Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **447,** 2683–2695 (2015) doi:10.1093/mnras/stu2632

Halo abundances within the cosmic web

D. Alonso,[1⋆] E. Eardley[2] and J. A. Peacock[2]

[1]*Astrophysics, University of Oxford, DWB, Keble Road, Oxford OX1 3RH, UK*
[2]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*

Accepted 2014 December 10. Received 2014 November 17; in original form 2014 June 16

ABSTRACT

We investigate the dependence of the mass function of dark-matter haloes on their environment within the cosmic web of large-scale structure. A dependence of the halo mass function on large-scale mean density is a standard element of cosmological theory, allowing mass-dependent biasing to be understood via the peak-background split. On the assumption of a Gaussian density field, this analysis can be extended to ask how the mass function depends on the geometrical environment: clusters, filaments, sheets and voids, as classified via the tidal tensor (the Hessian matrix of the gravitational potential). In linear theory, the problem can be solved exactly, and the result is attractively simple: the conditional mass function has no explicit dependence on the local tidal field, and is a function only of the local density on the filtering scale used to define the tidal tensor. There is nevertheless a strong *implicit* predicted dependence on geometrical environment, because the local density couples statistically to the derivatives of the potential. We compute the predictions of this model and study the limits of their validity by comparing them to results deduced empirically from *N*-body simulations. We have verified that, to a good approximation, the abundance of haloes in different environments depends only on their densities, and not on their tidal structure. In this sense we find relative differences between halo abundances in different environments with the same density which are smaller than $\sim$13 per cent. Furthermore, for sufficiently large filtering scales, the agreement with the theoretical prediction is good, although there are important deviations from the Gaussian prediction at small, non-linear scales. We discuss how to obtain improved predictions in this regime, using the 'effective-universe' approach.

Key words: large-scale structure of Universe.

1 INTRODUCTION

Environmental modifications of the properties of galaxies and galaxy systems is one of the central issues in the study of the formation of galaxies and large-scale structure (e.g. Abell 1965; Mo & White 1996; Peacock & Smith 2000; Berlind & Weinberg 2002; Berlind et al. 2003). The different observed degrees of clustering of galaxies of different types can be traced to this root cause, and the variation of galaxy bias with scale and with redshift can be understood quantitatively in this way. This standard piece of cosmological theory has two distinct elements, which are the galaxy contents of different dark-matter haloes, and how the properties of the halo population itself can vary with location within the overall density field. In the majority of published work, 'galaxy environment' refers purely to the former aspect, and is a shorthand for properties such as the mass of the halo hosting a given set of galaxies (e.g. Muldrew et al. 2012). In this paper, we shall be concerned with the larger-scale question – in effect defining environment via the density on scales of a few Mpc, rather than $<$100 kpc. Independent of the extent to which different haloes may generate different galaxy contents, a detailed understanding of the large-scale modulation of halo properties is a necessary preliminary. It has the virtue that it can be studied theoretically using *N*-body simulations of dark matter only. Furthermore, observational analogues of haloes can be obtained directly via catalogues of galaxy groups, so in principle there is much that can be learned about this aspect of cosmological structure without having to address the full complexity of galaxy formation.

There has been increasing recent interest in focusing the study of environmental modifications on what might be termed 'geometrical environment' – i.e. the location of a given galaxy or halo within the cosmic web (Doroshkevich 1970; Hahn et al. 2007; Forero-Romero et al. 2009; Hoffman et al. 2012; Cautun et al. 2014; Falck et al. 2014; Nuza et al. 2014; Metuki et al. 2015). This term captures the visually striking way in which the non-linear cosmic density field separates into distinct structures with different dimensionality, and normally four cases are considered: 3D voids and knots at either

