

# Two is better than one: joint statistics of density and velocity in concentric spheres as a cosmological probe

C. Uhlemann,<sup>1</sup> S. Codis,<sup>2</sup>★ O. Hahn,<sup>3</sup> C. Pichon<sup>4,5</sup> and F. Bernardeau<sup>4,6</sup>

<sup>1</sup>*Institute for Theoretical Physics, Utrecht University, Princetonplein 5, NL-3584 CC Utrecht, the Netherlands*

<sup>2</sup>*Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto ON M5S 3H8, Canada*

<sup>3</sup>*Laboratoire Lagrange, Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Blvd de l'Observatoire, F-06304 Nice, France*

<sup>4</sup>*Sorbonne Universités, UPMC Paris 6 & CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98 bis bd Arago, F-75014 Paris, France*

<sup>5</sup>*Korea Institute of Advanced Studies (KIAS) 85 Hoegiro, Dongdaemun-gu, Seoul 02455, Republic of Korea*

<sup>6</sup>*CNRS & CEA, UMR 3681, Institut de Physique Théorique, F-91191 Gif-sur-Yvette, France*

Accepted 2017 April 20. Received 2017 April 19; in original form 2016 November 30

## ABSTRACT

The analytical formalism to obtain the probability distribution functions (PDFs) of spherically averaged cosmic densities and velocity divergences in the mildly non-linear regime is presented. A large-deviation principle is applied to those cosmic fields assuming their most likely dynamics in spheres is set by the spherical collapse model. We validate our analytical results using state-of-the-art dark matter simulations with a phase-space resolved velocity field finding a 2 per cent level agreement for a wide range of velocity divergences and densities in the mildly non-linear regime ( $\sim 10 \text{ Mpc } h^{-1}$  at redshift zero), usually inaccessible to perturbation theory. From the joint PDF of densities and velocity divergences measured in two concentric spheres, we extract with the same accuracy velocity profiles and conditional velocity PDF subject to a given over/underdensity that are of interest to understand the non-linear evolution of velocity flows. Both PDFs are used to build a simple but accurate maximum likelihood estimator for the redshift evolution of the variance of both the density and velocity divergence fields, which have smaller relative errors than their sample variances when non-linearities appear. Given the dependence of the velocity divergence on the growth rate, there is a significant gain in using the full knowledge of both PDFs to derive constraints on the equation of state-of-dark energy. Thanks to the insensitivity of the velocity divergence to bias, its PDF can be used to obtain unbiased constraints on the growth of structures ( $\sigma_8, f$ ) or it can be combined with the galaxy density PDF to extract bias parameters.

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

## 1 INTRODUCTION

The kinematics of the large-scale structure of the Universe should allow astronomers to put very tight constraints on cosmological models. Deep and wide spectroscopic surveys, like the recently proposed MSE<sup>1</sup> (McConnachie et al. 2016), will soon allow us to study fundamental physics from the distribution of galaxies, in particular the origin of cosmic acceleration. However to reach per cent precision on the equation of state-of-dark energy, astronomers are facing various challenges. In particular, luminous matter is not a fair tracer of the total amount of matter in the Universe (Kaiser 1984) and is subject to systematic effects like redshift space distortions.

Unlike the density field, peculiar velocities are much less affected by galaxy biasing (Elia et al. 2011; Nusser, Branchini & Davis 2012; Baldauf, Desjacques & Seljak 2015) and therefore represent a more direct probe of the dynamics. Their power spectrum (Jennings 2012) and skewness (Bernardeau et al. 1995) have therefore been proposed as unbiased cosmological probes.

However, in order to extract as much non-linear information as possible, it has become necessary to investigate alternative estimators to the standard hierarchy of  $N$ -point correlation functions of both densities and velocities. In this context, Bernardeau, Pichon & Codis (2014), Bernardeau, Codis & Pichon (2015) and Codis et al. (2016a) have shown how the full 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 standard statistics. Indeed, the zero

\* E-mail: [codis@cita.utoronto.ca](mailto:codis@cita.utoronto.ca)

<sup>1</sup> <http://mse.cftht.hawaii.edu/project>

variance limit of the cumulant generating functions (CGFs) yields estimates of the joint probability distribution function (PDF) that seems to match simulations in the regime of variances of the order of unity (Balian & Schaeffer 1989; Bernardeau 1992; Juszkiewicz, Bouchet & Colombi 1993; Valageas 2002; Bernardeau, Pichon & Codis 2014; Bernardeau, Codis & Pichon 2015). Following this work, Uhlemann et al. (2016) recently presented fully analytic expressions for the PDF of the density field in spheres using a saddle-point approximation applied to the logarithm of the density.

This formalism can be extended to also include the velocity divergence and obtain the joint statistics of the density and velocity field while relying on large-deviation statistics and spherical collapse. This is of interest since velocities probe more directly the total gravitational field and are more sensitive to the growth rate, unlike e.g. galaxy count in cells that are biased with respect to the underlying dark matter density field. The purpose of this paper is therefore to generalize the above-cited works to the joint statistics of the density field  $\rho$  and the velocity divergence  $\theta$  in multiple concentric spheres. Conditional statistics of velocities in over- and underdense environments can be deduced and are of particular interest in the context of void dynamics. We will also use these statistics to build a maximum likelihood toy model for the estimation of cosmic parameters.

The PDFs of the density and peculiar velocity divergence in a sphere have been obtained and compared to simulations first in Bernardeau & van de Weygaert (1996) and Juszkiewicz et al. (1995) using perturbation theory together with an Edgeworth expansion. Scoccimarro & Frieman (1996b) investigated when loop corrections for unsmoothed fields start to dominate over tree-level contributions (hence signalling a breakdown of perturbation theory), finding that, while this happens for a variance  $\sigma \simeq 0.5$  for the density, it does not occur until  $\sigma \simeq 1$  for the velocity divergence. The velocity divergence field and its cumulants have been studied for different local collapse models in Scoccimarro & Frieman (1996b), Fosalba & Gaztanaga (1998), Ohta, Kayo & Taruya (2003) where the tidal contributions to the reduced cumulants have been calculated and observed to cancel out for smoothed fields, thus rendering the spherical collapse model successful.

Here, we refine the theoretical formalism for obtaining the PDF of the velocity divergence field and use state-of-the-art numerical simulations that are able to accurately extract the velocity field. Indeed, while peculiar velocities are notoriously difficult to measure in real data, recent progress in estimating the divergence of the velocity (Abel, Hahn & Kaehler 2012; Shandarin, Habib & Heitmann 2012; Hahn, Angulo & Abel 2015) in simulations now allow us to accurately determine these joint PDF and compare them to analytical predictions. Furthermore, we extend earlier theoretical works by Bernardeau (1994) to not only one but multiple concentric spheres and we revisit the non-linearity and stochasticity in the relation of the density and velocity divergence that has been investigated in Bernardeau et al. (1999).

The outline of the paper is the following. Section 2 shortly reviews the general formalism to obtain the PDFs (Section 3) for the density and the velocity divergence in concentric spheres based on spherical collapse dynamics and large-deviation statistics. Section 4 compares the predictions from the fully analytical saddle-point approximation to state-of-the-art numerical simulations with a phase-space resolved velocity field. Section 5 presents possible applications of this framework for constraining dark energy and structure formation parameters. It also discusses briefly the potential to infer halo bias from joint density and velocity measurements and the connection to measurements of peculiar velocities that pose

the main observational challenge to applying our formalism. Finally, Section 6 wraps up and gives an outlook. Appendix A presents the code we publicly release with this paper, while Appendix B provides a deeper theoretical background to our formalism. Appendix C discusses the scatter of the density–velocity relation, while the multiscale cumulants of the density and velocity divergence are derived in Appendix D.

## 2 DENSITY AND VELOCITY IN SPHERES

This paper aims to predict the one-point statistics of the velocity divergence from first principles using the spherical collapse dynamics in the large-deviation regime following the seminal work of Bernardeau (1992). In this section, we will first introduce the spherical collapse model that is at the heart of the construction and allows us to connect both the spherically averaged density and the velocity divergence to the initial density field. Then, we will shortly explain the algorithm of the simulation that we use to assess the relationship between density and velocity divergence numerically. To assess the validity of the relation between density and velocity divergence in a sphere, we will determine their joint PDF from the simulation and discuss their mean relation and dispersion in the context of the spherical collapse model and perturbation theory.

### 2.1 Spherical collapse dynamics

The spherical collapse dynamics relates the final density  $\rho$  (normalized to unit mean, hence  $\rho = 1 + \delta$ ) and velocity divergence  $\theta = \nabla \cdot \mathbf{u}/H_0$ , with  $H_0$  the present-day Hubble constant, in a sphere of size  $R$  to the initial density contrast  $\tau$  in a sphere of size  $r = R\rho^{1/3}$  (mass conservation). This mapping in general cannot be written analytically. Parametric solutions can be found in some regimes but, here for simplicity, we rely on a very accurate approximate expression given by

$$\rho_{\text{SC}}(\tau) = \left(1 - \frac{\tau}{\nu}\right)^{-\nu}, \quad (1a)$$

$$\theta_{\text{SC}}(\tau) = -f(\Omega) \frac{d \log \rho(\tau)}{d \log \tau} = -f(\Omega) \tau (1 - \tau/\nu)^{-1}, \quad (1b)$$

with the growth rate  $f(\Omega) = d \log D / d \log a \approx \Omega^{0.6}$  given in terms of the growth factor  $D$  that is related to the matter density parameter  $\Omega$ .

From equation (1b), it is clear that the velocity divergence  $\theta_{\text{SC}}(\tau)$  changes more gently with the initial density  $\tau$  than the density  $\rho_{\text{SC}}(\tau)$ , which makes it an interesting target for further investigation. In the linear regime, the velocity divergence is proportional to the density contrast,  $\theta = -f(\Omega)\delta$ , which is why we will use the rescaled velocity divergence  $\tilde{\theta} = -\theta/f(\Omega)$  as a variable instead of  $\theta$  in the following.

The joint spherical collapse mapping from equation (1) can be employed to obtain a relation between the non-linearly evolved velocity divergence and the density in a sphere, giving the relation (see e.g. Bernardeau & van de Weygaert 1996; Bernardeau et al. 1999; Nadkarni-Ghosh 2013)

$$\tilde{\theta}_{\text{SC}}(\rho) = \nu (\rho^{1/\nu} - 1), \quad (2)$$

which has been first given in Bernardeau (1992) for  $\nu = 3/2$ . Note that since the density is manifestly positive  $\rho > 0$  we have  $\tilde{\theta} > \tilde{\theta}_{\text{min}} = -\nu \simeq -1.5$ . In what follows, we will use  $\nu = 21/13$  instead of  $3/2$  because it gives both a better match to the tree-order skewness of the density field and the relation between the density and velocity divergence (see also Uhlemann et al. 2016).

## 2.2 Specifications and algorithm of the simulation

The simulation used in this work was presented in Hahn et al. (2015): it has a box size  $1 \text{ Gpc } h^{-1}$ , while the number of particles is  $1024^3$  particles. The gravitational evolution between redshift  $z = 100$  and  $z = 0$  has been performed using the tree-PM code L-GADGET 3 from Angulo et al. (2012) using a  $2048^3$  mesh for the PM force and a force softening of  $35 \text{ kpc } h^{-1}$ .

When estimating densities, we apply a simple cloud-in-cell (CIC) deposit of all particles to a grid of  $1024^3$  cells, followed by a convolution with a spherical top-hat kernel (cf. equation 17). We do not deconvolve with the CIC kernel since for our measurements the top-hat is always well resolved at this resolution.

Following Hahn et al. (2015), we measure the velocity field properties in this  $N$ -body simulation based on a tessellation of the dark matter sheet (Abel et al. 2012; Shandarin et al. 2012). In this approach, since dark matter can be assumed to be perfectly cold, the unperturbed particle lattice in Lagrangian space can be decomposed into a tetrahedral covering. The evolution of the tetrahedra as dictated by the motion of their vertex particles then gives information about the dark matter flow everywhere in space. The single-stream density  $\rho_s$ , e.g. is just the inverse of the volume of each tetrahedron. In multi-stream regions, after collapse, several tetrahedra will overlap a given point in Eulerian space so that the total density is given by the sum over all streams overlapping that point. The mean velocity field, in which we are interested in this paper, is then given by

$$\langle \mathbf{v} \rangle(\mathbf{x}) = \frac{\sum_{s \in S(\mathbf{x})} \mathbf{v}_s(\mathbf{x}) \rho_s(\mathbf{x})}{\sum_{s \in S(\mathbf{x})} \rho_s(\mathbf{x})}, \quad (3)$$

where  $S(\mathbf{x})$  is the set of all streams  $s$  (i.e. tetrahedra) that contain point  $\mathbf{x}$  and a subscript  $s$  indicates the value of a field at  $\mathbf{x}$  on a given sheet. The field  $\langle \mathbf{v} \rangle$  can be sampled on a regular mesh, of resolution  $1024^3$  in our case, for further analysis by calculating all intersections. The new method for resolving velocity fields has been shown to remove the shot noise that was present in previous approaches, e.g. when comparing with fig. 13 in Hahn et al. (2015).

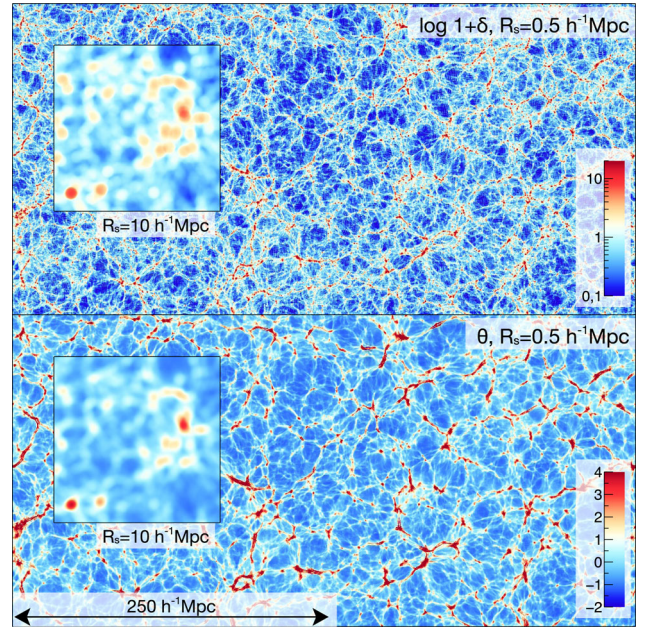
In Fig. 1, we show a slice through the simulation for the density field and the velocity divergence. The colour map of the velocity divergence has been inverted with respect to that of the density field to allow for a more easy comparison in terms of the anticorrelation between the two fields in linear theory. For a more detailed description of this approach, we refer to Hahn et al. (2015). Since we are interested in the velocity divergence smoothed on some scale, we next turn to Fourier transforms to calculate

$$\tilde{\theta} = i \mathbf{k} \cdot \langle \mathbf{v} \rangle \tilde{W}_R, \quad (4)$$

where a tilde here indicates a Fourier transformed field, and  $W_R$  is a spherical top-hat in coordinate space of width  $R$  (see equation 17).

