $R = 10$, $d = 20$ Mpc/h are well described by spherical collapse with $\nu = 1.6$ and a sphere-like mapping of the axis ratio as in (8).

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

The principles of large deviation statistics yield a formula for the PDF of finding a certain density in a symmetric volume given the initial conditions once it is assumed that the most likely dynamics that relates the initial to the final configurations is given by the appropriate collapse model. To achieve this goal, the main ingredient is the decay-rate function which encodes the exponential decay of the PDF. This formalism can be applied to any number of concentric symmetric volumes (e.g spheres, cylinders) as the symmetry is preserved in this configuration. For Gaussian initial conditions, which is assumed here[4], the initial PDF of density contrasts τ_k in N concentric cylinders with characteristic dimensions $\mathcal{R}_k = \{R_k, d\}$ can be written as

$$\mathcal{P}^{\mathrm{ini}}_{\{\mathcal{R}_k\}}(\{\tau_k\}) = \sqrt{\det\left[\frac{\partial^2 \Psi^{\mathrm{ini}}_{\{\mathcal{R}_k\}}}{\partial \tau_i \partial \tau_j}\right]} \frac{\exp\left[-\Psi^{\mathrm{ini}}_{\{\mathcal{R}_k\}}(\{\tau_k\})\right]}{(2\pi)^{N/2}}, \quad (9)$$

where $\mathcal{R}_k^{\mathrm{ini}} = (R_k^{\mathrm{ini}}, d_k^{\mathrm{ini}})$ encodes the characteristic dimensions of the initial cylinders. The initial decay-rate function is given by the usual quadratic form in the initial density contrasts τ_k in cylinders with characteristic dimensions $\mathcal{R}_k$

$$\Psi^{\mathrm{ini}}_{\{\mathcal{R}_k\}}(\{\tau_k\}) = \frac{1}{2}\sum_{i,j} \Xi_{ij}(\{\mathcal{R}_k^{\mathrm{ini}}\})\, \tau_i \tau_j, \quad (10)$$

where Ξ_{ij} is the inverse of the linear covariance matrix, $\sigma_L^2(\mathcal{R}_i^{\mathrm{ini}}, \mathcal{R}_j^{\mathrm{ini}})$ obtained from equation (3) and encoding all dependency with respect to the initial power spectrum and the volume which is averaged over. Note that equation (9) is merely a rewriting of a Gaussian distribution, emphasizing the central role of the decay-rate function (10).

2.2.2   Large deviation statistics PDF for an evolved field

The final decay-rate function is obtained from re-expressing the initial decay-rate function in terms of the final densities ρ and characteristic dimensions of the cylinder $\mathcal{R} = \{R, d\}$

$$\Psi_{\{\mathcal{R}_k\}}(\{\rho_k\}) = \frac{1}{2}\sum_{i,j} \Xi_{ij}(\{\mathcal{R}_k^{\mathrm{ini}}(\mathcal{R}_k, \rho_k)\})\tau_{\mathrm{SC}}(\rho_i)\tau_{\mathrm{SC}}(\rho_j), \quad (11)$$

where the mapping between initial and final densities $\tau_{\mathrm{SC}}(\rho)$ is given by the cylindrical collapse[5] from equation (6). The non-linear covariance is approximated by rescaling the linear covariance from equation (3) using a reference scale , $\mathcal{R}_p$

$$\sigma(\mathcal{R}_1^{\mathrm{ini}}, \mathcal{R}_2^{\mathrm{ini}}) \simeq \frac{\sigma(\mathcal{R}_p, \mathcal{R}_p)}{\sigma_L(\mathcal{R}_p, \mathcal{R}_p)}\sigma_L(\mathcal{R}_1^{\mathrm{ini}}, \mathcal{R}_2^{\mathrm{ini}}), \quad (12)$$

while the relation between the initial and final characteristic dimensions of the cylinders $\mathcal{R}^{\mathrm{ini}}(\mathcal{R}, \rho)$ is given by equations (7) for long cylinders. The non-linear variance at reference scale – also known as driving parameter in the context of large deviations theory – can be either measured directly from the densities in cylinders extracted

from the simulation, derived following equation (3) using the non-linear power spectrum or found by a best fit to the measured one-point PDF (Codis et al. 2016a).

The PDF of evolved densities in concentric cylinders, $\mathcal{P}_{\{\mathcal{R}_k\}}(\{\rho_k\})$, at one point can then be obtained in full analogy to the case for densities in concentric spheres as discussed in Uhlemann et al. (2016). Therein, it has been shown that the saddle-point technique provides an excellent approximation to the exact result from large-deviations statistics if a suitable variable is chosen. For a single cylinder the suitable variable is the log-density $\mu = \log \rho$ while for concentric cylinders one should use logarithmically mapped combinations of the densities, see Section 3.3 for the case of two cylinders. The PDF of those mapped densities is then given by

$$\mathcal{P}^{\mathrm{cyl}}_{\mu, \{\mathcal{R}_k\}}(\{\mu_k\}) = \sqrt{\det\left[\frac{\partial^2 \Psi_{\{\mathcal{R}_k\}}}{\partial \mu_i \partial \mu_j}\right]} \frac{\exp\left[-\Psi_{\{\mathcal{R}_k\}}\right]}{(2\pi)^{N/2}}, \quad (13a)$$

where Ψ is now given by equation (11). The PDF of the log-density can be related to the PDF of the density via a simple change of variables

$$\mathcal{P}^{\mathrm{cyl}}_{\{\mathcal{R}_k\}}(\{\rho_k\}) = \mathcal{P}^{\mathrm{cyl}}_{\mu, \{\mathcal{R}_k\}}[\{\mu_k(\{\rho_i\})\}] \left|\det\left[\frac{\partial \mu_i}{\partial \rho_j}\right]\right|. \quad (13b)$$

Equation (13a) assumes that the mean of μ_j vanishes independently of the variance. For a generic non-linear mapping, it will translate into a mean density which can deviate from one as σ grows. In order to avoid this effect, one has to consider the shifted PDF

$$\hat{\mathcal{P}}^{\mathrm{cyl}}_{\mu, \{\mathcal{R}_k\}}(\{\mu_k\}) = \mathcal{P}^{\mathrm{cyl}}_{\mu, \{\mathcal{R}_k\}}(\{\tilde{\mu}_k = \mu_k - \langle\mu_k\rangle\}), \quad (13c)$$

with the shifts $\langle\mu_k\rangle$ chosen such that the resulting mean densities are one $\langle\rho_i\rangle = 1\,\forall i = 1, \cdots, n$. Furthermore, since the saddle-point method yields only an approximation to the exact PDF, the PDF obtained from equation (13) is not necessarily perfectly normalized (Uhlemann et al. 2017c). In practice, this can be corrected for by considering

$$\hat{\mathcal{P}}^{\mathrm{cyl}}_{\mathcal{R}}(\{\rho_k\}) = \mathcal{P}^{\mathrm{cyl}}_{\mathcal{R}}(\{\rho_k\})/\langle 1 \rangle, \quad (13d)$$

with the shorthand notation $\langle 1 \rangle = \prod_k \int_0^\infty \mathrm{d}\rho_k\, \mathcal{P}^{\mathrm{cyl}}_{\mathcal{R}}(\{\rho_k\})$. Note that this effect is however very minor as the deviation of $\langle 1 \rangle$ from unity is typically at the sub-percent level.

2.3   The two-point cylinder bias in the large deviation regime

Let us finally close this section by turning to the two-point statistics of densities in cylinders and in particular to the two-point cylinder bias from large deviation statistics.

2.3.1   Kaiser bias for an initial Gaussian field

The initial two-point bias used in previous works on densities-in-spheres (see e.g. Bernardeau 1996; Codis et al. 2016b; Uhlemann et al. 2017c) can also be straightforwardly generalised to cylinders. Cylinder bias is defined as the ratio between the conditional mean density contrast induced by a given initial density contrast τ_0 in a cylinder of characteristic dimensions $\mathcal{R}_{\mathrm{ini}}$ at separation $r_\perp$ (perpendicular to the cylinder axis) and the average correlation

$$b_{\mathrm{ini}}(\tau_0, r_\perp) = \frac{\langle \tau(\mathbf{x} + \mathbf{r}_\perp)|\tau(\mathbf{x}) \equiv \tau_0\rangle}{\langle \tau(\mathbf{x} + \mathbf{r}_\perp)\tau(\mathbf{x})\rangle} = \frac{\int \mathrm{d}\tau\, \tau\, \mathcal{P}^{2\mathrm{pt}}_{\mathcal{R}_{\mathrm{ini}}}(\tau_0, \tau; r_\perp)}{\mathcal{P}_{\mathcal{R}_{\mathrm{ini}}}(\tau_0)\xi_L(\mathcal{R}_{\mathrm{ini}}, r_\perp)},$$

and can be expressed in terms of the joint PDF $\mathcal{P}^{2\mathrm{pt}}_{\mathcal{R}_{\mathrm{ini}}}$ of densities in cylinders of equal characteristic dimensions $\mathcal{R}_{\mathrm{ini}}$ at separation

---

[4] See Uhlemann et al. (2017b) for an extension of the formalism to include primordial non-Gaussianity.

[5] This is a result of the so-called contraction principle which can be derived rigorously in the zero variance limit (Bernardeau & Reimberg 2016). It states that the path integral over all possible initial realisations $\{X_{\mathrm{ini}}\}$ that lead to a given final realisation Y_{f} is dominated by the least suppressed realisation, $\Psi(Y_{\mathrm{f}}) = \inf_{\{X_{\mathrm{ini}}\} \to Y_{\mathrm{f}}} \Psi(X_{\mathrm{ini}})$, which in our case corresponds to symmetric collapse.