⋆E-mail: david.alonso@astro.ox.ac.uk

extreme of overdensity, connected by 2D sheets and 1D filaments. A variety of different techniques have been proposed for constructing this decomposition. Some are theoretically meaningful, but hard to apply to galaxy survey data, such as the use of the peculiar velocity shear tensor (e.g. Libeskind et al. 2013; Metuki et al. 2015). At the opposite extreme, there are empirical methods that are connected very closely to the nature of the data in a given survey (e.g. Alpaslan et al. 2014), but which are hard to relate to a priori theory. There is also the question of whether we should be trying to study the 'skeleton' of the pattern of the cosmic web, as in Sousbie, Colombi & Pichon (2009), or providing a classification analysis of space – so that each point and each astronomical object can be allocated to a unique geometrical environment. We take the latter approach, and will also focus on the effect of this environment on the most primitive cosmic structures: the dark-matter haloes. The properties of galaxies within the cosmic web are perhaps more readily observed, but the predictions are more model dependent; however, a framework such as the halo model is widely used to predict the galaxy distribution once the spatial distribution of haloes is known, so it makes sense to start with this fundamental problem.

Accordingly, we summarize in Section 2 some standard results concerning the conditional mass function of haloes, and how this is modulated in regions of different large-scale overdensity; Section 3 introduces the classification of geometrical environment via the Hessian of the gravitational potential; Section 4 derives the impact of this definition of the cosmic web on the halo mass function; Section 5 contrasts these predictions with results from simulations and in Section 6 we present and discuss the main conclusions of this paper.

2 THE CONDITIONAL HALO MASS FUNCTION

Most of the information about the non-linear accretion of dark-matter haloes is encoded in the mass function $n(M)\,\mathrm{d}M$, the comoving number density of haloes with mass $M \in (M, M + \mathrm{d}M)$. Press & Schechter (1974, PS hereafter) developed a theoretical framework for calculating the mass function based solely on Gaussian statistics and non-linear spherical collapse. Their main result is that the collapsed fraction $F(>M)$ (the fraction of the Universe collapsed into virialized structures of masses larger than M) can be written as a universal (cosmology-independent) function of the variable

$$\nu_\mathrm{c} \equiv \frac{\delta_\mathrm{c}}{\sigma(M)}, \tag{1}$$

where $\delta_\mathrm{c} \simeq 1.686$ is the linear threshold for spherical collapse and $\sigma^2(M)$ is the variance of the density contrast filtered on a scale $R = (3M/4\pi\rho_0)^{1/3}$, given in linear perturbation theory by

$$\sigma^2(M) = \frac{1}{2\pi^2}\int_0^\infty P_k\,|W(kR)|^2\,k^2\,\mathrm{d}k, \tag{2}$$

where $W(kR)$ is the transform of a top-hat window function and P_k is the matter power spectrum. The PS result is

$$F(>M) = 1 - \mathrm{erf}\left(\frac{\nu_\mathrm{c}}{\sqrt{2}}\right), \tag{3}$$

from which the mass function can be calculated as

$$n(M)\,\mathrm{d}M = \frac{\rho_0}{M}\left|\frac{\mathrm{d}F(>M)}{\mathrm{d}\ln M}\right|\mathrm{d}\ln M. \tag{4}$$

Note that it is often convenient to express this in the form of the multiplicity function: $M^2 n(M)/\rho_0$, which is the fraction of the mass of the Universe carried by haloes in unit range of $\ln M$.

2.1 The excursion set formalism

The PS result was given a more solid mathematical foundation by Bond et al. (1991) and by Lacey & Cole (1993) through the so-called excursion set formalism. In this framework the density contrast at a given point as a function of the smoothing scale, $\delta(R)$, forms a random walk in δ–R space. The collapsed fraction can then be calculated as the fraction of all those random walks that, starting at $\delta(R \to \infty) \to 0$, make a first crossing above the δ_c threshold at some $R > R(M)$. In its original form, this formalism has one important caveat: for the walks to be completely random (i.e. with uncorrelated steps) the density contrast must be filtered using a top-hat window function in k-space, whereas the original PS result assumed top-hat filters in configuration space. There exist a few approaches in the literature to take into account the non-zero correlations between steps when using different window functions (Peacock & Heavens 1990; Maggiore & Riotto 2010; Musso, Paranjape & Sheth 2012).