## 2.3 The mean velocity divergence and density relations

This paper relies on the fact that the velocity divergence on the one hand and the density (both filtered with a spherical top-hat filter) on the other hand are closely related to each other, so that the mean trend of their relationship can be captured by constrained averages, which encode the mean velocity divergence given a certain density and vice versa. Let us review possible ways to obtain these mean relationships and assess the amplitude of the scatter around this mean, using a state-of-the-art simulation that resolves the velocity field well.



**Figure 1.** Slice of dimension  $250 \times 500 \text{ (Mpc}/h)^2$  with a projection depth of  $15.6 \text{ Mpc } h^{-1}$  through the fields of the log-density (upper panel) and the velocity divergence (lower panel) and the smoothed versions that are under investigation here (insets) of the simulation from Hahn et al. (2015).

In the linear regime the (scaled) velocity divergence is proportional to the density contrast:  $\tilde{\theta} = \delta$ . In the non-linear regime, this linear relationship is modified through the coupling of density and velocity that also induces a scatter in the linear one-to-one relation. As long as the departure from the linear regime is small, generic features can be inferred from perturbation theory (Bernardeau 1992), which yields an expansion of the mean density given the velocity divergence (and vice versa) in terms of the fields themselves

$$\langle \delta \rangle_{\tilde{\theta}} = a_1 \langle \sigma_{\tilde{\theta}}^2 \rangle \tilde{\theta} + a_2 \langle \tilde{\theta}^2 - \sigma_{\tilde{\theta}}^2 \rangle + a_3 \tilde{\theta}^3. \quad (5)$$

To have a consistent expansion up to third order,  $a_1 = 1 + \mathcal{O}(\sigma^2)$  has to be calculated at next-to-leading order, whereas the other coefficients can be computed at their leading order in perturbation theory. The explicit expressions for the coefficients  $a_{1,2,3}$  in terms of higher order cumulants of the density and velocity field as computed in Bernardeau (1994) are given in Bernardeau (1992) and Bernardeau et al. (1999). More generally, the coefficients  $a_n$  were also derived by Chodorowski & Lokas (1997) and Chodorowski et al. (1998), respectively, and their values were explicitly calculated for a Gaussian window function. For a top-hat filter, the values of the higher coefficients  $a_{n \geq 2}$  turn out to be independent of the spectral index of the initial power spectrum and to be very close to those obtained from Taylor-expanding equation (2).

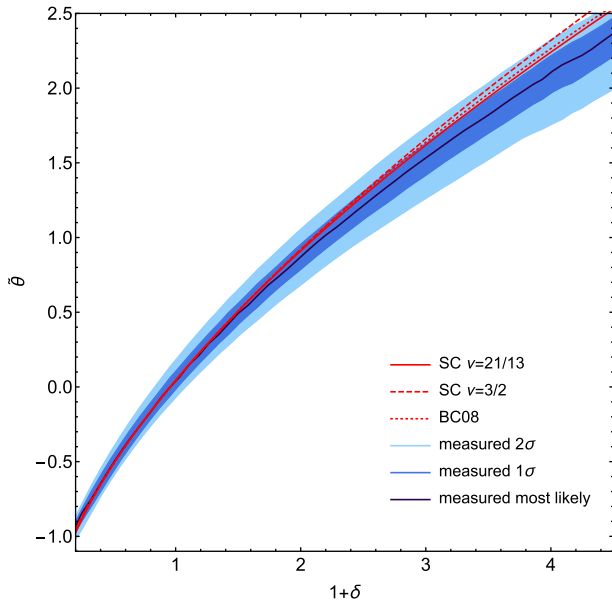
Bilicki & Chodorowski (2008) pointed out that deviations from the mean law found from spherical collapse are expected for larger density contrasts and provided the following fitting formula:

$$\tilde{\theta}_{\text{SC}}(\rho) \stackrel{\rho \geq 1}{\approx} 3 [\rho^{1/2} - \rho^{1/6}], \quad (6a)$$

$$\tilde{\theta}_{\text{SC}}(\rho) \stackrel{\rho \leq 1}{\approx} (\rho - 1) + (1 + \tilde{\theta}_{\text{min}}) [\rho \ln \rho - (\rho - 1)], \quad (6b)$$

where the minimum value was parametrized in terms of  $\Omega_m$  as  $\tilde{\theta}_{\text{min}} = -1 - 0.5 \Omega_m^{0.12 - 0.06 \Omega_m}$ . This formula agrees at the per cent level with the spherical collapse prediction on equation (2) for moderate-density contrast  $\rho \in [0.1, 10]$  that is of interest here,





**Figure 2.** One-sigma (dark blue) and two-sigma (light blue) contours of the conditional PDF of the velocity divergence  $\tilde{\theta}$  given the density  $\rho$  in a sphere of identical radius  $R = 15 \text{ Mpc } h^{-1}$  as measured in the simulation. The most likely value of  $\tilde{\theta}$  given  $\rho$  is shown with the solid black line. Those measurements are compared to the one-to-one mapping of the spherical collapse dynamics defined in equation (2) for different values for the parameter  $\nu = 3/2$  (dashed red) and  $21/13$  (solid red) and to the fit from equation (6) (dotted red). Note that for any fitting functions displayed in this figure, we subtract the mean so that  $\langle \tilde{\theta} \rangle$  is always zero by construction. Note also that instead of plotting the most likely value of  $\tilde{\theta}$  given  $\rho$ , we could have displayed the mean value of  $\tilde{\theta}$  given  $\rho$  but the difference is negligible.

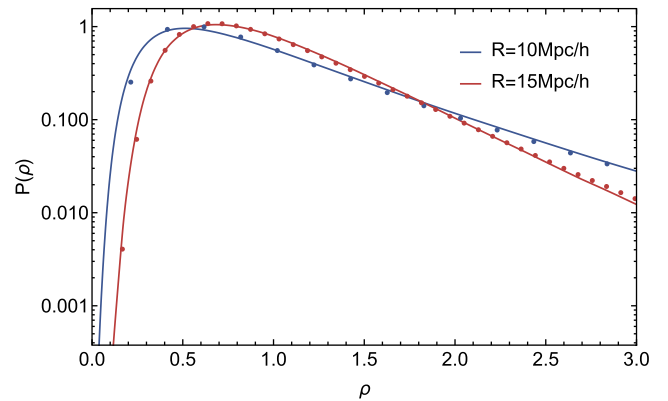
but shows about 10 per cent deviations for strongly under- or over regions  $\rho \in [0.01, 0.1]$  or  $\rho \in [10, 100]$ . The validity of the spherical collapse mapping to connect the velocity divergence and the density is demonstrated in Fig. 2. This figure shows that the approximate expression of the spherical collapse given by equation (2) provides a very good fit to the mean relation between density and velocity divergence for moderate-density contrasts and velocity divergences that is the main focus of this paper. For a description of the (very slight) dark energy dependence of the density–velocity relation, see Nadkarni-Ghosh (2013) where it has also been shown to stay weak even in the non-linear regime.

Using this mapping, the PDF of the velocity divergence can be used to reconstruct the PDF of the density (and vice versa) as illustrated in Fig. 3. Here, we simply map the measured PDF of  $\tilde{\theta}$  to the PDF of the density field using equation (2) and subtracting the mean. In practice, it means that we consider the following change of variable

$$\tilde{\theta} \rightarrow \rho_{\text{SC}}(\tilde{\theta}) = 1 + \left(1 + \frac{\tilde{\theta}}{\nu}\right)^{\nu} - \left\langle \left(1 + \frac{\tilde{\theta}}{\nu}\right)^{\nu} \right\rangle, \quad (7)$$

$$\mathcal{P}_{\tilde{\theta}}(\tilde{\theta}) \rightarrow \mathcal{P}_{\rho}(\rho) = (\rho_{\text{SC}}^{-1})'(\rho) \mathcal{P}_{\tilde{\theta}}(\rho_{\text{SC}}^{-1}(\rho)), \quad (8)$$

where the mean  $\langle (1 + \tilde{\theta}/\nu)^{\nu} \rangle$  is estimated from the measured PDF of the velocity divergence. This estimate of the density PDF is successfully compared to the measured PDF of the density field in Fig. 3, which shows that the spherical collapse dynamics can be used to relate the PDF of the density and velocity divergence field for a broad range of field values and level of non-linearities. The rest of the paper is devoted to properly derive the velocity PDF from first principles using the dynamics of the spherical collapse in the



**Figure 3.** Comparison of the density PDF from a direct measurement of the density (dots) and the reconstruction from the normalized velocity divergence  $\tilde{\theta}$  using the spherical collapse mapping equation (2) with  $\nu = 21/13$  (solid line) for spheres with radii  $R = 10 \text{ Mpc } h^{-1}$  (blue) and  $15 \text{ Mpc } h^{-1}$  (yellow). Note that we also subtract the mean in equation (2). The accuracy of the construction in the range of densities from 0.5 to 3 is around 7 per cent.

large-deviation regime, not only for one but also for two concentric spheres.

For a discussion of the dispersion around the mean relation between velocity divergence and density, we refer the reader to Appendix C.

### 3 CONSTRUCTION OF THE PDFs

Extending upon Bernardeau (1994) and Uhlemann et al. (2016), let us now present the general formalism of how the joint PDF of densities and velocity divergences in concentric spheres can be obtained using large-deviation statistics and spherical collapse dynamics, relying on a logarithmic remapping together with a saddle approximation. As described in Section 2, the spherical collapse model provides us with a direct mapping between densities and velocity divergences that will be used to determine the PDF of the velocity divergence using the predictions for the density PDF.

#### 3.1 Large-deviation statistics

We are interested in the joint statistics of the density  $\rho$  and the velocity divergence  $\theta$  in  $N$  concentric spherical cells of radii  $\{R_k\}$ , indexed by  $1 \leq k \leq N$  in which the mean density and velocity divergence are denoted by  $\Phi = \{\Phi_k\}_{1 \leq k \leq N} = \{(\rho_k, \tilde{\theta}_k)\}_{1 \leq k \leq N}$ .

In the following, we introduce the essential ingredients that are needed to construct the PDFs of densities and velocities in spheres from initially Gaussian statistics. In the main text, we limit ourselves to present the result obtained for the density field (Uhlemann et al. 2016) and use it to obtain a formula for the velocity divergence using the mean relation discussed in Section 2.3. See Appendix B for a more general and comprehensive description of the three basic steps for large-deviation statistics that allow us to map the initially Gaussian statistics to the final ones. For more details or mathematical rigour we refer to Bernardeau et al. (2014, 2015), Bernardeau & Reimberg (2016) and Uhlemann et al. (2016).

*Initial conditions:* The initial decay-rate function  $\Psi(\{\tau_k\})$  describes the exponential decay of the PDF of the initial density contrasts

$\{\tau_k\}$ . We assume Gaussian initial conditions that makes the decay-rate function a quadratic form

$$\Psi_i(\{\tau_k\}) = \frac{1}{2} \sum_{i,j} \Xi_{ij}(\{r_k\}) \tau_i \tau_j, \quad (9)$$

where the precision matrix,  $\Xi_{ij}$ , is the inverse of the cross-correlation (or covariance) matrix  $\Sigma_{ij} = \langle \tau_i \tau_j \rangle$ .

*Spherical collapse map from initial to final densities:* To predict the statistics of spherically averaged quantities one can take advantage of spherical collapse dynamics. This idea is rooted in the so-called contraction principle from large-deviation statistics saying that any large deviation follows the least unlikely of all unlikely transformations between the initial and final state because the PDF is exponentially suppressed when departing from the most likely way. Hence, one can use the one-to-one mapping between the initial density contrast  $\tau_k$  in a sphere of radius  $r_k$  and the final density  $\rho_k = \rho_{\text{SC}}(\tau_k)$  within radius  $R_k$  (with  $r_k = \rho_k^{1/3} R_k$  due to mass conservation) to determine the final decay-rate function as

$$\Psi_f(\{\rho_k\}) = \Psi_i(\{\tau_k : \rho_k = \rho_{\text{SC}}(\tau_k)\}). \quad (10)$$

Note that, while mapping the decay-rate functions with spherical collapse is allowed, one cannot simply in practice map the PDFs with spherical collapse because the relation between initial and final radii arising from mass conservation mixes scales such that fixing a final radius is not identical to fixing the initial radius. To avoid the problem of directly mapping PDFs, this decay-rate function can be used to derive the CGF and from there reconstruct the PDF. One can find a saddle-point approximation, applied to a suitably mapped set of density variables  $\mu_j(\{\rho_k\})$  to implement this procedure efficiently

$$\mathcal{P}(\{\rho_k\}) = \left| \det \left[ \frac{\partial \mu_i}{\partial \rho_j} \right] \right| \sqrt{\det \left[ \frac{\partial^2 \Psi_f}{\partial \mu_i \partial \mu_j} \right]} \frac{\exp[-\Psi_f]}{(2\pi)^{N/2}}. \quad (11)$$

*Spherical collapse map from densities to velocity divergences:* Thanks to the one-to-one mapping between final density and velocity divergence, as shown in more detail in Appendix B, one can simply rely on the results obtained for the density PDFs in Uhlemann et al. (2016) to obtain the PDFs of the velocity divergence and joint PDFs by a simple remapping

$$\mathcal{P}(\{\rho_k\}, \{\tilde{\theta}_k\}) = \mathcal{P}(\{\rho_k\}, \{\rho_k = \rho_{\text{SC}}(\tilde{\theta}_k)\}) \prod_k \rho'(\tilde{\theta}_k). \quad (12)$$

In the main text, we only discuss the validity of the prediction for the PDF. For a discussion of the accuracy of the tree-order perturbation theory prediction for cumulants of the velocity divergence and density field, we refer the reader to Appendix D. Note that the assumption that the mapping between the density and velocity divergence is one-to-one restricts the application of these formulas to cases where the radii  $R_k$  and  $R_{\tilde{k}}$  are sufficiently different from each other, since in the limit  $R_{\tilde{k}} \rightarrow R_k$  one would obtain a delta function centred around the spherical collapse relation which is not realistic as shown in Fig. C1. In practice, our predictions work reasonably well for radii of  $R_1 = 10$  and  $R_2 = 15 \text{ Mpc } h^{-1}$ , while the resolution of the simulation does not allow us to make a definite statement for the joint PDF with  $R_1 = 10$  and  $R_2 = 11 \text{ Mpc } h^{-1}$ .