$r_\perp$ and their correlation function $\xi(\mathcal{R}_{\rm ini}, r_\perp) = \langle \tau(\mathbf{x} + \mathbf{r}_\perp)\tau(\mathbf{x})\rangle$. The initial correlation function is obtained from the linear power spectrum as

$$\xi_L(\mathcal{R}_{\rm ini}, r_\perp) = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} P_L(k)\tilde{W}^2(\mathbf{k}, \mathcal{R}_{\rm ini}) \exp(i\mathbf{k}_\perp \cdot \mathbf{r}_\perp). \quad (14)$$

For Gaussian initial conditions, it is straightforward to determine and diagonalise the linear covariance matrix to find that the cylinder bias predicted by large deviation theory is independent of separation $r_\perp$ and reads

$$b_{\rm ini}(\tau, r_\perp) \equiv b_{\rm ini}(\tau) = \frac{\tau}{\sigma_L^2(\mathcal{R}_{\rm ini})}, \quad (15)$$

which is proportional to the initial overdensity τ as expected from Kaiser (1984). Since the cylinder bias is independent of the separation $r_\perp$, the density dependence of the two-point statistics of densities in cylinders can be factorised at large separations $r_\perp \gg 2R$ according to

$$\frac{\mathcal{P}_{\mathcal{R}_{\rm ini}}^{\rm 2pt}(\tau, \tau'; r_\perp)}{\mathcal{P}_{\mathcal{R}_{\rm ini}}(\tau)\mathcal{P}_{\mathcal{R}_{\rm ini}}(\tau')} \simeq 1 + \xi_L(\mathcal{R}_{\rm ini}, r_\perp)b_{\rm ini}(\tau)b_{\rm ini}(\tau'). \quad (16)$$

2.3.2 *Large deviation statistics bias for an evolved field*

The cylinder bias can be defined in the evolved density field in complete analogy to the initial one from equation (14) and reads

$$b_{\mathcal{R}}^{\rm cyl}(\rho_0, r_\perp) = \frac{\langle \rho(\mathbf{x} + \mathbf{r}_\perp) | \rho(\mathbf{x}) \equiv \rho_0 \rangle - 1}{\langle \rho(\mathbf{x} + \mathbf{r}_\perp)\rho(\mathbf{x})\rangle - 1}. \quad (17)$$

Following the same prescription as for the PDF for replacing initial and final densities and characteristic dimensions of the cylinder, large deviation theory predicts that the final cylinder bias is independent of the transverse separation and given by

$$b_{\mathcal{R}}^{\rm cyl}(\rho) = \frac{\sigma_L^2(\mathcal{R})}{\sigma_\mu^2(\mathcal{R})}\frac{\tau_{\rm SC}(\rho)}{\sigma_L^2(\mathcal{R}_{\rm ini}(\mathcal{R}, \rho))}. \quad (18a)$$

Equation (18a) must also be shifted and normalised (Uhlemann et al. 2017c) according to

$$\hat{b}_{\mathcal{R}}^{\rm cyl}(\rho) = \frac{b_{\mathcal{R}}^{\rm cyl}(\rho) - \langle b_{\mathcal{R}}^{\rm cyl}(\rho)\rangle}{\langle \rho b_{\mathcal{R}}^{\rm cyl}(\rho)\rangle - \langle b_{\mathcal{R}}^{\rm cyl}(\rho)\rangle}. \quad (18b)$$

where the averages denoted by $\langle \cdot \rangle$ are computed as integrals with the one-point PDF from equation (13). As before, the prediction depends on the assumed mapping between initial and final radii and depths and the choice of the filter function for obtaining the linear variance σ_L.

3 VALIDATION FOR DARK MATTER DENSITIES

Let us now compare the theoretical predictions for the one-point PDF and the two-point clustering of dark matter densities in cylinders to the cosmological N-body simulation Horizon Run 4. To avoid confusion between dark matter densities considered here and tracer densities that will be used in Section 4, from now on $\rho_{\rm m}$ stands for matter densities and $\rho_{\rm h}$ for halo densities.

3.1 Horizon Run 4 simulation

The Horizon Run 4 simulation (HR4, Kim et al. 2015) is a massive N-body simulation, evolving 6300^3 particles in a 3.15 Gpc/h box using the GOTPM TreePM code (Dubinski et al. 2004). It assumes a WMAP-5 cosmology, with $(\Omega_{\rm m}, \Omega_\Lambda, \Omega_{\rm b}, h, \sigma_8, n_{\rm s}) =$

$z = 0.7$	R \ d	3	5	7	10
$\hat{\sigma}_\rho$	150	0.330	0.274	0.242	0.209
$\hat{\sigma}_\mu$	150	0.300	0.259	0.232	0.203
$\hat{\sigma}_\rho$	300	0.235	0.196	0.173	0.150
$\hat{\sigma}_\mu$	300	0.22	0.190	0.169	0.148
$\hat{\sigma}_\mu$	450	0.193	0.161	0.142	0.123
$\hat{\sigma}_\mu$	450	0.185	0.157	0.140	0.122
$\hat{\sigma}_\rho$	700	0.154	0.129	0.114	0.099
$\hat{\sigma}_\mu$	700	0.150	0.127	0.113	0.099

Table 1. Variances of the cylindrical density ρ and log-density $\mu = \log \rho$ for different lengths d [Mpc/h] and radii R [Mpc/h] at redshift $z = 0.7$ as measured from the HR4 simulation.

$(0.26, 0.74, 0.044, 0.72, 0.79, 0.96)$, yielding a particle mass of $9 \times 10^9 h^{-1}\, M_\odot$. The initial conditions were generated at $z = 100$ using the second order Lagrangian perturbation theory, which ensures accurate power spectrum and halo mass function at redshift 0 (see L'Huillier et al. 2014, for details). For the purpose of this paper, cylinders of various radii and lengths were extracted from the dark matter and halo catalogues at fixed redshift and used to measure the statistical properties of their encompassed mean densities. Error bars for the PDFs are determined from the standard error on the mean that is obtained by dividing the full simulation cube into 8 subsamples and computing the corresponding histograms. A spherical slice of width 60 Mpc/h was also extracted at redshift 0.36 and projected onto the sky using the HealPix[6] equal area scheme to produce Fig. 9.

3.2 One-point PDF of density in single cylinders

The explicit formula for the one-point PDF for dark matter densities $\rho_{\rm m}$ within a cylinder of radius R and depth d at redshift z is expressed as

$$\mathcal{P}_{\mathcal{R}}^{\rm cyl}(\rho_{\rm m}) = \sqrt{\frac{\Psi_R''(\rho_{\rm m}) + \Psi_R'(\rho_{\rm m})/\rho_{\rm m}}{2\pi\sigma_\mu^2}} \exp\left(-\frac{\Psi_R(\rho_{\rm m})}{\sigma_\mu^2}\right), \quad (19)$$

where the prime denotes a derivative with respect to $\rho_{\rm m}$ and

$$\Psi_{\mathcal{R}}(\rho_{\rm m}) = \frac{\tau_{\rm SC}^2(\rho_{\rm m})\sigma_L^2(\mathcal{R})}{2\sigma_L^2(\mathcal{R}_{\rm ini}(\mathcal{R}, \rho_{\rm m}))}. \quad (20)$$

Note that $\sigma_\mu \equiv \sigma_\mu(\mathcal{R}, z)$ is the non-linear variance of the log-density in a cylinder with characteristic dimensions $\mathcal{R} = (R, d)$, which enters because the formula has been derived from an analytic approximation based on the log-density $\mu_{\rm m} = \log \rho_{\rm m}$, while σ_L is the linear variance. Table 1 reports the values for the variance of the density and log-density for different radii and depths of the cylinders. In Equation (20), $\tau_{\rm SC}(\rho_{\rm m})$ is the linear density contrast which can be mapped to the non-linearly evolved density $\rho_{\rm m}$ using the collapse model (6) where the parameter $\nu = 1.4$ characterises the dynamics of cylindrical collapse and matches the high-redshift skewness obtained from perturbation theory. The initial (Lagrangian) characteristic dimensions of the cylinder $\mathcal{R}_{\rm ini}(\mathcal{R}, \rho_{\rm m})$ are mapped using equation (7) for long cylinders. To ensure a unit mean density and the correct normalization of the PDF, as described in Sec-

---

[6] http://healpix.jpl.nasa.gov/

tion 2.2 one can simply evaluate the PDF obtained from equation (19) according to

$$\hat{\mathcal{P}}^{\rm cyl}_{\mathcal{R}}(\rho_{\rm m}) = \mathcal{P}^{\rm cyl}_{\mathcal{R}} \left(\rho_{\rm m} \frac{\langle \rho_{\rm m} \rangle}{\langle 1 \rangle} \right) \cdot \frac{\langle \rho_{\rm m} \rangle}{\langle 1 \rangle^2} , \qquad (21)$$

with the shorthand notation $\langle f(\rho_{\rm m}) \rangle = \int_0^\infty {\rm d}\rho_{\rm m}\, f(\rho_{\rm m}) \mathcal{P}^{\rm cyl}_{\mathcal{R}}(\rho_{\rm m})$.

Fig. 1 compares the theoretical prediction for long cylinders with fixed depth $d = 150$ Mpc/h and various radii R to the simulation results. The upper panel gives an overall view on the exponential decay of the one-point PDF for large deviations from the mean density. It is complemented by the lower panel which shows the residuals around the maximum of the PDF, and displays the 2-σ_μ region around the mean for the log-density since the PDF of the log-density is close to Gaussian and hence allows for a fair comparison of under- and over-densities. An equivalent plot with fixed radius $R = 5$ Mpc/h and varying depth d is shown in Fig. A2. Given the few percent-level agreement between the theoretical prediction and the simulation[7], a future application of counts-in-cells statistics to photometric surveys is in order. The description is next generalised to concentric cylinders and two-point clustering that allow to consider density-split statistics. To connect the results for dark matter to tracer densities, tracer bias is also discussed in Section 4 following closely Uhlemann et al. (2017a) .