The power of the excursion set formalism is that it can be extended to study the merging history of haloes as well as the conditional distribution of haloes in environments with different large-scale densities (Mo & White 1996), since both problems can be addressed by studying random walks crossing two barriers of different heights. Assuming Gaussian statistics one can compute the conditional probability for the density contrast smoothed over R_h to have a value δ_h given the value δ_e when smoothed on a scale R_e (here we will use the subscripts e for environmental quantities and h for haloes):

$$P(\delta_\mathrm{h}|\delta_\mathrm{e})\,\mathrm{d}\delta_\mathrm{h} = \frac{d\nu_\mathrm{h}}{\sqrt{2\pi(1-\varepsilon^2)}}\exp\left[-\frac{1}{2}\frac{(\nu_\mathrm{h} - \varepsilon\,\nu_\mathrm{e})^2}{1-\varepsilon^2}\right], \tag{5}$$

where $\nu_\mathrm{e} \equiv \delta_\mathrm{e}/\sigma_\mathrm{ee}$, $\nu_\mathrm{h} \equiv \delta_\mathrm{h}/\sigma_\mathrm{hh}$, $\epsilon \equiv \sigma^2_\mathrm{eh}/(\sigma_\mathrm{ee}\,\sigma_\mathrm{hh})$ and we have defined the covariance

$$\sigma^2_{ab} \equiv \frac{1}{2\pi^2}\int_0^\infty P_k\,W_a(kR_a)\,W_b^*(kR_b)\,k^2\mathrm{d}k. \tag{6}$$

Notice that different window functions may have been used for δ_h and δ_e. Equation (5) is in fact the same result found in the unconditional case with a rescaling of the variable

$$\nu_\mathrm{h} \to \frac{\nu_\mathrm{h} - \varepsilon\,\nu_\mathrm{e}}{\sqrt{1-\varepsilon^2}}, \tag{7}$$

and therefore the same rescaling applies to equation (3) to obtain the conditional collapsed fraction

$$F(>M|\delta_\mathrm{e}) = 1 - \mathrm{erf}\left(\frac{\nu_\mathrm{c} - \varepsilon\,\nu_\mathrm{e}}{\sqrt{2\,(1-\varepsilon^2)}}\right). \tag{8}$$

In both this and the unconditional problem, the random-walk approach solves the 'cloud-in-cloud' issue and yields a collapse fraction that is simply twice the area under the tail of the one-point density distribution; again, this simple result does not hold in the case of correlated trajectories. There exist other approaches in the literature that improve on this simple formalism to take into account the correlation between excursion steps as well as the peak nature of collapsed structures (Paranjape, Sheth & Desjacques 2013). We will not consider these here, since none of the trajectory-based approaches match 'exact' numerical results.

The original PS mass function in particular is well known to be a poor fit to N-body data, and more precise fitting formulae have been developed. These retain the main element of the PS philosophy by continuing to write the mass function with a universal dependence on the variable ν_c. In general, this works extremely well, although in detail slight deviations from this universal scaling do exist

(e.g. Sheth & Tormen 1999; Jenkins et al. 2001; Peacock 2007; Tinker et al. 2008; Watson et al. 2013). A more complex problem is what to do about the conditional mass function. Since the PS form for this uses the unconditional mass function with a change of variable from ν_c to $\nu_{\rm eff} \equiv (\nu_c - \varepsilon\,\nu_e)/\sqrt{1-\varepsilon^2}$, it is tempting to use the 'observed' functional dependence of the mass function on ν_c, substituting it by $\nu_{\rm eff}$. As we show explicitly in Section 5.3, this rescaling seems to be a good approximation only for large smoothing scales and mild environmental overdensities, and it breaks down for large values of δ_e. This comes as no surprise, since the same approach for the conditional mass function as a means to study the merging histories of haloes has been shown to fail to match N-body data (Cole et al. 2008).

As with the original PS approach, the excursion set method provides a framework for relating the linearized Gaussian density field to the abundance of virialized objects that form in the fully nonlinear density. For this reason, the environmental density entering equation (8) must be not the value of the local physical density Δ_e, but its linear extrapolation, which can be estimated using the prescription given by Bernardeau (1994):

$$1 + \Delta_e = \left(1 - \frac{\delta_e}{\delta_c}\right)^{-\delta_c}. \tag{9}$$

2.2 The effective-universe approach

The problem of the conditional mass function can also be approached from a different point of view. A well-known result of gravitational theory (which holds both in General Relativity and in the Newtonian limit) is that any spherically symmetric system behaves, at a fixed radius, in a manner equivalent to a spherical subregion of a homogeneous universe with some effective cosmological parameters. Thus, an overdense spherical region embedded in a Friedmann–Robertson–Walker (FRW) universe will have a radius-time history identical to that of a different effective FRW model (Weinberg 1972; Peebles 1993).