*Ensuring the correct normalization and mean:* Due to the remappings, the saddle-point PDFs have to be modified to ensure normalization and the correct mean and variance for both the density and the velocity divergence. Let us illustrate this procedure based

on the one-cell case where it can be easily implemented using equation (15).

Since the saddle-point method yields only an approximation to the exact PDF, the PDF obtained from equation (11) has to be normalized  $\hat{\mathcal{P}}_R(\rho) = \mathcal{P}_R(\rho) / \int d\rho \mathcal{P}_R(\rho)$ . Then, since we apply the saddle-point approximation to a transformed variable we have to impose the correct mean of this variable in order to ensure the correct mean of the density  $\langle \rho \rangle = 1$  and velocity divergence  $\langle \theta \rangle = 0$ . The formulas presented for the PDF so far assumed zero mean of the transformed variable and hence for the log-mapping  $\tilde{\rho} = \exp \mu$ . If we let the mean  $\tilde{\mu}$  unconstrained, we get instead  $\rho = \exp(\mu - \tilde{\mu})$ . From this equation, one can derive the PDF of the physical density  $\rho = \tilde{\rho} / \exp \tilde{\mu}$

$$\mathcal{P}_\rho(\rho) = \frac{\exp \tilde{\mu}}{\rho} \mathcal{P}_\mu(\mu = \log \tilde{\rho} = \log(\rho \times \exp \tilde{\mu})). \quad (13)$$

In this equation, one can either predict  $\tilde{\mu}$  using perturbation theory, measure it from – or fit it to – data, or impose the mean density.

For the density, the mean density should be set to one which requires to determine

$$\exp \tilde{\mu} = \langle \tilde{\rho} \rangle = \int \tilde{\rho} \mathcal{P}_\mu(\mu = \log \tilde{\rho}) d\tilde{\rho} \quad (14)$$

and use this as normalization for the mean in the PDF

$$\hat{\mathcal{P}}_R(\rho) = \mathcal{P}_R(\rho \times \langle \tilde{\rho} \rangle) \times \langle \tilde{\rho} \rangle. \quad (15)$$

In addition, for the velocity divergence we have to enforce zero mean,  $\langle \tilde{\theta} \rangle = 0$ , according to equation (7). Similarly, one can determine the correct normalization of the mean for the two-cell PDFs of the density or velocity divergence by imposing the means.

### 3.2 Parametrizing the covariance

In practice, in order to determine the decay-rate function and therefore the PDFs, one needs to compute the covariance matrix between initial densities in spheres of radii  $r_i$  and  $r_j$

$$\sigma^2(r_i, r_j) = \int \frac{d^3 k}{(2\pi)^3} P^{\text{lin}}(k) W_{3\text{D}}(kr_i) W_{3\text{D}}(kr_j), \quad (16)$$

where  $W_{3\text{D}}(k)$  is the shape of the top-hat window function in Fourier space,

$$W_{3\text{D}}(k) = 3 \sqrt{\frac{\pi}{2}} \frac{J_{3/2}(k)}{k^{3/2}}, \quad (17)$$

and  $J_{3/2}(k)$  the Bessel function of the first kind of order 3/2. Given an initial power spectrum, we can tabulate the full expression for the variance  $\sigma^2(R)$  and the covariance for two distinct radii  $\sigma^2(R_i, R_j)$  from the initial power spectrum by relying on LSSFAST (Codis et al. 2016a). The range of initial radii  $r$  that needs to be considered in the tabulation depends on the final radius and densities in the corresponding sphere according to mass conservation  $r = \rho^{1/3} R$ . For radii of  $R \simeq 10 - 15 \text{ Mpc } h^{-1}$  and densities in the range  $\rho \in [0.1, 5]$ , we typically need to cover  $r \in [5, 25] \text{ Mpc } h^{-1}$ . Alternatively, to obtain straightforward analytical expressions, it can be useful to parametrize the covariance matrix for a power-law linear power spectrum with spectral index  $n$  by

$$\sigma^2(r_i, r_i) = \sigma^2(R_p) \left( \frac{r_i}{R_p} \right)^{-n(R_p)-3}, \quad (18a)$$

$$\sigma^2(r_i, r_{j>i}) = \sigma^2(R_p) \mathcal{G} \left( \frac{r_i}{R_p}, \frac{r_j}{R_p}, n(R_p) \right), \quad (18b)$$

**Table 1.** Variances of the density  $\rho$ , normalized velocity divergence  $\bar{\theta}$ , the log-density  $\mu = \log \rho$  and the equivalent  $\mu_{\bar{\theta}} = \nu \log(1 + \theta/\nu)$  for different radii  $R$  at redshift  $z = 0$  as predicted from linear theory, the lognormal model and measured from the simulation.

$R$ (Mpc $h^{-1}$ )	5	7	10	11	15
$\sigma_{\text{lin}}$	1.060	0.865	0.683	0.638	0.506
$\sigma_{\mu, \text{lognorm}}$	0.868	0.747	0.618	0.584	0.477
$\hat{\sigma}_{\rho}$	1.339	1.012	0.751	0.689	0.529
$\hat{\sigma}_{\mu}$	0.835	0.735	0.619	0.585	0.481
$\hat{\sigma}_{\bar{\theta}}$	0.947	0.786	0.631	0.592	0.475
$\hat{\sigma}_{\mu_{\bar{\theta}}}$	0.810	0.704	0.585	0.556	0.459

where

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

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

with  $\alpha = 1 - n$ . In what follows, we will use the exact linear power spectrum of the simulation to predict one-cell PDFs but the parametric expressions for the two-cell case, while we have checked that the running of the power spectrum has no significant impact on the two-cell results.

The key parameter in the prediction of the PDF is the value of the variance at the pivot scale. In practice, we will treat the variance of the log-density at the pivot scale,  $\sigma^2(R_p)$ , as measured from or adjusted to the simulations as its linear prediction is not sufficient to get accurate predictions. Alternatively, the non-linear value of this parameter could be modelled e.g. following Repp & Szapudi (2017).<sup>2</sup> We report the results for the measured variance of the density  $\rho$ , velocity divergence  $\theta$ , the log-density  $\mu = \log \rho$  and the equivalent logarithm of the density-analogue for the velocity divergence  $\mu_{\bar{\theta}} = \nu \log(1 + \theta/\nu)$  in Table 1.

## 4 VALIDATION OF THE PDFS AGAINST SIMULATIONS

We will now validate our predictions for the one-cell PDF of the density and velocity divergence and the joint PDF of density and velocity divergence at two different scales using the simulation which relies on phase-space resolved velocity field. Thanks to improvements in both the theoretical formalism and the numerical simulation, we will present a precise comparison of residuals for the one-cell PDF of the velocity divergence together with first results on the joint PDF of densities and velocity divergences measured at two different scales.

### 4.1 One-cell PDF comparison

Fig. 4 shows a comparison of the measurements in the  $N$ -body simulation described in Section 2.2 at redshift  $z = 0$  for radii  $R = 10, 11, 15$  Mpc  $h^{-1}$  and the analytical saddle-point approximation for the PDF of the density and velocity divergence, respectively. The measured variances are given in Table 1.

<sup>2</sup> We find that while the prediction from linear theory systematically overestimates the non-linear variance, the lognormal model  $\sigma_{\mu}^2 = \log(1 + \sigma_{\text{lin}}^2)$  is much closer to the measured variances probed here.

For the density PDF, as already found in Uhlemann et al. (2016, 2017), the agreement is very good, below the 2 per cent level for all densities from 0.5 to 2.5 for all radii considered here. This result is impressive given that it is obtained at lower redshift and expected to improve significantly for higher redshifts where the variances (which are about 0.5 here) are reduced. We can also see that the predicted rare event tails agree qualitatively very well with the simulation results both at the large density and low-density ends.

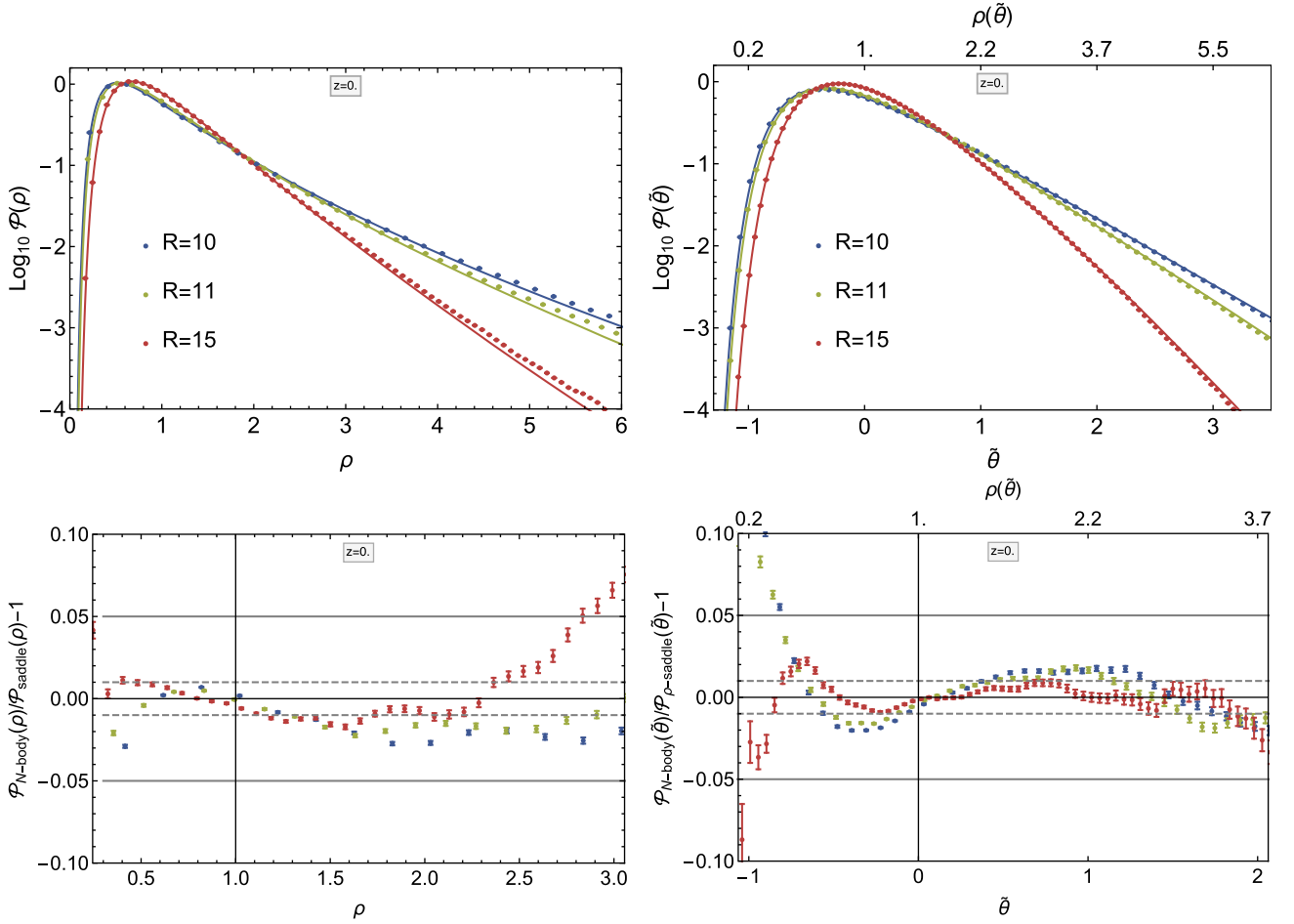
For the PDF of the velocity divergence, first note the significant improvement on the error bars of the simulation under consideration here compared to the results obtained in Bernardeau & van de Weygaert (1996) from a Voronoi and a Delaunay tessellation. This improvement in the accuracy of the velocity measurements allows us to much better test the theoretical predictions from large-deviation statistics. Similar to the density field, we find very good agreement between the analytical prediction and the simulation, at the 2 per cent level for reduced velocity divergences from  $-0.8$  to  $2$  that correspond to densities between 0.5 and 3.5. Also, the rare event tails of the distribution are well-described by the analytical formula with a good match for both large positive velocity divergences and strongly negative velocity divergences. Because of the very good agreement between the prediction and the measurements, we show the same result in Fig. 5 but extended to smaller radii  $R = 5$  and  $7$  Mpc  $h^{-1}$  finding agreement at the 5 per cent level even for those significantly smaller and hence more non-linear scales where the variances become of the order of 1. This shows that the PDF of the velocity divergence is even more robust against changes in the variance than the density, which allows us to probe smaller scales. Appendix D and especially Figs D1 and D2 show how this property also manifests on the reduced cumulants that remain close to the tree-order predictions obtained from perturbation theory at small radii where predictions for the density field deviate. This property of the velocity field was already pointed out in the early works of Kofman et al. (1994), Scoccimarro & Frieman (1996a), Fosalba & Gaztanaga (1998), in particular in Scoccimarro & Frieman (1996b) where it was explicitly stated that for the velocity divergence higher order corrections only start dominating over the tree-order result when the variance approaches unity.

### 4.2 Two-cell and conditional PDF comparison

Let us now present the first results for the joint PDFs of velocity divergences measured at different smoothing scales, the joint PDFs of the velocity divergence and the density measured at different smoothing scales and demonstrate how the joint PDF can be used to obtain predictions for the constrained PDF of the velocity divergence given an over- or underdense environment.