3.3 One-point PDF of densities in concentric cylinders

3.3.1 Two concentric cylinders

It is of interest to consider the joint PDF of concentric cylinders as it allows us to deduce conditionals which may e.g. act as proxies for peak or voids. To obtain such PDF for the densities ρ_1, ρ_2 in the cylinders of equal depth d and radii $R_1 < R_2$, one can also apply a logarithmic density mapping that relies on the sum and difference of mass in the two cylinders, (see for the analogous sphere case Uhlemann et al. 2016)

$$\mu_1 = \log \left(r_{12}^2 \rho_{2,\rm m} + \rho_{1,\rm m} \right), \ \mu_2 = \log \left(r_{12}^2 \rho_{2,\rm m} - \rho_{1,\rm m} \right), \quad (22)$$

where the relative shell thickness is $r_{12} = R_2/R_1 > 1$. The PDF $\mathcal{P}^{\rm cyl}_{R_1,R_2}(\rho_{1,\rm m}, \rho_{2,\rm m})$ can then be approximated via equation (13). Analogously to the one cell case, one still has to enforce the mean and normalization for the saddle point PDF following the procedure described in equations (13c)-(13d). It is convenient to express the joint density PDF $\mathcal{P}^{\rm cyl}_{R_1,R_2}(\rho_{1,\rm m}, \rho_{2,\rm m})$ in terms of the inner density $\rho_{1,\rm m}$ and the density in the outer shell $\rho_{12,\rm m} = (R_2^2 \rho_{2,\rm m} - R_1^2 \rho_{1,\rm m})/(R_2^2 - R_1^2)$. Note that for the sake of simplicity, a simpler scheme is used to compute the linear covariance matrix here, namely a power-law parametrisation of the power spectrum (A6) with index $n = n_{\rm cyl}(R, d)$ according to equation (A7) instead of the exact linear power spectrum, see Appendix A3.

The result is displayed in Fig. 2 and compared to measurements. Overall a very good agreement is found.

3.3.2 One-point PDF of densities in cylindrical rings

From the joint PDF of densities in concentric cylinders, one can obtain the PDF of a cylindrical shell and conditionals given inner
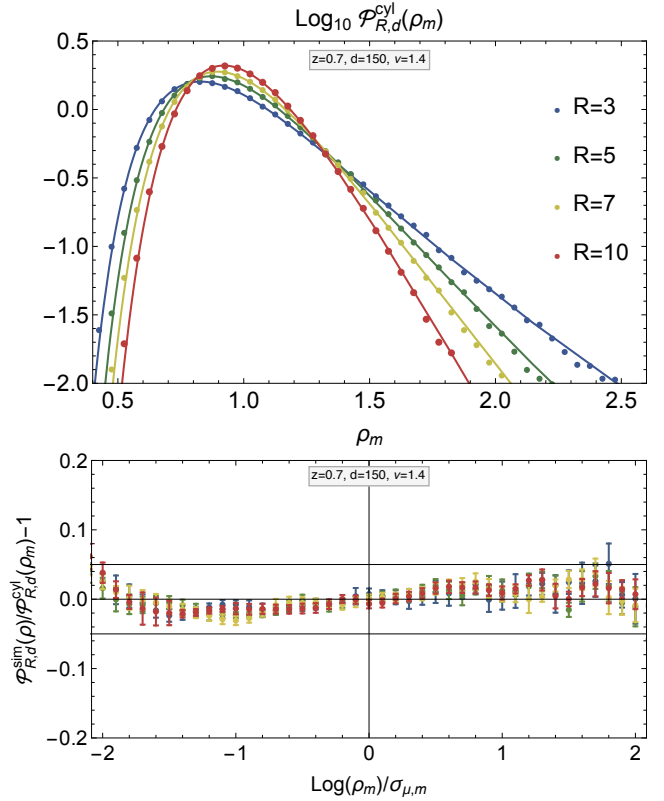

Figure 1. *(Upper panel)* PDF of the matter density field smoothed in cylindrical cells of length $d = 150$ Mpc/h and radii $R = 3, 5, 7, 10$ Mpc/h at redshift $z = 0.7$. Shown are the HR4 measurements (data points) against the theoretical predictions computed with the finite cylinder filter given by equations (1) and (7) together with the spherical collapse parametrized by $\nu = 1.4$ (solid lines). *(Lower panel)* Residuals of the theoretical prediction using finite cylinder filter against the measurements around the maximum of the PDF for log density within two-sigma from the mean.

over- or under-density via a simple marginalisation over the inner density

$$\mathcal{P}^{\rm shell}_{R_1,R_2}(\rho_{12,\rm m}) = \int {\rm d}\rho_{1,\rm m} \mathcal{P}^{\rm cyl}_{R_1,R_2}(\rho_{1,\rm m}, \rho_{12,\rm m}), \qquad (23)$$

$$\mathcal{P}^{\rm shell}_{R_1,R_2}(\rho_{12,\rm m}|\rho_{1,\rm m} \gtrless 1) \qquad (24)$$

$$= \frac{\int {\rm d}\rho_{1,\rm m} \mathcal{P}^{\rm cyl}_{R_1,R_2}(\rho_{1,\rm m}, \rho_{12,\rm m})\Theta(\pm(\rho_{1,\rm m}-1))}{\int {\rm d}\rho_{12,\rm m} \int {\rm d}\rho_{1,\rm m} \mathcal{P}^{\rm cyl}_{R_1,R_2}(\rho_{1,\rm m}, \rho_{12,\rm m})\Theta(\pm(\rho_{1,\rm m}-1))} .$$

A comparison between the marginal PDF and conditionals for over- and under-dense environments is shown in Fig. 3, while Fig. 4 focuses on the marginal PDF in cylindrical rings. For both statistics, very good agreement is found between the simulation measurements and the theoretical prediction, despite the simplifying assumption for the power spectrum. This is encouraging because statistics conditional on over- or underdensities have recently gained renewed interest in the context of weak lensing around galaxy troughs (Gruen et al. 2016) or more general density quantiles (Friedrich et al. 2017). While the statistics was here conditioned on the underlying dark matter density, in practice one has to rely on tracer densities. To this end, the inclusion of tracer bias is discussed in Section 4. The marginal PDF of the densities in cylindrical rings gives a qualitative impression of the ex-
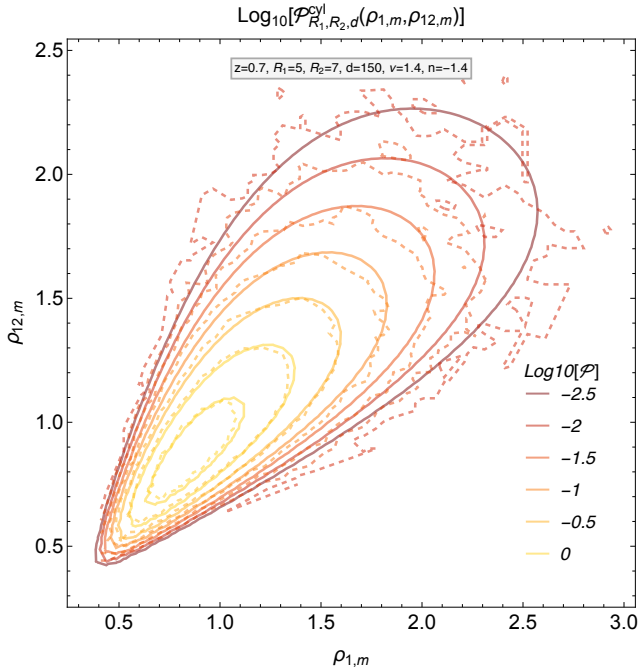
Figure 2. Joint PDF of the projected matter densities in concentric cylinders of depth $d = 150$ Mpc/h and radii $R_1 = 5$ Mpc/h and $R_2 = 7$ Mpc/h at redshift $z = 0.7$ from the saddle point approximation with $\nu = 1.4$ and the disk filter for a power-law spectrum with $n = -1.4$ (solid lines) compared to the HR4 measurements (dashed lines).

pected weak lensing signal that can probe the dark matter density up to a mass-sheet degeneracy, hence requires the use of compensated filters such as the difference of two cylinders. Note that the relation between conical cells and long cylinders can change the underlying analytical properties of the cumulant generating functions (see equation (34) in Bernardeau & Valageas 2000). However, large deviation statistics can be applied to many continuous nonlinear functionals of density profiles (Reimberg & Bernardeau 2017). Investigating approximations to compute PDFs for weak lensing convergence and cosmic shear fully analytically in the spirit of the log-density will be the subject of future work.

3.4 Two-point cylinder bias

In analogy to the 3D case leading to sphere bias described in Codis et al. (2016b); Uhlemann et al. (2017c), cylinder bias encodes the density dependence of large-scale angular clustering. This idea is related to the concept of sliced or marked correlation functions, see for example early works of Sheth & Tormen (2004); White & Padmanabhan (2009) as well as Neyrinck et al. (2016); White (2016). Cylinder bias is defined as the ratio between the mean density ρ' at separation $r_\perp$ (perpendicular to the axis of the cylinder) from a given density ρ in a cylinder of radius R and depth d and the average angular-correlation function, see equation (17). In the simulation, the cylinder bias is measured by using 4 neighbours at separation $r_\perp$ (along the two grid directions perpendicular to the cylinder axis) for every cylinder. A comparison between the prediction from equation (18) using a cylindrical filter and the measurement from the simulation is shown in Fig. 5 for fixed cylinder depth and varying radius. Except for the very low density end, there is a very good
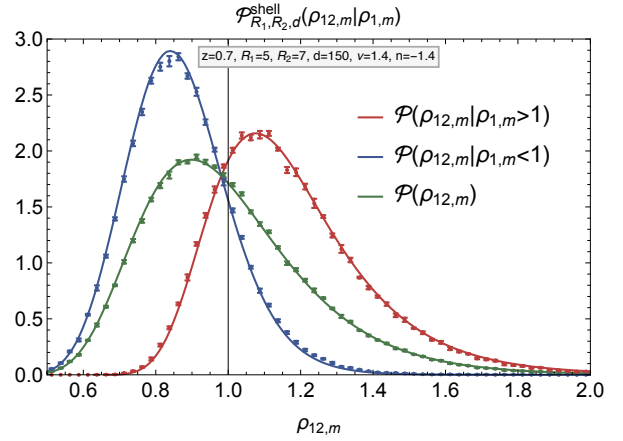