According to this result, and neglecting any error involved in applying it to non-spherical systems, we should be able to calculate an accurate conditional mass function for any degree of environmental overdensity, by exploiting the fact that the fitting formula for the unconditional mass function is claimed to be universal. Given a set of effective cosmological parameters, we should then be able to calculate the mass function in the effective universe – which is then equivalent to calculating the desired conditional mass function for environments of the same density. This approach has been discussed in previous works (Gottlöber et al. 2003; Goldberg & Vogeley 2004), and in fact has been shown to be equivalent to the excursion set method in the limit of large smoothing scales (Martino & Sheth 2009). We will outline here the main steps of this method, and refer the reader to Appendix A for a more detailed description.

Since the small-scale perturbations inside the effective universe have their origin in the same primordial power spectrum as perturbations in the background FRW model, the variance of the linear overdensity field in the effective universe is just a rescaled version of the variance outside:

$$\sigma_{\rm eff}(M) = D_g\,\sigma(M, R_e). \tag{10}$$

Here D_g is the ratio of the growth factor in the effective universe to the growth factor in the background cosmology, and $\sigma(M, R_e)$ is the variance of the density field in the background corrected for the size of the environment (see Appendix A, in particular equation A5, for further details). Thus, according to the PS theory, the mass function in this effective universe has the same functional form as the universal mass function (equation 3), using the rescaled variance in equation (10).

3 DEFINING THE COSMIC WEB

As we have seen, we can make predictions regarding the dependence of the mass function on the environment, when this is defined via the value of the density smoothed on a large scale – i.e. we can ask questions about the mass function in giant voids, and how this compares to the result within superclusters. For sharp k-space filtering, σ_{eh} is identical to σ_{ee}, and in the limit of large R_e, the correlation parameter ε becomes small. Thus the variable $\nu_{\rm eff}$ in the excursion set conditional mass function of equation (8) takes the value

$$\nu_{\rm eff} = \frac{\delta_h - \delta_e}{\sigma_{hh}}; \tag{11}$$

this is the peak-background split, in which the large-scale density fluctuation is treated as an external offset to the density. This shift makes it easier to attain the $\delta_h = \delta_c$ threshold and for more haloes to collapse in high-density regions. But although this phenomenon is of great interest in explaining the origin of biased clustering, it is not necessarily the best way of classifying environments if we actually want to observe the peak-background split in action. Examination of images of redshift surveys or N-body simulations shows that the majority of the mass condenses into geometrical features with low dimensionality and a small filling factor: the cosmic web. To understand biased clustering in detail, it will be interesting to look at how the halo population changes within these different environments. This means that we need to think about more than the mean density defined with some large isotropic filter; such an approach ignores e.g. the fact that filaments are relatively narrow features in some directions.

A number of approaches have been taken to the problem of making a geometrical classification of different parts of the cosmic web (Hahn et al. 2007; Forero-Romero et al. 2009; Hoffman et al. 2012); we have chosen to follow Forero-Romero et al. (2009) and use the tidal (or deformation) tensor. This is defined as $T_{ij} \equiv \partial_i \partial_j \tilde{\phi}$, i.e. the Hessian of the Newtonian potential,[1] and therefore the density contrast is directly given by Poisson's equation as its trace $\delta = T_{11} + T_{22} + T_{33}$. Note that we have normalized the usual definition of the Newtonian potential

$$\tilde{\phi} \equiv \phi/(4\pi G\bar{\rho}), \tag{12}$$

to make this relation simpler.