#### 4.2.1 The joint PDFs of velocity divergence and density

The joint PDFs of density and velocity divergence at two different scales obtained after the normalization described in Section 3.1 are shown in Fig. 6 for the case of mixed joint PDFs of density and velocity divergence and the pure joint PDF of the velocity divergence. For the mixed cases, we find a very good agreement in the central regions of the PDF that degrades towards the tails of the distribution. The accuracy of the result is, however, rather good and similar to the accuracy found for the joint statistics of the two-scale density field in previous works (Bernardeau et al. 2015; Uhlemann et al. 2016). The residuals found, especially in the tails and for the pure joint PDF, most likely have physical origins



**Figure 4.** Comparison of the PDF of the density  $\rho$  (left panel) and reduced velocity divergence  $\tilde{\theta}$  (right panel) for radii  $R = 10, 11, 15 \text{ Mpc } h^{-1}$  (upper panel) measured in the  $N$ -body simulation (points) with the saddle-point approximations for the log-density and the corresponding residuals for the measured variance (lower panel). Note that, as expected, the  $\theta$  PDF is well matched over a wider range of equivalent  $\rho$ .

such as correlations between scales beyond the spherical collapse model that require to incorporate tidal corrections. Note that, in order to rule out potential errors that stem from our assumptions (saddle-point approximation for the PDF and parametric expression for the variance), we performed two tests: (i) tabulating numerically the variance of the true linear power spectrum and (ii) computing numerically the inverse Laplace transform to get the PDF. In both cases, we did not find any evidence for a significant impact of respectively the running of the spectral index or the saddle-point approximation, at least within the error bars associated with those noisy methods. This suggests that while the PDF of the velocity divergence was more robust against changes in the variance than the density, especially the joint PDF of two velocity divergences is indeed more sensitive to changes in the variance than the joint PDF of two densities that has been compared to simulations in Uhlemann et al. (2017). Again, we demonstrate this property in Appendix D based on joint cumulants obtained from perturbation theory that are shown in Figs D3 and D4. Given the importance of tidal corrections that has been demonstrated in Scoccimarro & Frieman (1996b), Fosalba & Gaztanaga (1998), Ohta et al. (2003) for the unsmoothed velocity divergence, it would be not surprising that those tidal effects can alter the joint cumulants, even if they tend to cancel for simple cumulants of density and velocity fields smoothed at one scale. For future work, it would be interesting to study the impact of tidal corrections on those joint cumulants for

unsmoothed fields and investigate the potential subsequent erasure of those when smoothed fields are considered.

#### 4.2.2 The constrained PDF for velocity divergence given density

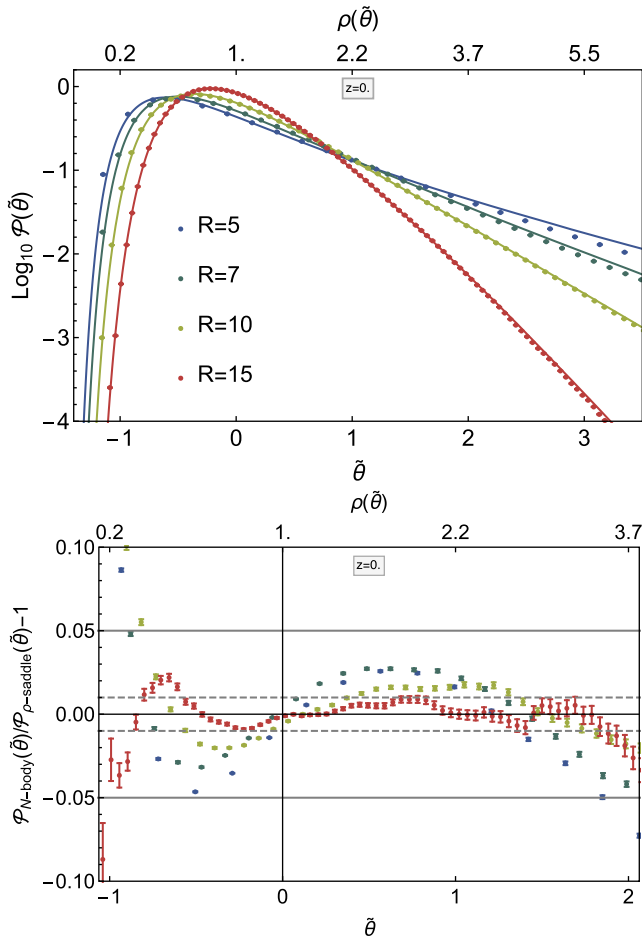
The constrained PDF of the velocity divergence given an over- or underdense environment (or equivalently positive or negative velocity divergence) is obtained from the two-cell PDF by the following marginalization:

$$\mathcal{P}(\tilde{\theta}_1 | \tilde{\theta}_2 \geq 0) = \frac{\int_{-v}^{\infty} d\tilde{\theta}_2 \Theta(\pm\tilde{\theta}_2) \mathcal{P}(\tilde{\theta}_1, \tilde{\theta}_2)}{\int_{-v}^{\infty} d\tilde{\theta}_1 \int_{-v}^{\infty} d\tilde{\theta}_2 \Theta(\pm\tilde{\theta}_2) \mathcal{P}(\tilde{\theta}_1, \tilde{\theta}_2)},$$

$$\mathcal{P}(\tilde{\theta}_1 | \rho_2 \geq 1) = \frac{\int_0^{\infty} d\rho_2 \Theta(\pm(\rho_2 - 1)) \mathcal{P}(\tilde{\theta}_1, \rho_2)}{\int_{-v}^{\infty} d\tilde{\theta}_1 \int_0^{\infty} d\rho_2 \Theta(\pm(\rho_2 - 1)) \mathcal{P}(\tilde{\theta}_1, \rho_2)}, \quad (19)$$

where  $\Theta$  is the Heavyside step function. The result that is obtained from the joint saddle-point PDF by performing the integration in equation (19) numerically is shown in Fig. 7. The one-cell marginal  $\mathcal{P}(\tilde{\theta}_1)$  shown here agrees well with the direct prediction for one-cell that is shown in the upper-right panel of Fig. 4. The agreement between the prediction of the conditional PDFs (solid lines) and





**Figure 5.** Comparison of the velocity divergence PDF as shown in the right panel of Fig. 4 but down to smaller radii  $R = 5, 7, 10, 15 \text{ Mpc } h^{-1}$  (upper panel) as measured in the  $N$ -body simulation (points) with the saddle-point approximations for the log-density and the corresponding residuals for the measured variance (lower panel).

the measurements in our simulation (dots) is very good in particular for similar velocity divergences. We observe, however, some departures for large slopes corresponding to large positive values of  $\theta_1$  when  $\theta_2$  is negative and vice versa. This feature was not observed for the density field and could come from the scatter of the velocity–density relation that is not accounted for in the spherical collapse model.

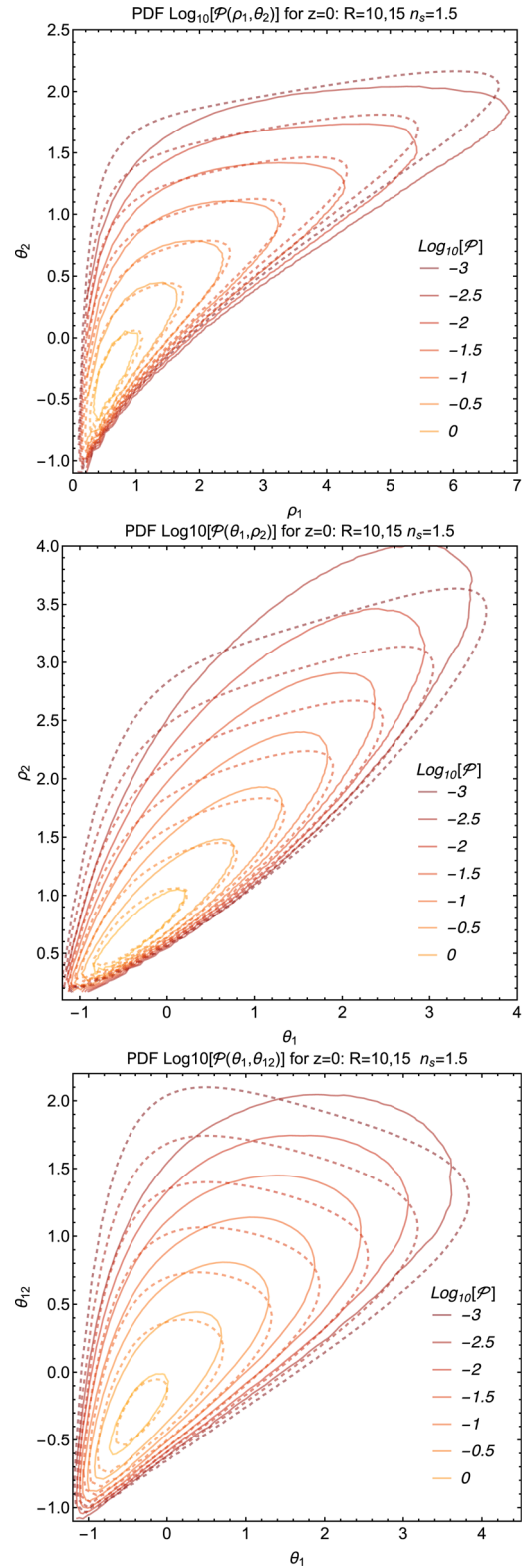
#### 4.2.3 Conditional velocity profile given density

The saddle-point approximation can be used to determine the conditional velocity profile, completely analogous to the density profile obtained in Bernardeau et al. (2014), by using the stationary condition for the decay-rate function

$$0 = \left. \frac{\partial \Psi(\theta_1, \theta_2)}{\partial \theta_2} \right|_{\theta_2 = \tilde{\theta}_2(\theta_1)}, \quad (20)$$

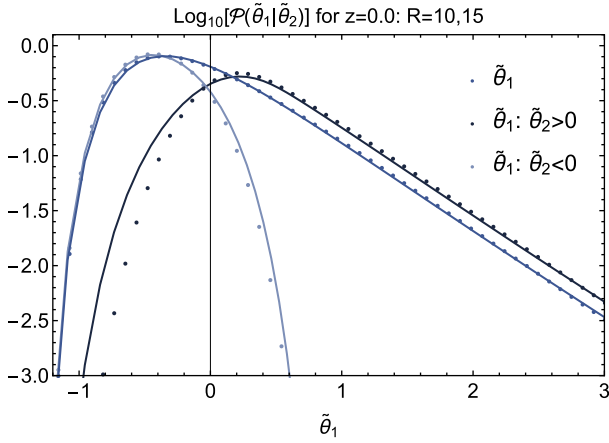
where the constrained variance is given by

$$\langle \theta_2^2 \rangle_{\theta_1} - \langle \theta_2 \rangle_{\theta_1}^2 = \left[ \frac{\partial^2 \Psi(\theta_1, \theta_2)}{\partial \theta_2^2} \right]^{-1} \Bigg|_{\theta_2 = \tilde{\theta}_2(\theta_1)}. \quad (21)$$

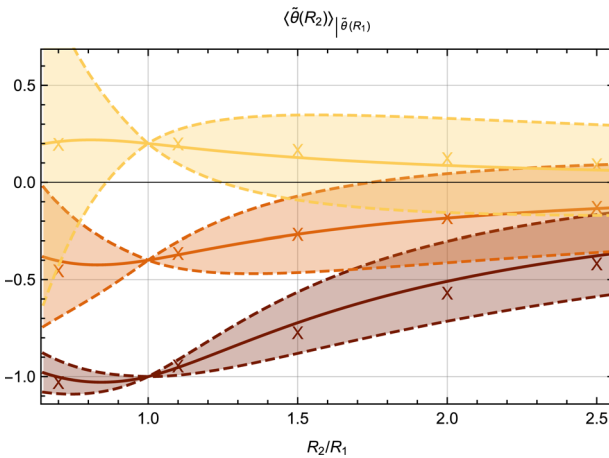


**Figure 6.** Comparison of the joint PDF of the density  $\rho_1$  and velocity divergence  $\theta_2$  (upper panel), the velocity divergence  $\theta_1$  and the density  $\rho_2$  (middle panel) as well as two velocity divergences  $\theta_1$  and  $\theta_{12} = \theta_1 - \theta_2$  (bottom panel), for two different radii  $R_1 = 10$  and  $R_2 = 15 \text{ Mpc } h^{-1}$  measured in the  $N$ -body simulation (thin lines) with the saddle-point approximations for  $\mu_\rho = \log \rho(\theta)$  computed with a power-law initial power spectrum with index  $n = -1.5$  (thick dotted lines). Contour lines scan decades of the probability distribution as labelled.





**Figure 7.** Conditional PDF of the velocity divergence  $\mathcal{P}(\theta_1|\theta_2 \gtrless 0)$  in an overdense (dark blue line) or underdense (light blue line) environment compared to the unconstrained case (blue line) derived from the two-cell saddle-point approximation in comparison to the measurements (data points).



**Figure 8.** Mean conditional profile  $\langle \tilde{\theta}_2 \rangle_{\tilde{\theta}_1} \simeq \langle \tilde{\theta}_2 \rangle_{\rho_1}$  including the variance computed from the saddle-point approximation equation (20) as a function of the ratio of radii  $R_2/R_1$  where  $R_1 = 10 \text{ Mpc } h^{-1}$  for values of  $\tilde{\theta}_1 = \{-1.0, -0.4, 0.2\}$  that corresponds to  $\rho_1 \approx \{0.2, 0.6, 1.2\}$  in comparison to the measured values for  $R_2 = \{7, 15, 20, 25\} \text{ Mpc } h^{-1}$  (crosses).

The result of this prediction is shown in Fig. 8 in comparison to the measurements from the simulation that are in very good agreement with the theory. Note that the conditional profile  $\langle \theta_2 \rangle_{\theta_1}$  corresponds to  $\langle \theta_2 \rangle_{\rho_1}$  if  $\rho_1 = \rho_{\text{sc}}(\theta_1)$ . However, the variance obtained from equation (21) for the mixed case  $\Psi(\rho_1, \theta_2)$  becomes problematic because contributions beyond tree level are important when similar radii are considered due to the scatter around the mean relation between the density and velocity divergence.

## 5 APPLICATION: COSMIC PARAMETER ESTIMATION

The fully analytical probability density function for the mildly non-linear cosmic velocity field within spherical cells described in Section 3 can be used jointly with the corresponding density prediction to build a simple but accurate maximum likelihood estimate for the redshift evolution of the variance of both density [as first investigated by Codis et al. (2016a)] and velocity divergence, which can also be shown to have smaller relative errors than their sample variances. Note that our formalism in principle allows us to get

a maximum likelihood estimate of all higher-order cumulants of the density and velocity divergence fields that would have again a smaller sample variance (see e.g. Seto & Yokoyama 2000, for an estimate of the sample variance of the velocity kurtosis).