Figure 3. Marginalised PDF of the matter densities in cylindrical rings of depth $d = 150$ Mpc/h and radii $R_1 = 5$ Mpc/h and $R_2 = 7$ Mpc/h at redshift $z = 0.7$ from the saddle point approximation with $\nu = 1.4$ and the disk filter for a power-law spectrum with $n = -1.4$ for arbitrary inner densities (green line), overdensities (red line) and underdensities (blue line) compared to the HR4 measurements (data points).
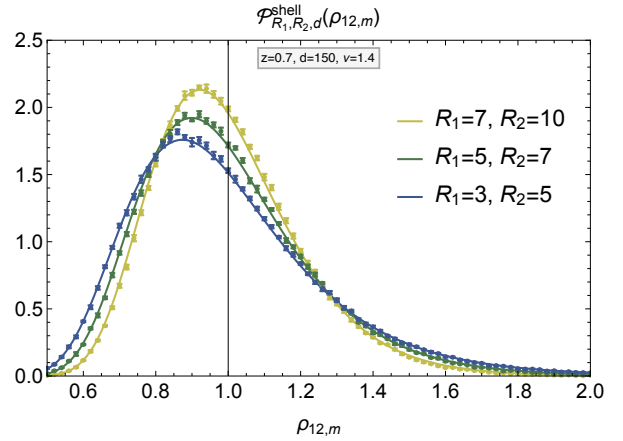

Figure 4. Marginalised PDF of the matter densities in cylindrical shells of depth $d = 150$ Mpc/h and radii $(R_1, R_2) = \{(3, 5), (5, 7), (7, 10)\}$ Mpc/h at redshift $z = 0.7$ from the saddle point approximation with $\nu = 1.4$ and the disk filter for a power-law spectrum with $n = -\{1.5, 1.4, 1.3\}$ (colored lines) compared to the HR4 measurements (points with error bars).

agreement between the theoretical prediction and the measurement in the simulation.

The cylinder bias can be used to approximate the two-point PDF of densities in cylinders at large separation $r_\perp > 2R$ according to

$$\mathcal{P}_\mathcal{R}^{2\text{pt}}(\rho_\text{m}, \rho'_\text{m}, r_\perp) = \mathcal{P}_\mathcal{R}^\text{cyl}(\rho_\text{m}) \mathcal{P}_\mathcal{R}^\text{cyl}(\rho'_\text{m}) \qquad (25)$$
$$\times \left[1 + \xi_\text{m}(r_\perp) b_\mathcal{R}^\text{cyl}(\rho_\text{m}) b_\mathcal{R}^\text{cyl}(\rho'_\text{m}) \right],$$

and is hence also useful to quantify correlations between neighbouring cylinders that contribute to the error budget of PDFs (Codis et al. 2016b).
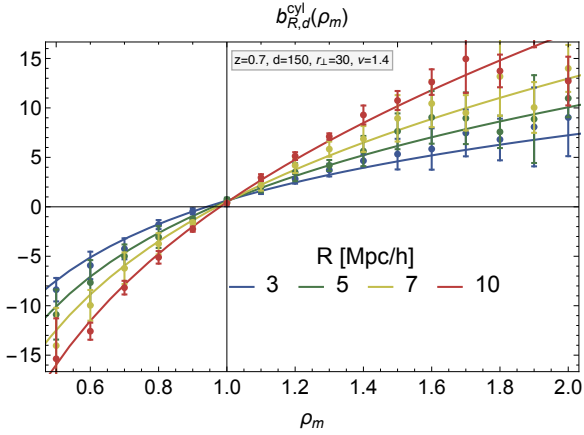
Figure 5. Two-point cylinder bias of dark matter densities in cylinders of depth $d = 150$ Mpc/h and varying radii R at separation $r_\perp = 30$ Mpc/h and redshift $z = 0.7$. The HR4 measurement (data points) is compared to the theoretical prediction from equation (18) using the cylindrical filter.

4 APPLICATION TO TRACER DENSITIES

Let us now investigate tracer counts using a local bias relation.

4.1 Halo identification

The haloes in HR4 were detected using Ordinary Parallel Friends-of-Friends (OPFOF, Kim & Park 2006), a massively parallel implementation of the friends-of-friends (FoF) algorithm, using a canonical linking length of 0.2 mean particle separations. Subhaloes were detected by the Physically Self-Bound algorithm (PSB, Kim & Park 2006), which finds the density peaks within each FoF halo, removes unbound particles, similarly to the SUBFIND halo finder, and additionally truncates the subhaloes to their tidal radius. All subhaloes with more than 30 particles were considered, yielding a masses from $2.7 \times 10^{11} h^{-1} M_\odot$ to $4.2 \times 10^{15} h^{-1} M_\odot$. Following the observations made in Seljak (2009); Hamaus et al. (2010); Jee et al. (2012) and Uhlemann et al. (2017a), let us consider a mass-weighted halo density (instead of number-weighted) because this makes the bias relation much tighter and considerably reduces the scatter. This can be intuitively understood as the mass-weighted halo densities resemble the overall dark matter density more closely than halo number. Note however that the mass-weighted densities of subhaloes are expected to be very similar to the mass weighted density of haloes (with no substructure) as the mass is almost preserved from haloes to subhaloes. Note finally that the subhaloes can be related to galaxies using for instance abundance matching techniques (Kravtsov et al. 2004; Vale & Ostriker 2004).

4.2 A local bias model for halo densities in cylinders

The treatment of mass-weighted halo densities in cylinders as biased tracers of the dark matter closely follows the procedure presented in Uhlemann et al. (2017a) for densities in spheres. In order to map the dark matter statistics to the halo statistics, let us rely on an 'inverse' quadratic bias model $\mu_\mathrm{m}(\mu_\mathrm{h})$ (writing the matter densities ρ_m as a function of the tracer densities ρ_h) for the log-densities $\mu_a = \log \rho_a$ (a=m,h) which reads

$$\mu_\mathrm{m} = \sum_{n=0}^{n_\mathrm{max}} b_n \mu_\mathrm{h}^n, \quad \text{with} \quad n_\mathrm{max} = 2. \quad (26)$$

	variance		correlation		bias		
R	$\sigma_{\mu,\mathrm{m}}$	$\sigma_{\mu,\mathrm{h}}$	$\xi_{\rho,\mathrm{m}}$	$\xi_{\rho,\mathrm{h}}$	b_0	b_1	b_2
3	0.300	0.683	0.0088	0.0040	0.0199	0.4643	0.052
5	0.259	0.546	0.0081	0.0041	0.0184	0.4921	0.0562
7	0.232	0.472	0.0075	0.0042	0.0177	0.5067	0.0583
10	0.203	0.402	0.0067	0.0043	0.0175	0.5177	0.0627

Table 2. Collection of simulation results for different radii R and depth $d = 150$ Mpc/h at redshift $z = 0.7$. The measured nonlinear variances σ_μ of the log-density $\mu = \log \rho$ and the correlation ξ of the density ρ at separation $r_\perp = 30$ Mpc/h of both dark matter (m) and haloes (h) in real space along with the bias parameters obtained from fitting the quadratic model (26) to the CDF bias function (27).

Following the idea of Sigad et al. (2000); Szapudi & Pan (2004), a direct way to obtain the mean bias relation is to use the properties of the cumulative distribution functions (CDFs), defined as $\mathcal{C}_a(\rho_a) = \int_0^{\rho_a} d\rho' \mathcal{P}_a^\mathrm{cyl}(\rho')$ with index $a = \mathrm{m}$ for matter and $a = \mathrm{h}$ for halos, so that

$$\mathcal{C}_\mathrm{m}(\rho_\mathrm{m}) = \mathcal{C}_\mathrm{h}(\rho_\mathrm{h}) \Rightarrow \rho_\mathrm{m}(\rho_\mathrm{h}) = \mathcal{C}_\mathrm{m}^{-1}(\mathcal{C}_\mathrm{h}(\rho_\mathrm{h})). \quad (27)$$

This is used to verify the accuracy of the polynomial log-bias model by fitting the bias parameters in equation (26) to the parametrisation-independent bias function obtained from the CDFs and comparing the two resulting functions.

Fig. 6 presents a scatter plot showing ρ_h as a function of ρ_m for redshift $z = 0.7$, depth $d = 150$ Mpc/h and radius $R = 5$ Mpc/h in order to assess how well bias models characterise the halo density bias. The lines correspond to the mean bias obtained in a parametrisation-independent way from the CDF method (red line) and fits based on a quadratic bias model for the log-densities (dashed and solid orange line) according to equation (26). The corresponding values of the best-fit bias parameters are given in Table 2 for different radii at $z = 0.7$. The second-order bias model for the logarithmic densities based on equation (26) agrees almost perfectly with the parametrisation-independent way of inferring bias using CDFs as in equation (27) and matches simulation results very well. An approximation of the standard linear bias parameter between matter and halo densities can be obtained by expanding the logarithmic relation to get $\delta_h/\delta_m \simeq [\exp(b_0)b_1]^{-1}$ which gives values close to 2 and slightly decreases with increasing radius.