We will classify a given point in space as belonging to one of the different elements of the cosmic web according to the number of eigenvalues of $\hat{T}$ above a given threshold λ_{th} at that point. Thus, let $\lambda_1 \geq \lambda_2 \geq \lambda_3$ be the three eigenvalues of $\hat{T}$. We define:

(i) **Voids**: all eigenvalues below the threshold ($\lambda_1 < \lambda_{th}$).
(ii) **Sheets**: only one eigenvalue above the threshold ($\lambda_1 > \lambda_{th}$, $\lambda_2 < \lambda_{th}$).
(iii) **Filaments**: two eigenvalues above the threshold ($\lambda_2 > \lambda_{th}$, $\lambda_3 < \lambda_{th}$).
(iv) **Knots**: all eigenvalues above the threshold ($\lambda_3 > \lambda_{th}$).

[1] In this paper all spatial derivatives ∂_i are defined with respect to physical coordinates $\boldsymbol{r}$ and not comoving ones $\boldsymbol{q} \equiv \boldsymbol{r}/a$. This removes the proportionality factor a^2 from the relation between $\tilde{\phi}$ and δ and makes the notation less cumbersome.

In this approach, the eigenvalue threshold is an arbitrary quantity. One might be tempted to set it to zero and classify based on the sign of the eigenvalues, but this is not very attractive from the point of view of gravitational collapse. We have seen that the density contrast is the sum of the eigenvalues, and we know that non-linear collapse involves a linear contrast of order unity. It is therefore natural to consider $\lambda_{\rm th}$ of order unity; we refine this choice below.

A different prescription, the so-called V-web approach (Hoffman et al. 2012), uses instead the eigenvalues of the velocity shear tensor $\Sigma_{ij} \equiv -(\partial_i v_j + \partial_j v_i)/2\,H$. Although this prescription allows for a finer discrimination on very small scales, Σ_{ij} converges to $f_{\rm g}\,T_{ij}$ in linear theory[2] (where $f_{\rm g} \equiv {\rm d}\ln\delta/{\rm d}\ln a$), and therefore similar results should hold. Note that in practice we would need to define the web using a redshift survey. If we work from the galaxy density and use a quasi-potential obeying $\nabla^2\tilde{\phi} = -1 + \rho_{\rm g}/\langle\rho_{\rm g}\rangle$, the Hessian approach can still be used – but the velocity shear tensor is not observable.

It is useful to check that the above description makes intuitive sense, via a couple of analytic examples. First let us consider a spherical model with $\tilde{\phi} = -a/(r^2+b^2)^{1/2}$. For this, $T_{ij} = a(r^2+b^2)^{-5/2}[(r^2+b^2)\delta_{ij} - 3x_i x_j]$, and we can assess the eigenvalues by looking at the diagonal form in the frame $(x, y, z) = (r, 0, 0)$. For $r < b/\sqrt{2}$, there are three positive eigenvalues that coincide at $r = 0$, so there is a knot there if a is large enough (otherwise it is a void everywhere). Since the first diagonal entry is smaller, this spherical knot must be surrounded by a shell where the field is classified as a filament, with classification changing to a void at large enough r (even though the density perturbation is positive everywhere). The prediction of a filament region is a little surprising, but sensible: away from the centre of the perturbation, tidal forces act on an extended object by 'squeezing' it in the two transverse directions and stretching it in the radial one. More intuitive results arise from lower dimensional versions of the same model: $\tilde{\phi} = -a/(xr + x^2 + b^2)^{1/2}$ makes a transition from a filament to a void, whereas $\tilde{\phi} = -a/(x^2+b^2)^{1/2}$ makes a transition from a sheet to a void. In general terms, tidal forces act on an extended object through the second derivatives of $\tilde{\phi}$:

$$\ddot{\delta x}_i = -T_{ij}\,\delta x_j, \tag{13}$$

and hence this prescription classifies different regions according to the number of directions in which tidal forces tend to contract or stretch an extended object.