Indeed, a dark energy probe may directly attempt to estimate the so-called equation of state parameters ( $w_a, w_p$ ) from the two PDFs while relying on the cosmic model for the growth rate and the growth factor (Blake & Glazebrook 2003),

$$D(z|w_p, w_a) = \frac{5\Omega_m H_0^2}{2} H(a) \int_0^a \frac{da'}{a'^3 H^3(a')}, \quad (22)$$

$$f(z|w_p, w_a) = \frac{d \log D}{d \log a}, \quad (23)$$

$$H^2(a) = H_0^2 \left[ \frac{\Omega_m}{a^3} + \Omega_\Lambda \exp \left( 3 \int_0^z \frac{1+w(z')}{1+z'} dz' \right) \right], \quad (24)$$

with  $\Omega_m, \Omega_\Lambda$  and  $H_0$ , respectively, the dark matter and dark energy densities and the Hubble constant at redshift  $z = 0$ ,  $a \equiv 1/(1+z)$  the expansion factor, and with the equation of state  $w(z) = w_p + w_a/(1+z)$ .

### 5.1 A joint fiducial dark energy experiment

Let us extend the fiducial experiment presented in Codis et al. (2016a) to account for the joint fit of both PDFs. Let us consider a survey with redshifts between 0.1 and 1 binned so that the comoving distance of one bin is  $40 \text{ Mpc } h^{-1}$ , and let us draw regularly spheres of radius  $R = 10 \text{ Mpc } h^{-1}$  separated by  $d = 40 \text{ Mpc } h^{-1}$  (hence, we ignore neighbouring spheres and assume that the spatial correlations are negligible as shown by Codis, Bernardeau & Pichon 2016b). For a 15 000 square degree survey in a flat sky approximation,<sup>3</sup> it yields 50 bins of redshift ( $z_i$ ) with a number of spheres ranging from about  $N_1 = 800$  (at  $z_1 = 0.1$ ) to  $N_{50} = 45\,000$  (at  $z_{50} = 1$ ) for a total of almost 900 000 supposedly independent spheres. In this experiment, we assume that the model for the density and divergence PDFs are exact and that the variances (which are free parameters) are related to the growth rate and the growth factor by linear theory. At each redshift, we can reconstruct the variance of each PDF of  $x$  (either  $\rho$  or  $\theta$ ) by measuring the full PDF

$$\hat{\sigma}_{\text{ML}}(z_j) = \arg \max_{\sigma} \left\{ \prod_{i=1}^{N_j} \mathcal{P}(x_{i,j}|\sigma) \right\}. \quad (25)$$

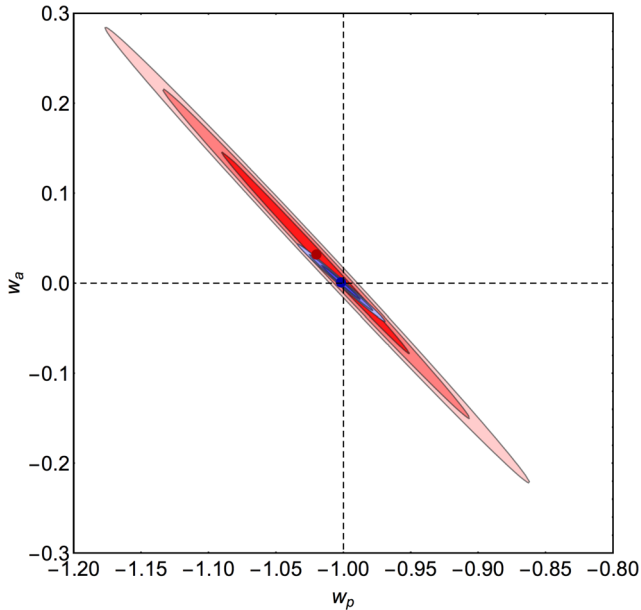
A percent of precision on  $\sigma$  is typically found for each PDF. As mentioned in Codis et al. (2016a), the reconstruction is more accurate at higher redshift where the accessible volume and therefore the number of spheres is larger.

In order to get constraints on the equation of state-of-dark energy, we compute the log-likelihood of the  $\sim 900\,000$  measured densities and velocities  $\{\rho_{i,j}, \theta_{i,j}\}_{1 \leq i \leq N_j, 1 \leq j \leq 50}$  given models for which  $w_p$  and  $w_a$  vary

$$\mathcal{L}(\{x_{i,j}\}|w_p, w_a) = \sum_{j=1}^{50} \sum_{i=1}^{N_j} \log \mathcal{P}_x(x_{i,j}|z_j, w_p, w_a),$$

where  $\mathcal{P}_\rho(\rho|z, w_p, w_a)$  and  $\mathcal{P}_\theta(\theta|z, w_p, w_a)$  are, respectively, the theoretical density PDF for the density  $\rho$  and the divergence  $\theta$  at redshift  $z$  for a cosmological model with dark energy equation of

<sup>3</sup> Here, we use the flat sky approximation and we draw spheres regularly on a grid for the sake of simplicity. To get more precise estimates, one could account for the curvature and check that configuration of spheres maximizes their total number while preventing them from being closer than  $40 \text{ Mpc } h^{-1}$  from their neighbours.



**Figure 9.** Constraints on the equation of state of dark energy in an Euclid-like survey containing about 900 000 spheres of comoving radius  $10 \text{ Mpc } h^{-1}$  regularly drawn between redshift 0.1 and 1. Contours at one, two and three sigmas are displayed with shaded areas from dark to light red when measurements of densities are done and dark to light blue for velocity divergence measurements. The dark red and blue dots correspond to the recovered most likely solution, the input being  $(w_p, w_a) = (-1, 0)$ . The resulting precision on  $w_p$  and  $1 + w_a$  is, respectively, at about the 6 per cent and 10 per cent level for densities and 2 per cent and 2 per cent level for velocities.

state (e.o.s) parametrized by  $w_p$  and  $w_a$ . Optimizing the probability of observing densities and divergences  $\{\rho_{i,j}, \theta_{i,j}\}_{1 \leq i \leq N_j, 1 \leq j \leq 50}$  at redshifts  $\{z_j\}_{1 \leq j \leq 50}$  with respect to  $(w_p, w_a)$  yields a maximum likelihood estimate for the dark energy equation of state parameters

$$(\hat{w}_p, \hat{w}_a) = \arg \max_{w_p, w_a} \{ \mathcal{L}(\{x_{i,j}\} | w_p, w_a) \}. \quad (26)$$

The resulting  $\alpha = 1, 2, 3$  sigma contours are shown in Fig. 9 and correspond to the models for which  $\mathcal{L}(\{x_{i,j}\} | w_p, w_a) = \max_{w_p, w_a} \mathcal{L}(\{x_{i,j}\} | w_p, w_a) + \log(1 - \text{Erf}(\alpha/\sqrt{2}))$ . Modulo our assumptions, this maximum likelihood method allows for constraints on  $w_p$  at a few per cent and  $1 + w_a$  at about a 10 per cent level when the PDF of the density is used, making it a competitive tool for the analysis of future Euclid-like surveys as already pointed out by Codis et al. (2016a). Fig. 9 compares results when only densities are used (in red) to the sole use of velocity data (blue). As expected, for a same survey volume, the velocity divergence PDF allows for much tighter constraints on the equation of state of dark energy than the density PDF alone (by a factor of more than 10!).

In practice, surveys mapping the peculiar velocity field have a much smaller volume than those targeting the galaxy distribution but despite this drawback, they should still carry a competitive cosmological information. It is expected that various uncertainties will degrade the accuracy of the proposed method (uncertainties on the model itself at low-redshift, galaxy biasing, redshift space distortions, observational biases, etc.). Accounting for these effects requires further works, beyond the scope of this paper. Our main conclusion should still hold once all those effects are accounted for, namely that there is a significant gain in using the full knowledge of both PDFs – therefore relying on a maximum likelihood analysis – when contrasted to the direct measurements of cumulants (variance,

skewness, etc.). Large-deviation theory will therefore allow us to get tighter cosmological constraints.

## 5.2 $f(\Omega)$ and $\sigma_8$ from velocity PDF

Alternatively, and to avoid uncertainties due to galaxy biasing, we propose another experiment in which we use velocity data alone. Our model for the PDF of the velocity divergence then depends on the spherical collapse dynamics that is assumed to be fixed here,  $f(\Omega)$  – which is now allowed to vary – and the linear power spectrum whose amplitude will be denoted by  $\sigma_8$  as usual and redshift evolution is governed by dark energy following equation (24). Measuring the velocity divergence PDF in the local Universe then provides low-redshift constraints on  $\sigma_8$  and  $f(\Omega)$  directly (here, we will assume that dark energy is a cosmological constant). To illustrate this point, we measure the mean velocity divergence in 10 000 spheres of  $20 \text{ Mpc } h^{-1}$  at  $z = 0$ . The corresponding constraints on  $\sigma_8$  and  $f(\Omega)$  are shown in the left-hand panel of Fig. 10 which demonstrates that using the full statistics of the velocity divergence field (beyond the sole skewness as proposed e.g. by Bernardeau et al. 1995) can allow us to get interesting constraints. Adding a *Planck* prior on  $\Omega$  in addition allows us to get an estimate of  $\gamma$  and therefore could be used to constraint modified gravity models.

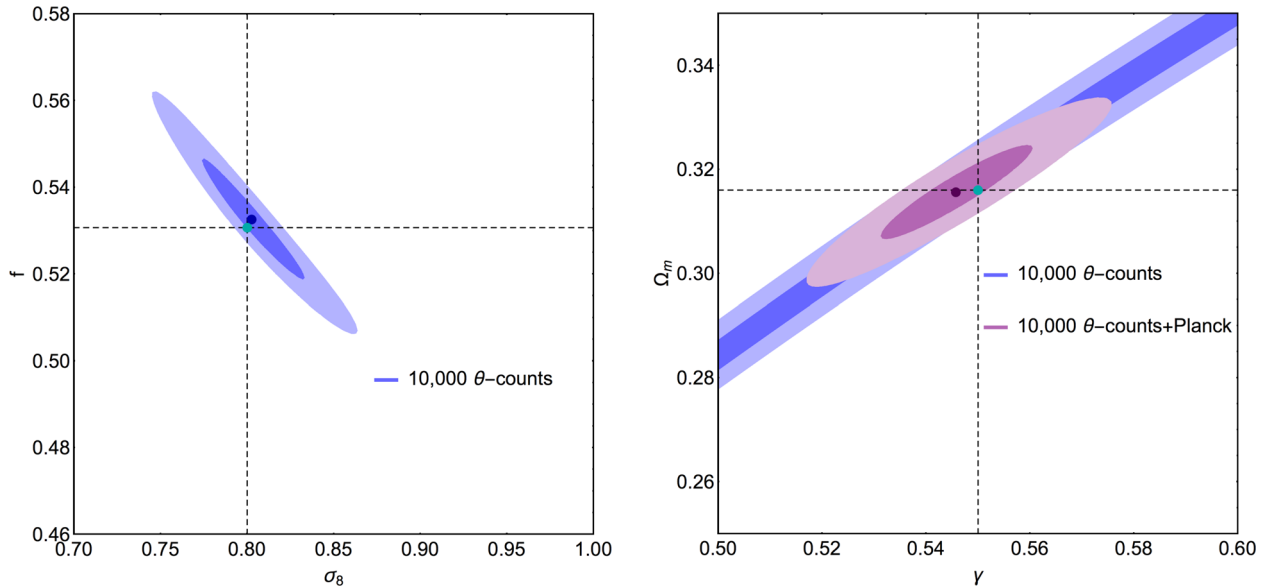
In this work, we have assumed Gaussian initial conditions, but it should be straightforward to include non-Gaussian initial conditions. In this case, the velocity divergence PDF could allow us to measure precisely the departure from Gaussian initial conditions, improving upon previous studies that focused on a subset of cumulants like the skewness (Protogeris & Scherrer 1997) by gathering information from the full PDF (i.e. all cumulants of the field).

## 5.3 Velocity divergence and the extraction of bias

If we assume that the count-in-cell statistics of both the density and the velocity divergence are well-described by large-deviation statistics together with the spherical collapse mapping, we obtain an effective mapping between the two PDFs  $\mathcal{P}_{\tilde{\rho}_m} \xleftrightarrow{\text{SC}} \mathcal{P}_{\rho_m}$ . In Figs 2 and 3, we already demonstrated that the density PDF can be reconstructed from the PDF of the velocity divergence using the spherical collapse mapping. For the simulation under consideration, at redshift  $z = 0$  and radius  $R = 10 \text{ Mpc } h^{-1}$ , the accuracy of this reconstruction is within 10 per cent for densities between 0.3 and 3. Having established this relation between the statistics of the density and velocity field for dark matter now gives us a handle on comparing the statistics of the density of dark matter and haloes. By assuming that velocities between matter and haloes are unbiased  $\mathcal{P}_{\tilde{\theta}_h} \stackrel{\tilde{\theta}_h = \tilde{\theta}_m}{=} \mathcal{P}_{\tilde{\theta}_m}$ , we can obtain an effective measurement of the dark matter density PDF by measuring halo velocities. On the other hand, we can directly measure the halo density PDF that finally allows us to determine the bias between matter and haloes. The idea can be summarized as follows:

$$\mathcal{P}_{\tilde{\theta}_h} \stackrel{\tilde{\theta}_h = \tilde{\theta}_m}{=} \mathcal{P}_{\tilde{\theta}_m} \xleftrightarrow{\text{SC}} \mathcal{P}_{\rho_m} \stackrel{\delta_h = \sum_i b_i \delta_m^i}{\longleftrightarrow} \mathcal{P}_{\rho_h}. \quad (27)$$

Since the reconstruction of the density from the velocity dispersion is accurate around the peak of the distribution, which is significantly transformed due to the effect of bias (Uhlemann et al., in preparation), we expect that measuring the PDF of the velocity divergence along with the PDF of the biased dark matter field will allow us to extract properties of bias.



**Figure 10.** Left-hand panel: constraints on  $\sigma_8$  and  $f(\Omega)$  obtained from velocity measurements in 10 000 spheres of  $20 \text{ Mpc } h^{-1}$  at  $z = 0$ . Right-hand panel: same as left-hand panel when we add a prior from *Planck* on  $\Omega$  to get constraints on  $\gamma$ . Input parameters are shown with the cyan dot while blue and purple dots show the maximum likelihood estimates. The one- and two-sigma contours are shown with the dark and light-shaded areas.