4.3 One-point PDF and two-point cylinder bias for tracers

Equipped with a bias model for the mean relation $\rho_\mathrm{m}(\rho_\mathrm{h})$, the halo PDF $\mathcal{P}_\mathrm{h}$ is now obtained from the dark matter PDF $\mathcal{P}_\mathrm{m}^\mathrm{cyl}$ in equation (13) by conservation of probability

$$\mathcal{P}_\mathrm{h}^\mathrm{cyl}(\rho_\mathrm{h}) = \mathcal{P}_\mathrm{m}^\mathrm{cyl}(\rho_\mathrm{m}(\rho_\mathrm{h})) |d\rho_\mathrm{m}/d\rho_\mathrm{h}|. \quad (28)$$

One can also obtain the modulation of the two-point correlation function, the cylinder bias $b_\circ$ for haloes, from the result for dark matter given in equation (18)

$$b_\mathrm{h}^\mathrm{cyl}(\rho_\mathrm{h}) = b_\mathrm{m}^\mathrm{cyl}(\rho_\mathrm{m}(\rho_\mathrm{h})) \sqrt{\xi_\mathrm{m}/\xi_\mathrm{h}}, \quad (29)$$

where the ratio of correlation functions is given by

$$\sqrt{\xi_\mathrm{h}/\xi_\mathrm{m}} = \left\langle \rho_\mathrm{h}(\rho_\mathrm{m}) b_\mathrm{m}^\mathrm{cyl}(\rho_\mathrm{m}) \right\rangle, \quad (30)$$

and can be approximated by expanding the log-bias relation to first order to obtain $\sqrt{\xi_\mathrm{m}/\xi_\mathrm{h}} \simeq \exp(b_0)b_1$. A validation of those theoretical predictions against the numerical simulation is shown in Fig. 7 for the one point PDF and in Fig. 8 for cylinder bias as function of halo density. The agreement is once again very good.
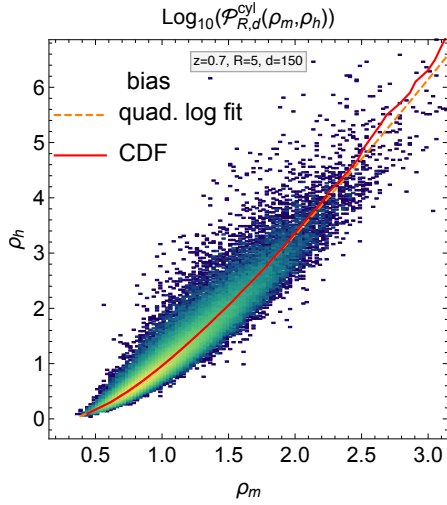
Figure 6. Density scatter plot of the halo density ρ_h with mass-weighting (blue-green region) versus the dark matter density ρ_m for a cylinder of radius $R = 5$ Mpc/h and depth $d = 150$ Mpc/h at redshift $z = 0.7$. The figure also shows the best-fit quadratic bias model for the log-density obtained from a fit to the scatter plot (dashed, orange) and to the CDF bias function (orange line, overlapping with the red line) which agrees very well with the parametrisation-independent bias obtained from the CDF (red line) in particular in the intermediate density region .
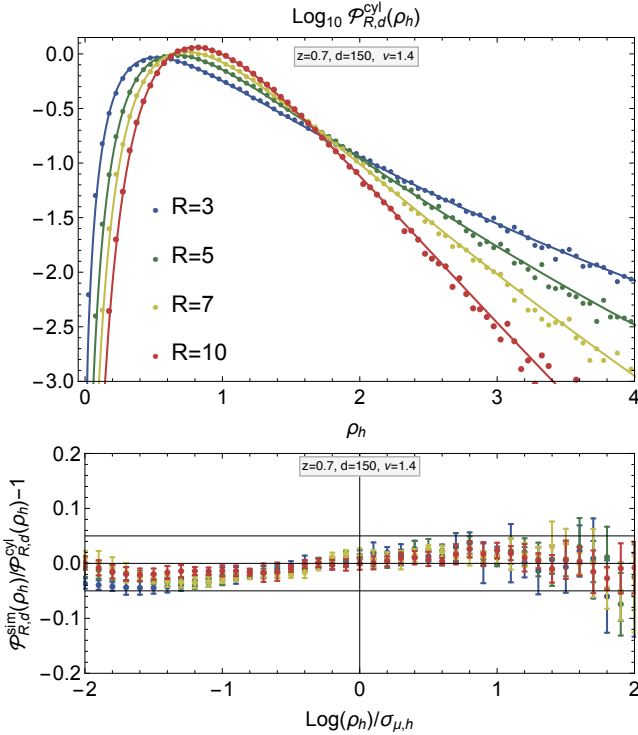
Figure 7. PDF of the mass-weighted halo density in cylindrical cells of length $d = 150$ Mpc/h and radii $R = 3, 5, 7, 10$ Mpc/h at redshift $z = 0.7$. Shown is a comparison of HR4 measurements (data points) and the prediction from equation (28) computed with the cylindrical filter (1) together with the spherical collapse parametrized by $\nu = 1.4$ (solid lines) and the fitted bias parameters from Table 2.
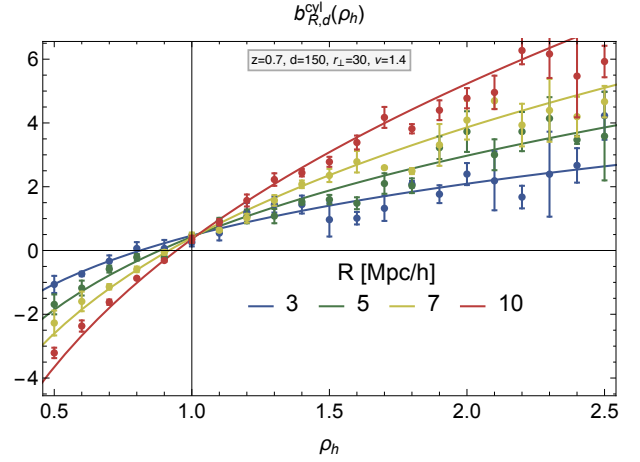
Figure 8. Two-point cylinders bias of mass-weighted halo densities in cylinders of depth $d = 150$ Mpc/h and radii $R = 3, 5, 7, 10$ Mpc/h at separation $r_\perp = 30$ Mpc/h and redshift $z = 0.7$. The HR4 measurement (data points) is compared to the theoretical prediction from equation (29) using the cylindrical filter and the best-fit bias parameters from Table 2.

4.4 Conditional PDF of DM density given tracer densities

The conditional PDF of dark matter density in cylindrical rings given a central tracer density is interesting because it can be used to probe the matter density profile (for example measured through weak lensing) around regions of fixed tracer density. The matter density profile can be related to the tangential shear or lensing convergence profile around under- and overdense regions as is done in Friedrich et al. (2017). The mixed halo matter joint PDF can be obtained from the joint PDF of dark matter using the one-to-one bias mapping

$$\mathcal{P}^{cyl}_{R_1,R_2}(\rho_{1,h}, \rho_{2,m}) = \mathcal{P}^{cyl}_{R_1,R_2}(\rho_{1,m}(\rho_{1,h}), \rho_{2,m}) \left| \frac{d\rho_{1,m}}{d\rho_{1,h}} \right| , \quad (31)$$

and then defining conditionals in analogy to equation (24). Similar to Figs. 2 and 4 showing the joint PDF of matter in concentric cylinders and the conditional PDFs of density in rings given an inner dark matter over- or under-density, Fig.s A4 and A5 show the case of a mixed joint PDF of dark matter around tracer density.

4.5 Prospects for dark energy experiments

Given the accuracy of the theoretical predictions for one and two-point statistics of halo densities in cylinders, counts-in-cells statistics seem promising for constraining cosmological parameters. In particular, measuring the tracer PDF in relatively thin redshift slices through photometric surveys gives access to the time evolution of clustering and hence can probe dark energy. Let us discuss briefly how the joint fit of estimators for counts-in-cells statistics presented in the previous subsection could be used to constrain dark energy. In this context the goal is to estimate the so-called equation of state of dark energy $w(z) = w_0 + w_a/(1 + z)$ with parameters (w_0, w_a) from the PDF while relying on the cosmic model for the growth rate (Blake & Glazebrook 2003),

$$D(z|w_0, w_a) = \frac{5\Omega_m H_0^2}{2} H(a) \int_0^a \frac{da'}{a'^3 H^3(a')} , \quad (32)$$

$$H^2(a) = H_0^2 \left[\frac{\Omega_m}{a^3} + \Omega_\Lambda \exp\left(3 \int_0^z \frac{1 + w(z')}{1 + z'} dz' \right) \right] , \quad (33)$$

with Ω_m, Ω_Λ and H_0 the dark matter and dark energy densities and the Hubble constant respectively at redshift $z = 0$ and $a \equiv 1/(1 + z)$ the expansion factor. If one had access to the projected dark matter PDF at different redshift slices directly (through weak lensing convergence), one could follow the fiducial experiment described in Codis et al. (2016a). The key idea is to take advantage of the fact that the variance is the only parameter in the dark matter PDF and that its redshift-dependence is given by $\sigma(z) \propto D(z|w)$ where D is given by equation (33).

When applying this idea to galaxy counts, the main difficulty is that disentangling bias parameters and variance (following Uhlemann et al. 2017a) requires both one- and two-point statistics. Those are difficult to extract accurately enough on a thin slice-by-slice basis (such as that shown on Fig. 9), because the number of cylinders is limited by the available cosmic volume and the average number of galaxies per cylinders. In this case, one has to circumvent this difficulty by parametrising the redshift evolution of the bias parameters (a possible parametrisation could be e.g. $b_n(z) = \sum a_{np} z^p$). Then, one could marginalise over some prior values for this parametrisation or use information at different scales (either varying radius or separation of the cylinders) to constrain those parameters jointly with w. To constrain dark energy, one could naturally compute the joint log-likelihood, $\mathcal{L}$ of the measured densities encoded in the halo PDF as

$$\mathcal{L}(\{(\rho_h)_{i,k}\}|w) = \sum_{k=1}^{N_z} \sum_{i=1}^{N_{cyl}(k)} \log \mathcal{P}_{\mathcal{R}}^{cyl}\left((\rho_h)_{i,k}|z_k, w\right), \quad (34)$$

where N_z is the number of redshift slices, $N_{cyl}(k)$ the number of cylinders per slice and $\mathcal{P}((\rho_h)_{i,k}|z_k, w)$ is the theoretical density probability of having halo densities $(\rho_h)_{i,k}$ in a cylinder of given radius at redshift-slice z_k for a cosmological model with dark energy e.o.s parametrized by $w = (w_0, w_a)$ and marginalized over possible values of the bias parametrization. Optimizing $\mathcal{L}$ in equation (34) with respect to w would yield a maximum likelihood estimate for the dark energy equation of state parameters. As mentioned earlier, the key ingredient is to parametrize the redshift bias evolution with a few parameters, so that each new redshift slice does not introduce new unknown parameters. One could also attempt taking correlations between neighbouring cylinders into account. The DE parameters could then be explored for example using an adapting (importance) sampling scheme where the sampling proposal could be a multi-variate Gaussian copula, see e.g. Benabed et al. (2009). An alternative strategy to address the tracer bias problem would be to use a joint count and lensing-in-cell analysis as done in Gruen et al. (2017); Friedrich et al. (2017), possibly applied to tomographic bins.