4 GAUSSIAN STATISTICS IN THE COSMIC WEB

It is interesting to study the statistics of the cosmic web in the limit that $\tilde{\phi}$ is a Gaussian random field. We will also assume that the potential is smoothed on a scale $R_{\rm e}$ with some window function $W_{\rm e}(kR_{\rm e})$. In this case the probability density function of the eigenvalues λ_i is given by

$$\begin{aligned} {\rm d}P(\lambda_i) &\equiv\; p(\rho,\theta,\nu)\,{\rm d}\rho\,{\rm d}\theta\,{\rm d}\nu \qquad (14)\\ &\equiv 225\sqrt{\frac{5}{2\pi}}\rho\,(\rho^2-\theta^2)\,\exp\left[-\frac{1}{2}(15\rho^2+5\theta^2)\right]\\ &\quad\times\frac{{\rm e}^{-\nu^2/2}}{\sqrt{2\pi}}\,{\rm d}\rho\,{\rm d}\theta\,{\rm d}\nu, \qquad (15)\end{aligned}$$

[2] Note also that the normalized potential $\tilde{\phi}$ is also linearly related to the potential of the displacement vector $\boldsymbol{\Psi} \equiv \nabla\phi_d$ via $\tilde{\phi} = a^2\phi_d$.

where we have defined the variables $\nu \equiv (\lambda_1+\lambda_2+\lambda_3)/\sigma_{\rm ee} \equiv \delta_{\rm e}/\sigma_{\rm ee}$, $\rho \equiv (\lambda_1-\lambda_3)/2\sigma_{\rm ee}$ and $\theta \equiv (\lambda_1 - 2\lambda_2 + \lambda_3)/2\sigma_{\rm ee}$. The derivation of this result is outlined in Appendix B and has been widely used in the literature (Doroshkevich 1970; Bardeen et al. 1986; Bond & Myers 1996; Rossi 2013). Note that these results are usually quoted in terms of the ellipticity e and prolateness p which are related to our variables by $e \equiv \rho/\nu$, $p \equiv \theta/\nu$.

The restriction $\lambda_1 \geq \lambda_2 \geq \lambda_3$ implies $\rho \in [0,\infty)$, $\theta \in [-\rho, \rho]$, and by selecting a given type of environment we further constrain the dynamical range of ν. Let $\lambda_{\rm th}$ be the eigenvalue threshold used for our classification, and let α denote the number of eigenvalues above this threshold for each case. Then the integration limits for ν are $f_1^\alpha(\rho,\theta) < \nu - \nu_{\rm th} < f_2^\alpha(\rho,\theta)$, with $\nu_{\rm th} \equiv 3\lambda_{\rm th}/\sigma_{\rm ee}$ and

$$f_1^\alpha(\rho,\theta) = \begin{cases} -\infty, & \alpha = 0\ \text{(voids)}\\ -3\rho-\theta, & \alpha = 1\ \text{(sheets)}\\ 2\theta, & \alpha = 2\ \text{(filaments)}\\ 3\rho-\theta, & \alpha = 3\ \text{(knots)}\end{cases}, \tag{16}$$

$$f_2^\alpha(\rho,\theta) = \begin{cases} -3\rho-\theta, & \alpha = 0\ \text{(voids)}\\ 2\theta, & \alpha = 1\ \text{(sheets)}\\ 3\rho-\theta, & \alpha = 2\ \text{(filaments)}\\ \infty, & \alpha = 3\ \text{(knots)}\end{cases}. \tag{17}$$

The volume fraction in environments of type $(\alpha, \nu_{\rm th})$ can be calculated as the probability for a random point in space to belong to that type:

$$F_{\rm V}(\alpha,\nu_{\rm th}) \equiv \int_0^\infty {\rm d}\rho \int_{-\rho}^{\rho} {\rm d}\theta \int_{\nu_{\rm th}+f_1^\alpha(\rho,\theta)}^{\nu_{\rm th}+f_2^\alpha(\rho,\theta)} {\rm d}\nu\; p(\rho,\theta,\nu). \tag{18}$$

Note that the volume fraction is a universal function of the normalized threshold $\nu_{\rm th}$.