#### 5.4 Measurements of peculiar velocities from real surveys

In galaxy redshift surveys, the observables that are related to the radial peculiar velocity (along the line of sight) are the redshifts and radial distances of galaxies  $cz = H_0 r - v_r$ . While the redshift can be obtained accurately from spectroscopy, extracting a peculiar velocity field also requires using empirical distance indicators such as the Tully–Fisher relation for spiral galaxies (Tully & Fisher 1977, relating the luminosity to the angular velocity  $L \propto v^4$ ), the Faber–Jackson relation (Faber & Jackson 1976, relating the luminosity to the central stellar velocity dispersion  $L \propto \sigma^4$ ) for elliptical galaxies and extended versions thereof such as the Fundamental Plane (Djorgovski & Davis 1987) or luminosity–distance relation for Type Ia supernovae (Tonry et al. 2003). Note that those traditional ways of estimating distances have a distant-dependent uncertainty and it gets increasingly difficult to estimate velocities for deep samples and large volumes. One can avoid this problem by using the CMB temperature distortion that is induced by the kinetic Sunyaev–Zeldovich effect (due to the scattering of CMB photons by the ionized gas in intervening structures that induces a Doppler shift depending on the line-of-sight peculiar velocity) that allows us to extract peculiar velocities for galaxy clusters (Hand et al. 2012), but due to a weakness in signal and the contamination by primary CMB fluctuations, this method does not seem to be promising for detecting peculiar velocities of individual galaxies.

All those methods have in common that they only give the line-of-sight component of the peculiar velocity field. However, the three-dimensional velocity field can be reconstructed using methods like POTENT (Bertschinger & Dekel 1989), Least Action (Peebles 1989) or Monge–Ampère–Kantorovich (MAK; Brenier et al. 2003). We believe that the underlying assumption that the velocity can be well approximated as a gradient field is not an obstruction in our case, given that our statistical estimator is for average velocity divergences in spheres such that the radial component of the peculiar velocity suffices and non-linear effects of vorticity on very small scales are irrelevant (Pichon & Bernardeau 1999).

## 6 CONCLUSIONS

The general formalism to describe the cosmic evolution of the joint statistics of the density and the velocity divergence fields in multiple concentric spheres was presented and validated against state-of-the-art numerical estimators for the velocity divergence.

In particular, we obtained fully analytical predictions for the PDF of the velocity divergence in a sphere that are accurate at the 2 per cent level for a wide range of velocity divergences that extend beyond the range in which the density PDF has a comparable accuracy. It was made possible through advances in the analytical theory for predicting PDFs as well as an improvement in simulations allowing to accurately resolve velocity fields. We also obtained the first theoretical prediction for the joint PDFs of densities and velocity divergences in spheres of two different radii and compared it to the simulation finding qualitatively good agreement. We used this result to compute the conditional PDF of finding a velocity divergence, given a surrounding over or underdense region and the corresponding conditional velocity profile that is of interest for studies of both haloes and voids.

To highlight the practical implications of our results, we built a maximum likelihood estimator for the joint redshift evolution of the variance of the density and the velocity divergence that allows us to constrain dark energy considerably better than using density fields alone. Finally, we provided an outlook of how the relationship between the density and velocity divergence in spheres and hence their PDFs might be used to estimate bias. For the future, it would be interesting to investigate the feasibility of a removal of redshift space distortions using the density–velocity relationship based on upcoming results for the PDF of the biased density fields (Uhlemann et al., in preparation). It could also be of interest to also revisit the two-point large separation limit for the bias function presented in Codis et al. (2016b) and Uhlemann et al. (2017) to predict the pairwise peculiar velocities function in the mildly non-linear regime.

**ACKNOWLEDGEMENTS**

This work is partially supported by the grants ANR-12-BS05-0002 and ANR-13-BS05-0010 of the French *Agence Nationale de la Recherche*. CU is supported by the Delta-ITP consortium, a programme of the Netherlands organization for scientific research (NWO) funded by the Dutch Ministry of Education, Culture and Science (OCW). OH acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 679145, project 'COSMO-SIMS'). We warmly thank Martin Feix, Juhan Kim and Dmitri Pogosyan for discussions and the anonymous referee for useful suggestions.

**REFERENCES**

- Abel T., Hahn O., Kaehler R., 2012, *MNRAS*, 427, 61  
 Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, *MNRAS*, 426, 2046  
 Baldauf T., Desjacques V., Seljak U., 2015, *Phys. Rev. D*, 92, 123507  
 Balian R., Schaeffer R., 1989, *A&A*, 220, 1  
 Bernardeau F., 1992, *ApJ*, 390, L61  
 Bernardeau F., 1994, *A&A*, 291, 697  
 Bernardeau F., Reimberg P., 2016, *Phys. Rev. D*, 94, 063520  
 Bernardeau F., van de Weygaert R., 1996, *MNRAS*, 279, 693  
 Bernardeau F., Juszkiewicz R., Dekel A., Bouchet F. R., 1995, *MNRAS*, 274, 20  
 Bernardeau F., Chodorowski M. J., Lokas E. L., Stompór R., Kudlicki A., 1999, *MNRAS*, 309, 543  
 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  
 Bertschinger E., Dekel A., 1989, *ApJ*, 336, L5  
 Bilicki M., Chodorowski M. J., 2008, *MNRAS*, 391, 1796  
 Blake C., Glazebrook K., 2003, *ApJ*, 594, 665  
 Brenier Y., Frisch U., Hénon M., Loeper G., Matarrese S., Mohayaee R., Sobolevskii A., 2003, *MNRAS*, 346, 501  
 Chodorowski M. J., Lokas E. L., 1997, *MNRAS*, 287, 591  
 Chodorowski M. J., Lokas E. L., Pollo A., Nusser A., 1998, *MNRAS*, 300, 1027  
 Codis S., Pichon C., Bernardeau F., Uhlemann C., Prunet S., 2016a, *MNRAS*, 460, 1549  
 Codis S., Bernardeau F., Pichon C., 2016b, *MNRAS*, 460, 1598  
 Djorgovski S., Davis M., 1987, *ApJ*, 313, 59  
 Elia A., Kulkarni S., Porciani C., Pietroni M., Matarrese S., 2011, *MNRAS*, 416, 1703  
 Faber S. M., Jackson R. E., 1976, *ApJ*, 204, 668  
 Fosalba P., Gaztanaga E., 1998, *MNRAS*, 301, 535  
 Hahn O., Angulo R. E., Abel T., 2015, *MNRAS*, 454, 3920  
 Hand N., Addison G. E., Aubourg E., Battaglia N., Battistelli E. S. et al., 2012, *Phys. Rev. Lett.*, 109, 041101  
 Jennings E., 2012, *MNRAS*, 427, L25  
 Juszkiewicz R., Bouchet F. R., Colombi S., 1993, *ApJ*, 412, L9  
 Juszkiewicz R., Weinberg D. H., Amsterdamski P., Chodorowski M., Bouchet F., 1995, *ApJ*, 442, 39  
 Kaiser N., 1984, *ApJ*, 284, L9  
 Kofman L., Bertschinger E., Gelb J. M., Nusser A., Dekel A., 1994, *ApJ*, 420, 44  
 McConnachie A. W. et al., 2016, preprint ([arXiv:1606.00060](https://arxiv.org/abs/1606.00060))  
 Nadkarni-Ghosh S., 2013, *MNRAS*, 428, 1166  
 Nusser A., Branchini E., Davis M., 2012, *ApJ*, 755, 58  
 Ohta Y., Kayo I., Taruya A., 2003, *ApJ*, 589, 1  
 Peebles P. J. E., 1989, *ApJ*, 344, L53  
 Pichon C., Bernardeau F., 1999, *A&A*, 343, 663  
 Protogeros Z. A. M., Scherrer R. J., 1997, *MNRAS*, 286, 223  
 Repp A., Szapudi I., 2017, *MNRAS*, 464, L21  
 Scoccimarro R., Frieman J., 1996a, *ApJ*, 473, 620

- Scoccimarro R., Frieman J., 1996b, *ApJS*, 105, 37  
 Seto N., Yokoyama J., 2000, *PASJ*, 52, 749  
 Shandarin S., Habib S., Heitmann K., 2012, *Phys. Rev. D*, 85, 083005  
 Tonry J. L., Schmidt B. P., Barris B., Candia P., Challis P. a. a., 2003, *ApJ*, 594, 1  
 Tully R. B., Fisher J. R., 1977, *A&A*, 54, 661  
 Uhlemann C., Codis S., Pichon C., Bernardeau F., Reimberg P., 2016, *MNRAS*, 460, 1529  
 Uhlemann C., Codis S., Kim J., Pichon C., Bernardeau F., Pogosyan D., Park C., L'Huillier B., 2017, *MNRAS*, 466, 2067  
 Valageas P., 2002, *A&A*, 382, 412

**APPENDIX A: LSSFAST PACKAGE**

The density PDF of  $\tilde{\theta}$  for power-law and arbitrary power spectra are made available in the updated LSSFAST package distributed online.

**APPENDIX B: CONSTRUCTION OF THE PDFS****B1 The large-deviation principle**

Using large-deviation theory, PDFs can be constructed following three steps.

*Step 1: Spherical collapse map to the final rate function.* The final decay-rate function can be obtained from the initial one using the so-called contraction principle saying that any large deviation follows the least unlikely of all unlikely transformations between the initial and final states. When considering statistics of spherically averaged quantities the leading order contribution to the time evolution can be identified as the dynamics of spherical collapse. This one-to-one mapping between the initial density contrast  $\tau_k$  in a sphere of radius  $r_k$  and the final density and velocity divergence  $(\rho_k, \theta_k) = (\rho_{SC}, \theta_{SC})(\tau_k)$  within radius  $R_k$  (with  $r_k = \rho_k^{1/3} R_k$  due to mass conservation) allows us to determine the final decay-rate function as

$$\Psi_f(\Phi = \{(\rho, \theta)_k\}) = \Psi_i(\{\tau_k : (\rho, \theta)_k = (\rho_{SC}, \theta_{SC})(\tau_k)\}). \quad (\text{B1})$$

*Step 2: Legendre transform to the cumulant generator.* The CGF  $\varphi$  is obtained as a Legendre transform of the decay-rate function  $\Psi(\Phi)$  according to Varadhan's theorem:

$$\varphi(\Lambda) = {}^t \Lambda \cdot \Phi - \Psi(\Phi), \quad (\text{B2})$$

where

$$\Lambda = \{\Lambda_k\} = \{\nabla_{\Phi_k} \Psi(\Phi)\}. \quad (\text{B3})$$

*Step 3: Inverse Laplace transform to the PDF.* Once the CGF  $\varphi(\Lambda)$  is known, the PDF  $\mathcal{P}(\Phi)$  can be computed from the usual inverse Laplace transform

$$\mathcal{P}(\Phi) = \int_{-\infty}^{+\infty} \frac{d\Lambda}{(2\pi i)^{2N}} \exp(-{}^t \Lambda \cdot \Phi + \varphi(\Lambda)), \quad (\text{B4})$$

which requires an integration in the complex plane.

**B2 Saddle-point approximation**

In general, the numerical integration in the complex plane of equation (B4) can be challenging because of numerical oscillations and singularities. Indeed, the CGF often presents poles and branch cuts as underlined e.g. by Bernardeau et al. (2014) due the fact that the



first derivative of the rate function has a maximum beyond which  $\Lambda$  is not defined. An alternative to the brute force numerical integrations performed in Bernardeau et al. (2014, 2015) is to rely on a saddle-point approximation when the singularity is sufficiently far away from the values of interest. Depending on the variable used in the saddle-point approximation, the domain of validity of this analytical approximation varies.

Let us now describe how the complex integral can be obtained from a saddle-point approximation that, using a well-chosen variable, provides analytical results as good as the numerical integration. Let us first call the appropriate set of variables for the general  $N$  cell case  $\{\mu_k\}$ ; we will later give the concrete variables that are suited to obtain good predictions for the joint density and velocity statistics. A saddle-point approximation of equation (B4) allows for a direct shortcut from step 1 to step 3. Indeed, applying the saddle-point approximation gives the following PDF

$$\mathcal{P}_{\mu, \{R_k\}}(\{\mu_k\}) = \sqrt{\det \left[ \frac{\partial^2 \Psi_{\{R_k\}}}{\partial \mu_i \partial \mu_j} \right]} \frac{\exp[-\Psi_{\{R_k\}}]}{(2\pi)^{N/2}}, \quad (\text{B5a})$$

which does not require the explicit calculation of the Legendre transform described in step 2. From equation (B5a), it is clear that the variables  $\{\mu_k\}$  have to be chosen to ensure the convexity of the decay-rate function ( $\partial^2 \Psi_{\{R_k\}} / \partial \mu_i \partial \mu_j > 0$ ) which in practice boils down to pushing the singularity back to infinity. Equation (B5a) can then be translated in the PDF of the density or velocity field ( $\rho, \theta$ ) =  $(\Phi^1, \Phi^2)$  via a simple change of variables

$$\mathcal{P}_{R_k}(\{\Phi_k^a\}) = \mathcal{P}_{\mu, \{R_k\}}[\{\mu_k(\{\Phi_i^a\})\}] \left| \det \left[ \frac{\partial \mu_i}{\partial \Phi_j^a} \right] \right|, \quad (\text{B5b})$$

where the Hessian of the decay-rate function  $\Psi_{\{R_k\}}$  after a change of variables  $\{\Phi_k^a\} \rightarrow \{\mu_k\}$  is given by

$$\frac{\partial^2 \Psi_{\{R_k\}}}{\partial \mu_i \partial \mu_j} = \frac{\partial \Phi_k^a}{\partial \mu_i} \times \frac{\partial^2 \Psi_{\{R_k\}}}{\partial \Phi_k^a \partial \Phi_l^a} \times \frac{\partial \Phi_l^a}{\partial \mu_j} + \frac{\partial^2 \Phi_k^a}{\partial \mu_i \partial \mu_j} \times \frac{\partial \Psi_{\{R_k\}}}{\partial \Phi_k^a}.$$

### B3 Optimizing the reach of the saddle-point approximation

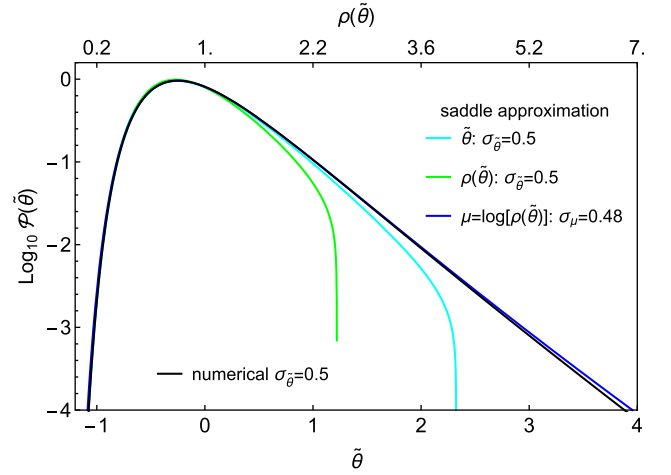
After introducing the general formula for the saddle-point approximation applied to a set of variables  $\{\mu_k\}$ , we now present exemplary results for different choices of variables for the one-cell case. This demonstrates the optimality of the log-density and log-mass mapping we use for the predictions and the comparison to the simulation in Section 4.