Implementing the idea to constrain dark energy on the light-cone of HR4 will be the topic of future work (but see Fig 10).

5 CONCLUSION AND OUTLOOK

Building upon recent works on large deviation statistics for cosmic densities in spherical cells, estimators for the one-point PDF of densities in concentric cylinders and for the density-dependence of their angular clustering (cylinder bias) – therefore their two-point statistics – were presented. These statistics were combined with a simple polynomial local bias model (in log-densities) relating dark matter and tracer densities in cylinders to produce estimators for 2D count-in-cell statistics and compensated filters. The validity of the parametrised bias model was established using a parametrisation-independent extraction of the bias function. All these estimators
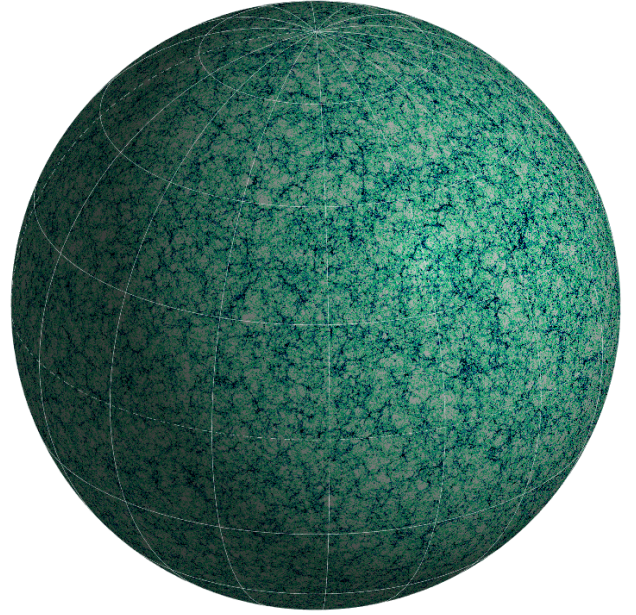

Figure 9. Logarithmic view on the mass weighted count-in-cell of sub-haloes in the HR4 lightcone at redshift $z = 0.36$ in a slice of 60 Mpc/h projected on the sphere using the HealPix scheme (here $\ell_{max} = 512$).
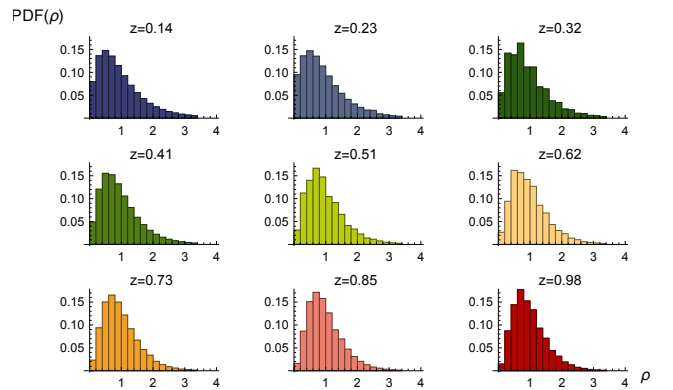

Figure 10. Subhalo counts from the (non-cylindrical) HealPix cells for $\ell_{max} = 256$ also displayed in Fig. 9 at various redshifts as labeled. As expected, the PDFs become more Gaussian at higher redshift. Fitting counts like this following the procedure presented in Fig 7 should allow us to put constraints on the cosmic evolution of σ.

were found to be in excellent agreement with state-of-the-art numerical simulations. The present formalism will allow to probe cosmology using photometric surveys containing billions of galaxies. It also provides the basis for a direct bias independent application to weak lensing mass aperture maps. The PDF for dark matter and biased tracers could also be used to rapidly generate mock top-hat smoothed maps for e.g. weak lensing or intensity mapping.

A possible extension of this work is a validation of the analysis on lightcone slices with virtual galaxies extracted from Horizon Run 4. This could be used to quantify the accuracy at which dark energy parameters can be extracted from 2D count-in-cells following the fiducial dark energy experiment sketched in Sec. 4.5. Another direction is to apply the large deviation statistics formalism to

weak lensing convergence and shear in order to directly extract information from the dark matter density and its relation to biased tracers. The PDF presented in equation (28) could also be used (while relying on a fast particle-mesh code such as Feng et al. 2016, to generate dark matter maps at the $\sim 5\,\mathrm{Mpc}/h$ scale with accurate two-point functions) to rapidly generate mock top-hat smoothed 2D maps for dark haloes which would by construction have the right one and two-point statistics.

Eventually, this formalism should be applied to constrain cosmology using density and 2D counts-in-cells statistics in ongoing or upcoming surveys like DES, Euclid, WFIRST, LSST, KiDs.

References

Agrawal A., Makiya R., Chiang C.-T., Jeong D., Saito S., Komatsu E., 2017, J. Cosmology Astropart. Phys., 10, 003
Bel J., Marinoni C., 2014, A&A, 563, A36
Bel J., et al., 2016, A&A, 588, A51
Benabed K., Cardoso J.-F., Prunet S., Hivon E., 2009, MNRAS, 400, 219
Bernardeau F., 1992, ApJ, 392, 1
Bernardeau F., 1994, A&A, 291, 697
Bernardeau F., 1995, A&A, 301, 309
Bernardeau F., 1996, A&A, 312, 11
Bernardeau F., Reimberg P., 2016, Phys. Rev. D, 94, 063520
Bernardeau F., Valageas P., 2000, A&A, 364, 1
Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1
Bernardeau F., Pichon C., Codis S., 2014a, Phys. Rev. D, 90, 103519
Bernardeau F., Pichon C., Codis S., 2014b, Phys. Rev. D, 90, 103519
Blake C., Glazebrook K., 2003, ApJ, 594, 665
Bouchet F. R., Strauss M. A., Davis M., Fisher K. B., Yahil A., Huchra J. P., 1993, ApJ, 417, 36
Clerkin L., et al., 2017, MNRAS, 466, 1444
Codis S., Pichon C., Bernardeau F., Uhlemann C., Prunet S., 2016a, MNRAS, 460, 1549
Codis S., Bernardeau F., Pichon C., 2016b, MNRAS, 460, 1598
Coles P., Jones B., 1991, MNRAS, 248, 1
DESI Collaboration et al., 2016, preprint, (`arXiv:1611.00036`)
Driver S. P., et al., 2016, MNRAS, 455, 3911
Dubinski J., Kim J., Park C., Humble R., 2004, New Astron., 9, 111
Efstathiou G., 1995, MNRAS, 276, 1425
Feng Y., Chu M.-Y., Seljak U., McDonald P., 2016, MNRAS, 463, 2273
Friedrich O., et al., 2017, preprint, (`arXiv:1710.05162`)
Gruen D., et al., 2016, MNRAS, 455, 3367
Gruen D., et al., 2017, preprint, (`arXiv:1710.05045`)
Guzzo L., et al., 2014, A&A, 566, A108
Hamaus N., Seljak U., Desjacques V., Smith R. E., Baldauf T., 2010, Phys. Rev. D, 82, 043515

Hilbert S., Hartlap J., Schneider P., 2011, A&A, 536, A85
Hurtado-Gil L., Martínez V. J., Arnalte-Mur P., Pons-Bordería M.-J., Pareja-Flores C., Paredes S., 2017, A&A, 601, A40
Jee I., Park C., Kim J., Choi Y.-Y., Kim S. S., 2012, ApJ, 753, 11
Kacprzak T. e. a., 2016, preprint, (`arXiv:1603.05040`)
Kaiser N., 1984, ApJ, 284, L9
Kayo I., Taruya A., Suto Y., 2001, ApJ, 561, 22
Kim J., Park C., 2006, ApJ, 639, 600
Kim J., Park C., L'Huillier B., Hong S. E., 2015, J. Korean Astron. Soc., 48, 213
Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35
L'Huillier B., Park C., Kim J., 2014, New Astron., 30, 79
LSST Dark Energy Science Collaboration 2012, preprint, (`arXiv:1211.0310`)
Laureijs R., et al., 2011, preprint, (`arXiv:1110.3193`)
McConnachie A., Babusiaux C., Balogh M., Driver S., Côté et al. 2016, preprint, (`arXiv:1606.00043`)
Munshi D., Jain B., 2000, MNRAS, 318, 109
Munshi D., Jain B., 2001, MNRAS, 322, 107
Munshi D., Valageas P., Barber A. J., 2004, MNRAS, 350, 77
Munshi D., Coles P., Kilbinger M., 2014, J. Cosmology Astropart. Phys., 4, 004
Neyrinck M. C., Szapudi I., McCullagh N., Szalay A., Falck B., Wang J., 2016, preprint, (`arXiv:1610.06215`)
Pajer E., van der Woude D., 2017, preprint, (`arXiv:1710.01736`)
Reimberg P., Bernardeau F., 2017, preprint, (`arXiv:1708.00252`)
Rykoff E. S., et al., 2016, ApJS, 224, 1
Seljak U., 2009, Physical Review Letters, 102, 021302
Sheth R. K., Tormen G., 2004, MNRAS, 350, 1385
Sigad Y., Branchini E., Dekel A., 2000, ApJ, 540, 62
Spergel D., et al., 2013, preprint, (`arXiv:1305.5422`)
Szapudi I., Pan J., 2004, ApJ, 602, 26
Takada M., et al., 2014, PASJ, 66, R1
Taruya A., Takada M., Hamana T., Kayo I., Futamase T., 2002, ApJ, 571, 638
Uhlemann C., Codis S., Pichon C., Bernardeau F., Reimberg P., 2016, MNRAS, 460, 1529
Uhlemann C., et al., 2017a, preprint, (`arXiv:1705.08901`)
Uhlemann C., Pajer E., Pichon C., Nishimichi T., Codis S., Bernardeau F., 2017b, preprint, (`arXiv:1708.02206`)
Uhlemann C., Codis S., Kim J., Pichon C., Bernardeau F., Pogosyan D., Park C., L'Huillier B., 2017c, MNRAS, 466, 2067
Valageas P., 2002, A&A, 382, 412
Vale A., Ostriker J. P., 2004, MNRAS, 353, 189
White M., 2016, preprint, (`arXiv:1609.08632`)
White M., Padmanabhan N., 2009, MNRAS, 395, 2381
Wolk M., McCracken H. J., Colombi S., Fry J. N., Kilbinger M., Hudelot P., Mellier Y., Ilbert O., 2013, MNRAS, 435, 2
Xavier H. S., Abdalla F. B., Joachimi B., 2016, MNRAS, 459, 3693
Yang A., Saslaw W. C., 2011, ApJ, 729, 123
Yoshisato A., Matsubara T., Morikawa M., 1998, ApJ, 498, 48
Yoshisato A., Morikawa M., Gouda N., Mouri H., 2006, ApJ, 637, 555
de Jong J. T. A., et al., 2013, The Messenger, 154, 44

APPENDIX A: DEPTH, COLLAPSE AND CUMULANTS

A1 Leading order cumulants depending on shape

While a closed-form spherical collapse solution is known for densities in spheres and densities in infinitely long cylinders (that correspond to two-dimensional disks on the sky), no solution for finite cylinders is known yet. To get an intuition about the finite-size effects related to cylinders, let us perturbatively compute the cumulants at leading order. Once the second order kernel from standard (Eulerian) perturbation theory is written as (see e.g. Bernardeau
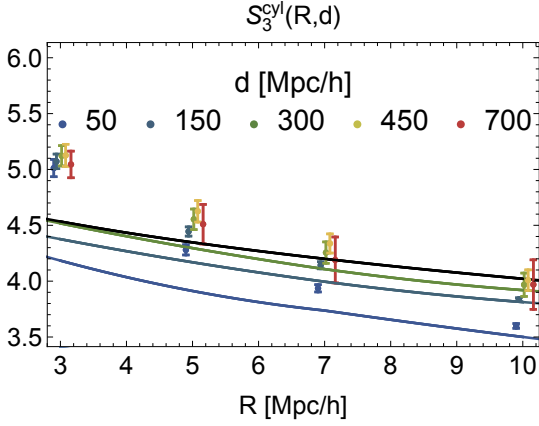
Figure A1. Skewness of the density in cylindrical cells with length $d = 50, 150, 300, 450, 700$ Mpc/h (lower blue to higher red) and radii $R = 3, 5, 7, 10$ Mpc/h at redshift $z = 0.7$ from the HR4 measurements (data points with error bars) and the tree-order skewness computed from equation (A2) for $d = 50, 150, 300$ as well as the infinitely long cylinder limit according to equation (A5) (black line).

et al. 2002)

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{1}{2} \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)(k_1^2 + k_2^2)}{k_1^2 k_2^2} + \frac{2}{7} \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2}, \quad (A1)$$

the leading-order skewness for a smoothing kernel of arbitrary shape can be computed as

$$\kappa_3(\mathcal{R}) = 3 \langle (\delta^{(1)})^2 \delta^{(2)} \rangle \quad (A2)$$

$$= 6 \int \frac{\mathrm{d}^3 k_1}{(2\pi)^3} \int \frac{\mathrm{d}^3 k_2}{(2\pi)^3} F_2(\mathbf{k}_1, \mathbf{k}_2) P_L(k_1) P_L(k_2) \quad (A3)$$
$$W(\mathbf{k}_1, \mathcal{R}) W(\mathbf{k}_2, \mathcal{R}) W(-(\mathbf{k}_1 + \mathbf{k}_2), \mathcal{R}).$$

The reduced skewness $S_3 = \kappa_3 / \sigma^4$ is obtained as the ratio of the skewness and the square of the variance which can be evaluated in linear theory according to equation (3). The result of a numerical integration for the tree-order skewness with a finite cylindrical filter is shown in Fig. A1. *Analytical computation of cumulants from symmetry.* When computing the cumulants for unsmoothed fields (hence $W = 1$), one has to compute an angular average of the F_2 kernel which gives for spherical symmetry and in an EdS universe the well-known result $S_3^{3D} = 34/7$. Using the approximate power law for spherical collapse from equation (6) can be matched to $S_3 = 3(1 + 1/\nu)$ giving $\nu_{3D} = 21/13$. In analogy, one obtains $S_3^{2D} = 36/7$ for the disk symmetry which holds for an infinite cylinder where on can switch to cylindrical coordinates where $\mathbf{k} = (\mathbf{k}_\perp, k_\parallel)$ and $0 \simeq k_\parallel \ll k_\perp$ and hence $\nu_{2D} = 7/5 = 1.4$. When including the top-hat filtering in real space, the effects of smoothing for a power law initial spectrum $P(k) \propto k^n$ are given in Bernardeau (1994, 1995) and read

$$S_3^{3D} = \frac{34}{7} + \gamma_1 \simeq \frac{34}{7} - (n_{3D} + 3), \quad \gamma_1 = \frac{\mathrm{d} \log \sigma_{\mathrm{sph}}^2}{\mathrm{d} \log R}, \quad (A4)$$

$$S_3^{2D} \simeq \frac{36}{7} - \frac{3}{2}(n_{2D} + 2) \quad (A5)$$

There is a slight Ω_m dependence in the constant term which is bracketed by $21/4 > S_3^{2D} > 36/7$ for $\Omega_m \in [0, 1]$.

For finite-size cylinders, one needs to take the component of the wave-vector parallel to the line of sight into account such that one cannot obtain a direct analytical expression but has to numerically evaluate equation (A2). Fig. A1 shows the reduced skew-
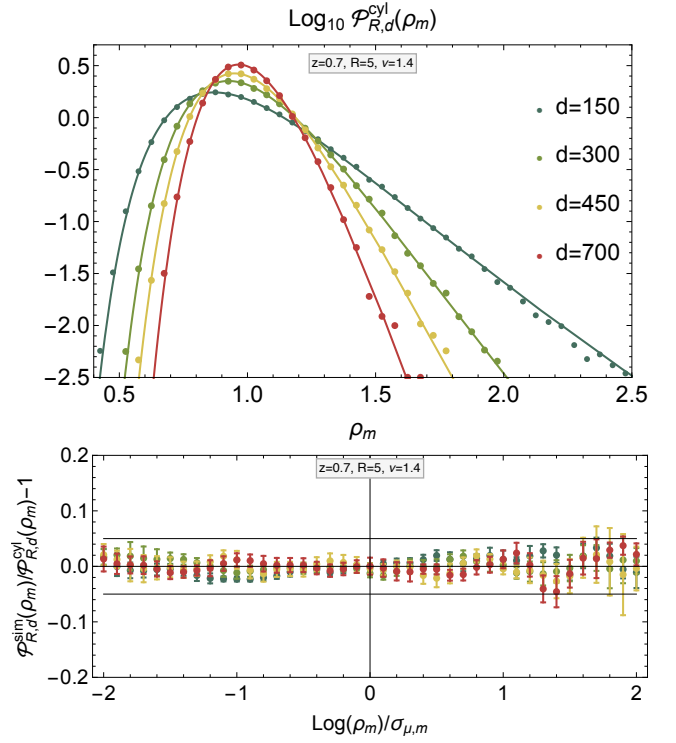

Figure A2. *(Upper panel)* PDF of the projected density in cylindrical cells of length $d = 150, 300, 450, 700$ Mpc/h and radius $R = 5$ Mpc/h at redshift $z = 0.7$. Shown is a comparison of HR4 measurements (data points) and the prediction computed with the cylindrical filter (1) with mass-splitting for long cylinders (7) together with the spherical collapse parametrized by $\nu = 1.4$ (solid lines). *(Lower panel)* Residuals of the theoretical prediction against the HR4 measurements with error bars.

ness measured from the HR4 simulation. As expected by cylindrical collapse, the reduced skewness becomes independent of depth if the axis ratio is large enough (typically for $2R/d \gtrsim 1/10$). Note that the nonlinear variance $\sigma^2(R, d) \propto 1/d$, see equation (31) in Bernardeau & Valageas (2000), such that the product $\sigma^2(R, d) \times d$ remains constant to a very good accuracy.

A2   Large deviation statistics with varying cylinder depth

Fig. A2 compares the theoretical prediction for infinitely long cylinders, hence effectively 2D-spherical densities when fixing the radius R and varying the depth d. The limit of essentially infinitely long cylinders is achieved at 10% accuracy in the PDF already for a depth to radius ratio of about $(2R)/d \simeq 1/5$ and quickly improves with increasing depth. This effect can be also seen in Fig. A1 which compares the measured skewness to the perturbation theory result for the corresponding hierarchical clustering ratio.