The reach of the saddle-point approximation applied to the density  $\rho$  and the log-density  $\mu_\rho = \log \rho$  has been extensively compared in Uhlemann et al. (2016), finding that the reach of the approximation for the log-density covers the whole range of densities that are of interest, while the application to the density leads to critical points that prevent its practical use. Inspired by this finding, we will generalize it to obtain the PDF of the velocity divergence  $\theta$ .

In analogy to the PDF of the density, the saddle-point approximation can be also applied with the velocity divergence as variable such that we have  $\mu = \theta$  and hence

$$\mathcal{P}_R(\tilde{\theta}) = \sqrt{\Psi_R''[\tilde{\theta}]/(2\pi)} \exp(-\Psi_R[\tilde{\theta}]). \quad (\text{B6})$$

As for the saddle-point approximation of the density PDF considered in Bernardeau et al. (2014, 2015) this approximation fails as soon as the decay-rate function  $\Psi_R(\tilde{\theta})$  becomes non-convex. Although the saddle-point approximation for the velocity divergence has a larger range of validity than the one for the density (see the



**Figure B1.** Comparison of the reach of the saddle-point approximation for the PDF of the velocity divergence  $\theta$ . The normalized saddle-point PDF from a logarithmic transform according to equation (B7) (blue line) gives the same result as the numerical integration in the complex plane for  $\tilde{\theta}$  (black line) in the whole range under consideration while the saddle-point approximations for the density (green line) and velocity divergence (cyan line) fail earlier.

cyan and green lines in Fig. B1), eventually criticality prevents its use for the whole range of values of interest.

However, as has been shown for the density in Uhlemann et al. (2016) the criticality of the decay-rate function can be circumvented by using a logarithmic mapping hence allowing for fully analytical predictions of the PDF based on a saddle-point approximation. Indeed, the PDF of the velocity divergence can be obtained by applying a saddle-point approximation to the log-density which is in turn related to the velocity divergence by a one-to-one mapping

$$\mu_{\rho(\tilde{\theta})} = \log(\rho_{\text{SC}}(\tilde{\theta})). \quad (\text{B7})$$

Note that this procedure is completely equivalent to using a logarithmic mapping for  $v + \tilde{\theta}$ .

In Fig. B1, we compare the range of validity for the saddle-point approximation applied to different variables and compare them to the direct result for the PDF obtained from a numerical integration of equation (B4). We find that while the saddle-point approximations applied to the density or velocity divergence fail early, the log-density is valid in the whole range of interest and in very good agreement with the numerical integration.

## APPENDIX C: DISPERSION AND DENSITY RELATION

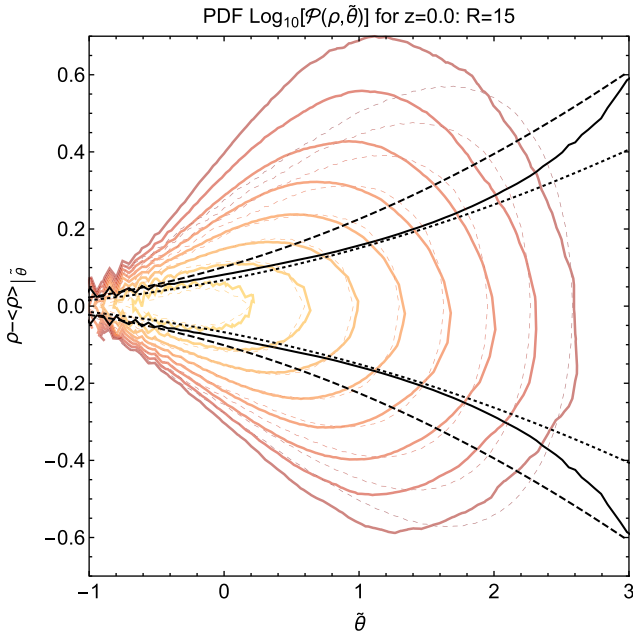
The scatter around the mean relation between density and velocity divergence can be obtained from conditional averages of

$$\sigma_{\rho|\tilde{\theta}}^2 = \langle (\rho - \langle \rho \rangle_{|\tilde{\theta}})^2 \rangle_{|\tilde{\theta}}. \quad (\text{C1})$$

They can also be expanded in perturbation theory up to third order

$$\sigma_{\rho|\tilde{\theta}}^2 = b_0 \sigma_{\tilde{\theta}}^4 + b_2 \sigma_{\tilde{\theta}}^2 \tilde{\theta}^2 + b_1 \sigma_{\tilde{\theta}}^4 \tilde{\theta}, \quad (\text{C2})$$

which has been computed at leading order  $\mathcal{O}(\sigma^4)$  in Chodorowski et al. (1998). Since  $b_2$  is identically zero for a top-hat filter, Bernardeau et al. (1999) argued that the next-to-leading order  $\mathcal{O}(\sigma^5)$  can be of importance and included it in a phenomenological



**Figure C1.** Joint PDF of the velocity divergence (bold coloured lines) and the scatter around the measured mean  $\langle \rho \rangle_{|\tilde{\theta}}$  together with the typical scatter from the simulation (black line) and the fit equation (C3) from Bernardeau et al. (1999) (black dashed line) together with our improved fit (black dotted line) and the corresponding Gaussian model for the joint PDF (thin coloured lines) from equation (C4). Note that the region  $\tilde{\theta} \sim -1$  is completely dominated by a binning effect as the bin we use is approximately  $\Delta\tilde{\theta} = 0.08$ .

fit against numerical simulations finding

$$\sigma_{\rho|\tilde{\theta}} = b_0 \left( 1 + \tilde{\theta} + \frac{2}{9}\tilde{\theta}^2 \right) \sigma_{\tilde{\theta}}^2, \quad (\text{C3})$$

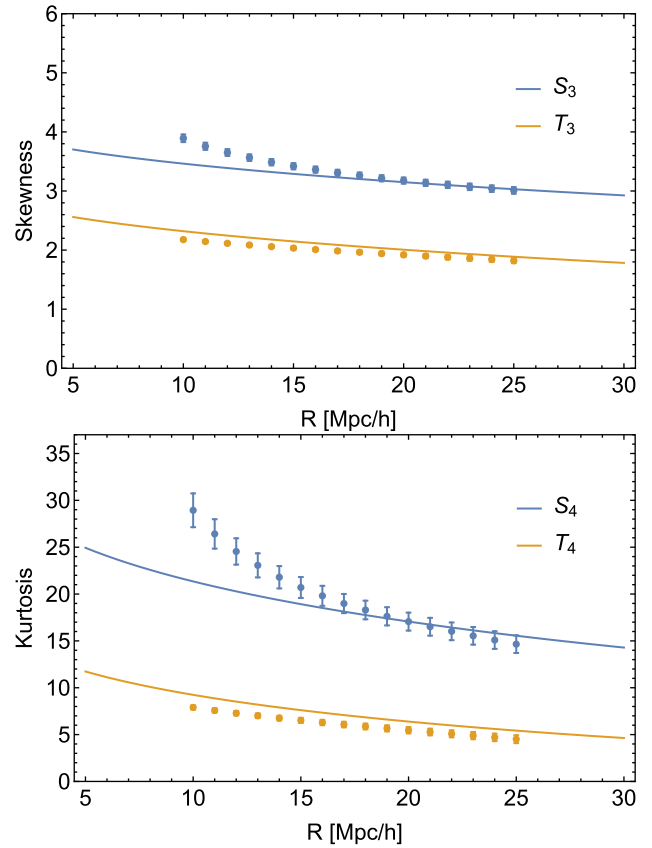
with  $b_0 = 0.45$ . The scatter around the mean found numerically is shown in Fig. C1 (black solid line) together with the fit from equation (C3) using the original coefficient  $b_0 = 0.45$  (dashed line) and our modified fit that uses  $b_0 = 0.3$  (dotted line) and has been checked to give good results for smoothing radii  $R = 10, 11$  and  $15 \text{ Mpc } h^{-1}$ . Interestingly,  $0.3^2 = 0.09$  is much closer to the value predicted by perturbation theory that, depending on the spectral index, is between  $b_0 \simeq 0.05 - 0.09$  for smoothing radii of interest here. If one assumes that the mean and the scatter completely characterize the joint distribution, one can make the following ansatz for the joint PDF of density and velocity divergence at the same scale  $R$

$$P_R(\rho, \tilde{\theta}) = \frac{1}{\sqrt{2\pi\sigma_{\rho|\tilde{\theta}}^2}} \exp \left[ -\frac{(\rho - \langle \rho \rangle_{|\tilde{\theta}})^2}{2\sigma_{\rho|\tilde{\theta}}^2} \right] P_R(\tilde{\theta}), \quad (\text{C4})$$

where the appropriate PT-inspired fit has to be inserted for  $\sigma_{\rho|\tilde{\theta}}$  according to equation (C3). We show the result in comparison to the simulation measurements in Fig. C1.

#### APPENDIX D: CUMULANTS OF THE DIVERGENCE

At tree-order (i.e. vanishing variance), the reduced cumulants  $T_n = \langle \tilde{\theta}^n \rangle_c / \sigma_{\tilde{\theta}}^{2(n-1)}$  of the velocity divergence field smoothed with a top-hat filter can all be obtained by applying a large-deviation principle to get the scaled CGF. In particular, for an Einstein–de



**Figure D1.** The first two reduced cumulants being the skewness (upper panel) and kurtosis (lower panel) of the density,  $S_{3|4}$ , and velocity divergence,  $T_{3|4}$  measured in the simulations (points error bars) compared to their tree-order prediction (solid lines) for different radii,  $R$ , of the top-hat filter.

Sitter Universe, its skewness and kurtosis read (Bernardeau 1994)

$$T_3 = \frac{26}{7} + \gamma_1(R), \quad (\text{D1})$$

$$T_4 = \frac{12088}{441} + \frac{338}{21}\gamma_1(R) + \frac{7}{3}\gamma_1(R)^2 + \frac{2}{3}\gamma_2(R), \quad (\text{D2})$$

where  $\gamma_m(R) = d^m \log \sigma^2(R) / d(\log R)^m$ . Let us also recall the result for the reduced cumulants  $S_n = \langle \rho^n \rangle_c / \sigma_\rho^{2(n-1)}$  of the top-hat filtered density field

$$S_3 = \frac{34}{7} + \gamma_1(R), \quad (\text{D3})$$

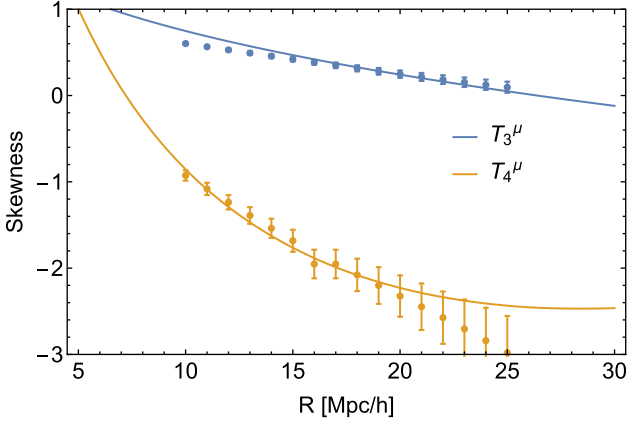
$$S_4 = \frac{60712}{1323} + \frac{62}{3}\gamma_1(R) + \frac{7}{3}\gamma_1(R)^2 + \frac{2}{3}\gamma_2(R). \quad (\text{D4})$$

The measured skewness and kurtosis of the density and velocity divergence fields are compared to their tree-order PT prediction in Fig. D1 that improves upon early results by Bernardeau & van de Weygaert (1996).

In this paper, we also consider the log-variable  $\mu = \log v + \tilde{\theta}$  (or equivalently  $\mu = \log(1 + \tilde{\theta}/v)$  as they only differ by their mean) whose first cumulants follow

$$T_3^\mu = vT_3 - 3 + \mathcal{O}(\sigma^2), \quad (\text{D5})$$

$$T_4^\mu = v^2T_4 - 12vT_3 + 20 + \mathcal{O}(\sigma^2), \quad (\text{D6})$$



**Figure D2.** Same as Fig. D1 for the skewness and kurtosis of the log-variable associated with the velocity divergence  $\log(v + \tilde{\theta})$ .

in complete analogy to the cumulants of the log-density  $\mu = \log \rho$

$$S_3^\mu = S_3 - 3 + \mathcal{O}(\sigma^2), \quad (\text{D7})$$

$$S_4^\mu = S_4 - 12S_3 + 20 + \mathcal{O}(\sigma^2), \quad (\text{D8})$$

once the cumulants of  $\rho$  are replaced by the cumulants of  $\tilde{\theta}/v$  which in turn can be expressed as a function of the cumulants of  $\tilde{\theta}$  and  $v$ . It has to be emphasized here that according to the approximation given by equation (1b),  $\rho \approx v \log(1 + \tilde{\theta}/v)$  so that we have the following approximate relations:

$$T_3^\mu \approx v S_3^\mu, \quad (\text{D9})$$

$$T_4^\mu \approx v^2 S_4^\mu, \quad (\text{D10})$$

which is true at tree order for the skewness (as expected since the  $v$ -approximation we use for the spherical collapse is designed to reproduce the zero variance limit of the skewness) but not the kurtosis. Despite those approximate relations, we stick to the variable  $\mu = \log(1 + \tilde{\theta}/v)$  instead of  $v\mu$  for aesthetic purposes.<sup>4</sup>

We compare the above-defined tree-order PT predictions of  $\log v + \tilde{\theta}$  to measurements in our simulation in Fig. D2. Similar to what was found in Uhlemann et al. (2016) for the log-density, the log-variable associated with the velocity divergence,  $\mu = \log v + \tilde{\theta}$ , has cumulants that are closer to the tree-order predictions. This motivates the choice to rely on a large-deviation principle for the log variable instead of the velocity divergence itself. In addition, we showed in the main text that it allows for fully analytical predictions for the PDF.

Note also that a perturbative expansion can be used to map the variance of the velocity divergence to the variance of the density thanks to the approximation of the spherical collapse dynamics (equation 2). It reads

$$\sigma_{\tilde{\theta}}^2 = \sigma_\rho^2 + \sigma_\rho^4 \left( \frac{88}{147} - \frac{8S_3}{21} \right). \quad (\text{D11})$$

In practice, this perturbative relation can be improved in the quasi-linear regime by using the fully non-linear relation given by equation (12) at the level of the PDFs.

<sup>4</sup>We note here that this simple link between the log-density and the log variable associated with the velocity divergence implies that applying the large-deviation principle to  $\log \rho$  or  $\mu = \log(1 + \tilde{\theta}/v)$  is equivalent in the  $v$ -approximation we use in this paper.

Beyond one-cell statistics, the large-deviation principle can be used to also compute mixed cumulants involving the field (density and velocity divergence) at different scales. To do so, let us first write down the exact spherical collapse mapping as

$$\rho_{\text{SC}}(\tau) = \sum_{i \geq 0} \frac{v_i}{i!} \tau^i, \quad (\text{D12})$$

$$\tilde{\theta}_{\text{SC}}(\tau) = \sum_{i \geq 1} \frac{\mu_i}{i!} \tau^i, \quad (\text{D13})$$

with  $v_i$  and  $\mu_i$  the usual spherically averaged PT kernels (Bernardeau et al. 2002)

$$v_i = i! \int d\Omega_1 \dots d\Omega_i F_i(\mathbf{k}_1, \dots, \mathbf{k}_i), \quad (\text{D14})$$

$$\mu_i = i! \int d\Omega_1 \dots d\Omega_i G_i(\mathbf{k}_1, \dots, \mathbf{k}_i). \quad (\text{D15})$$

In particular, one can check that in an Einstein–de Sitter Universe,  $v_1 = 1$ ,  $v_2 = 34/21$ ,  $v_3 = 682/189$ ,  $v_4 = 446440/43659$  and for the velocity divergence  $\mu_1 = 1$ ,  $\mu_2 = 26/21$ ,  $\mu_3 = 142/63$ ,  $\mu_4 = 236872/43659$ .

From equations (D12) and (D13), one can compute the rate function of  $\phi = \rho$  or  $\tilde{\theta}$  and by Legendre transform the corresponding one-cell CGF  $\varphi(\lambda)$ . In particular, one can check that the one-cell cumulants can be computed as

$$\langle \phi^{2+p} \rangle_c = \frac{\partial^{2+p} \varphi}{\partial \lambda^{2+p}}(\lambda = 0) = \left[ \frac{1}{\Psi''(\phi)} \frac{d}{d\phi} \right]^p \left( \frac{1}{\Psi''(\phi)} \right) \Big|_{\tau=0}$$

for all  $p \geq 0$ . Note that the condition  $\tau = 0$  will translate into either  $\rho = 1$  or  $\tilde{\theta} = 0$ . This procedure allows us to compute the successive cumulants of the density and the velocity divergence as a function of the spectral parameters  $\gamma_p$  and the spherical collapse parameters  $v_i$  and  $\mu_i$ . In particular, it allows us to recover the value of the skewness and kurtosis quoted in equations (D1)–(D4) and to compute all higher order cumulants.

The same procedure can be applied for multiple concentric cells in order to compute the mixed cumulants of densities and velocity divergences at different scales. Again relying on equations (D12) and (D13), the two-cell rate function and CGF  $\varphi(\lambda_1, \lambda_2)$  can be computed. The two-cell cumulants then follow

$$\begin{aligned} \langle \phi_1^{1+p} \phi_2^{1+q} \rangle_c &= \frac{\partial^{2+p} \varphi}{\partial \lambda^{2+p}}(\lambda_1 = 0, \lambda_2 = 0) \\ &= D_1^p D_2^q \left( \frac{-\Psi_{12}}{\det \Psi_{ij}} \right) \Big|_{\tau_i=0}, \end{aligned} \quad (\text{D16})$$

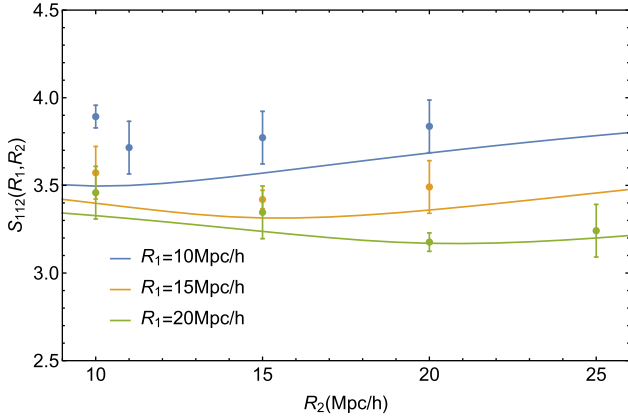
where the operators  $D_1$  and  $D_2$  are defined as follows:

$$D_1 = \left[ \frac{\Psi_{22}}{\det \Psi_{ij}} \frac{\partial}{\partial \phi_1} - \frac{\Psi_{12}}{\det \Psi_{ij}} \frac{\partial}{\partial \phi_2} \right], \quad (\text{D17})$$

$$D_2 = \left[ \frac{-\Psi_{12}}{\det \Psi_{ij}} \frac{\partial}{\partial \phi_1} + \frac{\Psi_{11}}{\det \Psi_{ij}} \frac{\partial}{\partial \phi_2} \right]. \quad (\text{D18})$$

In particular, following this approach, Bernardeau et al. (2014) showed that

$$\begin{aligned} \langle \rho_1^2 \rho_2 \rangle_c &= v_2 \sigma^2(R_1, R_2) (\sigma^2(R_1, R_2) + 2\sigma^2(R_1, R_1)) \\ &\quad + \frac{2}{3} \sigma^2(R_1, R_1) \frac{R_1 \partial}{\partial R_1} \sigma^2(R_1, R_2) \\ &\quad + \frac{1}{3} \sigma^2(R_1, R_2) \left( 2 \frac{R_2 \partial}{\partial R_2} \sigma^2(R_1, R_2) + \frac{R_1 d}{dR_1} \sigma^2(R_1, R_1) \right). \end{aligned}$$



**Figure D3.** Mixed third-order cumulants of the density field as defined in equation (D19) for  $R_1$  and  $R_2$  as labelled. The tree-order prediction is shown with solid lines and the measurements with error bars. Note that in this plot, we estimate the cumulants composed of two different scales from the measured PDF and we display fixed error bars of absolute size 0.15 as this is the typical error we make when using the PDF to estimate the skewness of the density rather than direct measurements in the simulation.

A similar result can now be obtained for the mixed cumulants of the velocity divergence field

$$\begin{aligned} \langle \tilde{\theta}_1^2 \tilde{\theta}_2 \rangle_c &= \mu_2 \sigma^2(R_1, R_2) (\sigma^2(R_1, R_2) + 2\sigma^2(R_1, R_1)) \\ &\quad + \frac{2}{3} \sigma^2(R_1, R_1) \frac{R_1 \partial}{\partial R_1} \sigma^2(R_1, R_2) \\ &\quad + \frac{1}{3} \sigma^2(R_1, R_2) \left( 2 \frac{R_2 \partial}{\partial R_2} \sigma^2(R_1, R_2) + \frac{R_1 \text{d}}{\text{d}R_1} \sigma^2(R_1, R_1) \right), \end{aligned}$$

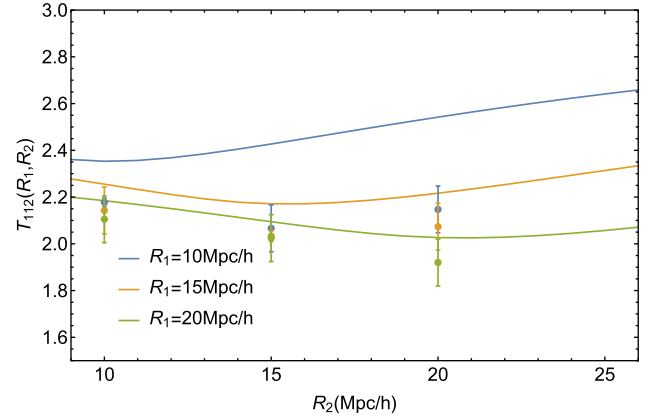
and for the mixed cumulants of the velocity divergence and density fields

$$\begin{aligned} \langle \rho_1 \tilde{\theta}_2^2 \rangle_c &= \sigma^2(R_1, R_2) (v_2 \sigma^2(R_1, R_2) + 2\mu_2 \sigma^2(R_2, R_2)) \\ &\quad + \frac{2}{3} \sigma^2(R_2, R_2) \frac{R_2 \partial}{\partial R_2} \sigma^2(R_1, R_2) \\ &\quad + \frac{1}{3} \sigma^2(R_1, R_2) \left( 2 \frac{R_2 \partial}{\partial R_2} \sigma^2(R_1, R_2) + \frac{R_1 \text{d}}{\text{d}R_1} \sigma^2(R_1, R_1) \right), \\ \langle \rho_1^2 \tilde{\theta}_2 \rangle_c &= \sigma^2(R_1, R_2) (\mu_2 \sigma^2(R_1, R_2) + 2v_2 \sigma^2(R_1, R_1)) \\ &\quad + \frac{2}{3} \sigma^2(R_1, R_1) \frac{R_1 \partial}{\partial R_1} \sigma^2(R_1, R_2) \\ &\quad + \frac{1}{3} \sigma^2(R_1, R_2) \left( 2 \frac{R_1 \partial}{\partial R_1} \sigma^2(R_1, R_2) + \frac{R_2 \text{d}}{\text{d}R_2} \sigma^2(R_2, R_2) \right). \end{aligned}$$

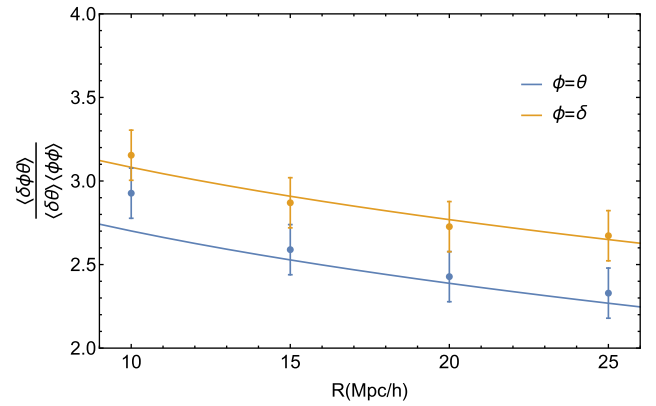
The above formulae allow us to compute the corresponding multi-scale reduced higher order moments. For instance, Fig. D3 compares measurements and tree-order predictions of the two-scale skewness of the density field defined as

$$S_{112} = \frac{3 \langle \rho_1^2 \rho_2 \rangle_c}{\sigma^2(R_1, R_2) (\sigma^2(R_1, R_2) + 2\sigma^2(R_1, R_1))}, \quad (\text{D19})$$

which is exactly the skewness,  $S_3$ , when  $R_1 = R_2$ . For radii  $R_1, R_2 \gtrsim 15 \text{ Mpc } h^{-1}$  at redshift 0, we find that measurements agree very well with the tree-order prediction. Below this scale, some departure starts to appear as non-linear corrections arise. Note that tree order predictions are computed by integrating the linear power spectrum such that



**Figure D4.** Same as Fig. D3 for the two-cell skewness of the velocity divergence. Here we assume error bars are of the order of 0.1 given the error we make in estimating the skewness from the PDF rather than from direct measurements.



**Figure D5.** Same as Fig. D3 for the mixed skewness of the velocity divergence and density at the same scale  $R$ . Here, we assume error bars are of the order of 0.15.

$$\begin{aligned} \sigma^2(R_1, R_2) &= \int \frac{\text{d}^3 k}{(2\pi)^3} P^{\text{lin}}(k) W_{3\text{D}}(kR_1) W_{3\text{D}}(kR_2), \\ R_1 \frac{\partial}{\partial R_1} \sigma^2(R_1, R_2) &= \int \frac{\text{d}^3 k}{(2\pi)^3} P^{\text{lin}}(k) \tilde{W}_{3\text{D}}(kR_1) W_{3\text{D}}(kR_2), \\ R_2 \frac{\partial}{\partial R_2} \sigma^2(R_1, R_2) &= \int \frac{\text{d}^3 k}{(2\pi)^3} P^{\text{lin}}(k) W_{3\text{D}}(kR_1) \tilde{W}_{3\text{D}}(kR_2), \end{aligned}$$

where the logarithmic derivative of the top-hat filter reads

$$\tilde{W}_{3\text{D}}(x) = -3 \left( W_{3\text{D}}(x) + \frac{\sin x}{x} \right). \quad (\text{D20})$$

The same comparison as in Fig. D3 is also performed for the velocity divergence field. The result is displayed in Fig. D4. The convergence towards the tree-order prediction seems slower for the velocity divergence than the density as we found consistent results for  $R_1, R_2 \gtrsim 20 \text{ Mpc } h^{-1}$  at redshift 0 in this case. This might come from the fact that velocities are more sensitive to previralization motions and therefore depart earlier from the tree-order expectations.

One can also compute the joint cumulants of velocities and densities in order to better assess the scatter in the density–velocity relation. As an illustration, we measure the third-order cumulants involving densities and velocities at the same scale and we compare them in Fig. D5 to their respective tree-order prediction:



$$\frac{\langle \rho \tilde{\theta} \tilde{\theta} \rangle_c}{\langle \delta \tilde{\theta} \rangle_c \langle \tilde{\theta} \tilde{\theta} \rangle_c} = v_2 + 2\mu_2 + \gamma(R) + \mathcal{O}(\sigma^2), \quad (\text{D21})$$

$$\frac{\langle \rho \rho \tilde{\theta} \rangle_c}{\langle \delta \tilde{\theta} \rangle_c \langle \rho \rho \rangle_c} = 2v_2 + \mu_2 + \gamma(R) + \mathcal{O}(\sigma^2). \quad (\text{D22})$$

Those mixed cumulants are very accurately fitted by their tree-order prediction [provided we divide by the right combination of variances in equations (D21) and (D22)].

All higher order mixed cumulants (for instance the two-scale kurtosis) can be obtained following the same scheme as used here for the two-scale skewness. Given the length of those formula, we do not show them in this paper.

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