A3   Scale-dependence of the linear covariance

The scale-dependence of the linear variance from equation (3) for varying radius and depth of the cylinder is shown in Fig. A3 and compared to the variance computed with a spherical filter of equal volume. For concentric cylinders of different radii R_i but identical depth d, for the sake of simplicity, the covariance matrix from equation (3) is parametrized in analogy to a power-law initial spectrum
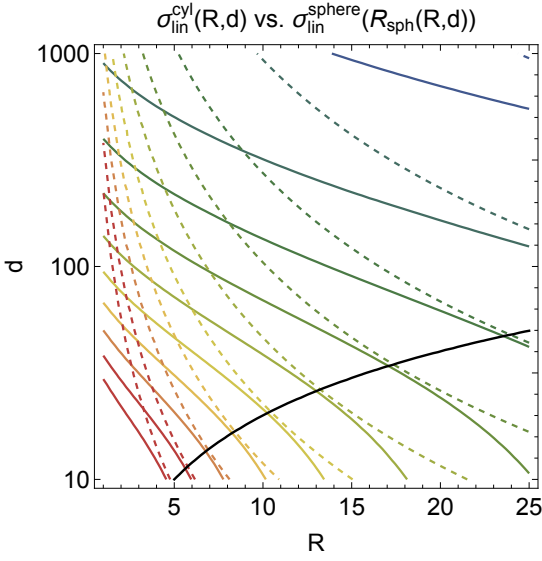
Figure A3. Scale dependence of the linear variance σ induced by a cylindrical filter with radius R and depth d, both in Mpc/h. The contours indicate variances from 0.1 (blue) to 1.0 (red) in steps of 0.1. The cylindrical filter (solid lines) is compared to a volume-equivalent spherical filter with $R_{\rm sph} = (3/4R^2 d)^{1/3}$ (dashed lines). As expected, they are approximately equal for an axis ratio $2R/d = 1$ (black line).

with spectral index $n_{\rm 2D}(R_p) \simeq n_{\rm cyl}(R_p, d)$ by

$$\frac{\sigma^2(R_i, R_i, d)}{\sigma^2(R_p, d)} = \left(\frac{R_i}{R_p}\right)^{-n_{\rm cyl}(R_p, d)-2} , \qquad \text{(A6a)}$$

$$\frac{\sigma^2(R_i, R_{j>i}, d)}{\sigma^2(R_p, d)} = \mathcal{G}\left(\frac{R_i}{R_p}, \frac{R_j}{R_p}, n_{\rm cyl}(R_p, d)\right) , \qquad \text{(A6b)}$$

where the 2D-spectral index is obtained as

$$n_{\rm cyl}(R, d) = -2 - \frac{d \log \sigma^2_{\rm cyl}(R, d)}{d \log R} . \qquad \text{(A7)}$$

and

$$\mathcal{G}(x, y, n) = \frac{\int {\rm d}^3 k \, k^n W_{\rm 2D}(kx) W_{\rm 2D}(ky)}{\int {\rm d}^3 k \, k^n W_{\rm 2D}(kR_p) W_{\rm 2D}(kR_p)} ,$$

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

with $\alpha = 1 - n$. The key parameter in the prediction of the PDF is the value of the variance at the pivot radius R_p and depth d which is measured in the simulation and use as an input to our theoretical model.

A4 Joint PDF of dark matter around halo density

Figs. A4 and A5 show the joint PDF of matter density in rings around halo density in cylinders and the conditional PDF of matter density in rings given tracer density. The curves of the theoretical prediction for the conditional look almost identical to Fig. 4, which is due to the fact that the best-fit bias model leads to a similar split $\rho_m(\rho_h = 1) \simeq 1.02$ between over- and underdensity
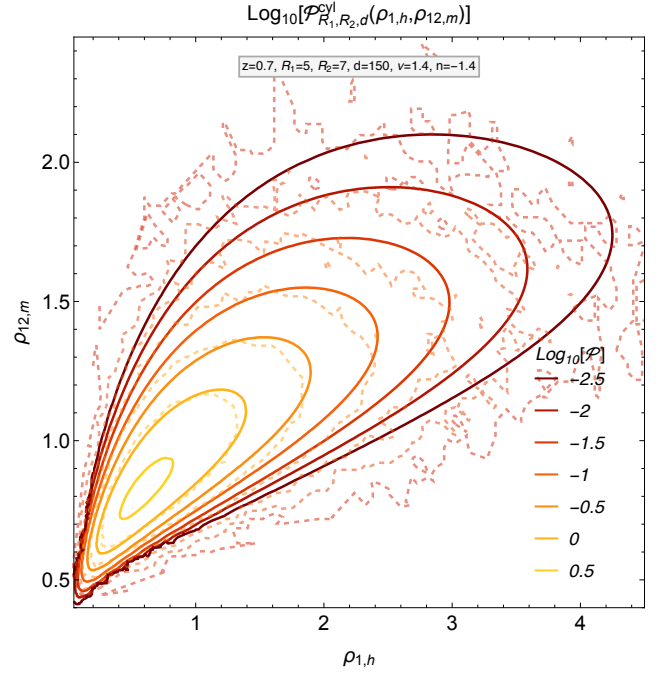

Figure A4. Joint PDF of the projected halo densities in a cylinder of depth $d = 150$ Mpc/h and radius $R_1 = 5$ Mpc/h and dark matter density in a surrounding cylindrical ring with $R_2 = 7$ Mpc/h at redshift $z = 0.7$ from the saddle point approximation with $\nu = 1.4$ and the disk filter for a power-law spectrum with $n = -1.4$ and the best fit bias parameters from Table 2 (solid lines) compared to the HR4 measurements (dashed lines).

for both matter and haloes. A possible cause of the residuals seen in the PDFs conditional-on-halo-density is the scatter around the mean bias relation which should increase the probability of unlikely events such as over-/underdense matter rings around under-/overdense halo cores that are not captured in a one-to-one bias model.

APPENDIX B: LOGNORMAL RECONSTRUCTION

One widespread phenomenological ansatz for the one-point PDF of either dark matter or its tracers is the lognormal distribution (Coles & Jones 1991). Assuming a lognormal density field, the one-point PDF is fully determined by the variance of the log-density and hence one can attempt a reconstruction of the density PDF using the functional form

$$\mathcal{P}_{\rm LN}(\rho \,|\, \sigma_\mu) = \frac{1}{\sqrt{2\pi}\sigma_\mu} \frac{1}{\rho} \exp\left[-\frac{(\log \rho + \sigma_\mu^2/2)^2}{2\sigma_\mu^2}\right] . \qquad \text{(B1)}$$

Fig. B1 shows a comparison of the lognormal model for the PDF of dark matter in cylinders with fixed depth and varying radius. When plotted as ratio between the measured PDF and the lognormal reconstruction, a clear residual skewness is visible which leads to deviations at the 5%-20% level for all densities of interest. Overall, the accuracy improves with increasing cylinder radius and cylinder depth.
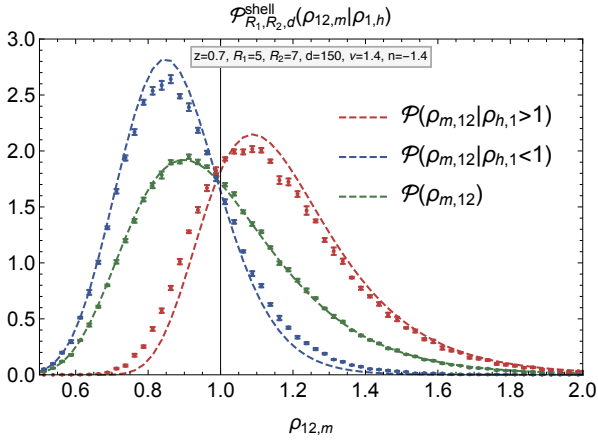
Figure A5. Marginalised PDF of the matter densities in cylindrical rings of depth $d = 150$ Mpc/h and radii $R_1 = 5$ Mpc/h and $R_2 = 7$ Mpc/h at redshift $z = 0.7$ from the saddle point approximation with $\nu = 1.4$ and the disk filter for a power-law spectrum with $n = -1.4$ together with the best fit bias model from Table 2 for arbitrary inner tracer densities (dashed, green line), overdensities in tracers (dashed red line) and underdensities in tracers (dashed blue line) compared to the HR4 measurements (data points). Comparing the theoretical prediction to the PDF given over- or underdensity in dark matter in Fig. 4, they look virtually identical.
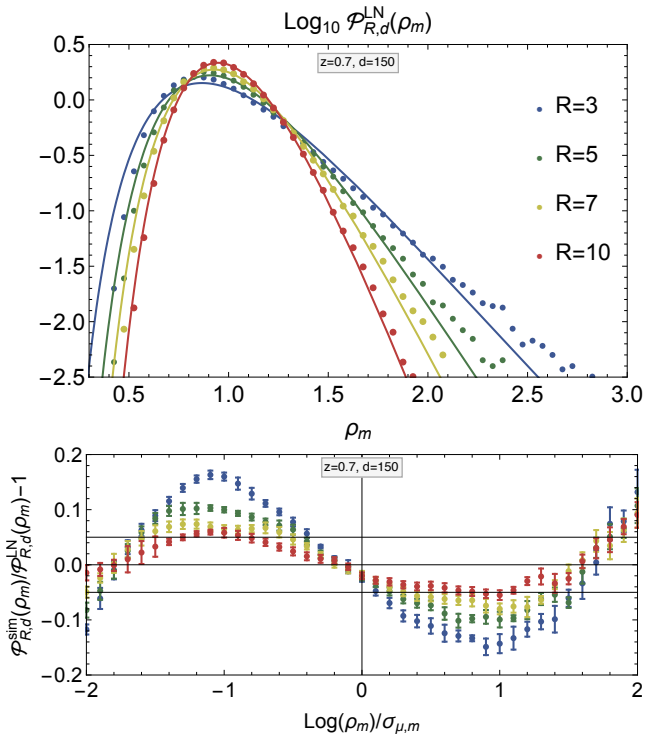

Figure B1. Comparison of lognormal PDF model from equation (B1) with measured log-variance as given in Table 1 for fixed depth $d = 150$ Mpc/h and varying radius against the measured PDF for dark matter. This plot is the analogue of Fig. 1 for the lognormal reconstruction.