4.1 Dependence on the environmental density

We aim to give a prediction for the abundance of haloes of different mass in different types of environment. In the excursion set approach, this abundance is given by the number of first upcrossings of the collapse threshold by the density contrast field smoothed over a scale $R_{\rm h}$, corresponding to the halo mass. Therefore we first need to study the joint probability for the density contrast $\delta_{\rm h}$ smoothed over this scale and the environmental tidal tensor eigenvalues, filtered on a scale $R_{\rm e}$. As shown in Appendix B, this distribution is

$$\frac{{\rm d}P(\delta_{\rm h},\lambda_i)}{{\rm d}\nu_{\rm h}\,{\rm d}\nu_{\rm e}\,{\rm d}\rho\,{\rm d}\theta} = \frac{p(\rho,\theta,\nu_{\rm e})}{\sqrt{2\pi(1-\varepsilon^2)}}\exp\left[-\frac{(\nu_{\rm h}-\varepsilon\,\nu_{\rm e})^2}{2\,(1-\varepsilon^2)}\right], \tag{19}$$

where $\nu_{\rm e} \equiv \delta_{\rm e}/\sigma_{\rm ee}$.

As a first step, we are interested in the abundance of haloes within environments of type $(\alpha, \nu_{\rm th})$ with local environmental density contrast $\delta_{\rm e}/\sigma_{\rm ee} \in (\nu_{\rm e}, \nu_{\rm e} + {\rm d}\nu_{\rm e})$. In this case we integrate over ρ and θ to find the joint distribution

$$\begin{aligned} P(\nu_{\rm h},\nu_{\rm e},\alpha,\nu_{\rm th})\,{\rm d}\nu_{\rm e}\,{\rm d}\nu_{\rm h} &= \frac{{\rm d}\nu_{\rm e}}{\sqrt{2\pi}}\,C(\nu_{\rm e})\,{\rm e}^{-\nu_{\rm e}^2/2}\\ &\quad\times\frac{{\rm d}\nu_{\rm h}}{\sqrt{2\pi(1-\varepsilon^2)}}\exp\left[-\frac{(\nu_{\rm h}-\varepsilon\,\nu_{\rm e})^2}{2\,(1-\varepsilon^2)}\right]. \qquad (20)\end{aligned}$$

Here

$$C(\nu_{\rm e}) \equiv 225\sqrt{\frac{5}{2\pi}}\iint_{\mathcal{A}(\nu_{\rm e})} {\rm d}\rho\,{\rm d}\theta\,\rho\,(\rho^2-\theta^2)\,{\rm e}^{-\frac{1}{2}(15\rho^2+5\theta^2)} \tag{21}$$

and $\mathcal{A}(\nu_e)$ is the region in the plane (θ, ρ) defined by the conditions

$$\rho \geq 0, \quad -\rho \leq \theta \leq \rho,$$
$$f_1^\alpha(\rho,\theta) \leq \nu_e - \nu_{th}, \leq f_2^\alpha(\rho,\theta). \tag{22}$$

On the other hand, the probability of finding such an environment is given by

$$P(\nu_e, \alpha, \nu_{th})\,d\nu_e = \frac{d\nu_e}{\sqrt{2\pi}}\,C(\nu_e)\,e^{-\nu_e^2/2}, \tag{23}$$

which cancels out when computing the conditional probability

$$P(\nu_h|\nu_e, \alpha, \nu_{th})\,d\nu_h = P(\delta_h|\delta_e)\,d\delta_h, \tag{24}$$

where $P(\delta_h|\delta_e)$ is given in equation (5).

From this we can extract a general prediction of this formalism, which is a key result of this paper: the only environmental quantity that determines the abundance of haloes is the local density $\delta_e \equiv \lambda_1 + \lambda_2 + \lambda_3$. This conclusion arises from the fact that our treatment is based on the spherical top-hat collapse, which disregards all couplings of the halo orientations with other combinations of the eigenvalues of the tidal tensor (i.e. the only non-zero, non-diagonal element of the covariance matrix is between ν_h and ν_e). While this assumption of zero coupling of gravitational collapse to tidal forces could be challenged in detail, we find it striking that the geometrical environment is not predicted to have a more direct influence on the properties of haloes at a given overdensity.

In the remainder of this paper, we therefore subject our result to the test of confrontation with numerical simulations. Any evidence for an explicit dependence on geometrical environment would be interesting as it would relate to the issue of halo 'assembly bias' (e.g. Gao & White 2007). This term stands for the possibility that halo properties have some dependence on parameters beyond the local overdensity; the concept may apply either to the final-state properties of the halo or to its merger history (which potentially influences the galaxy contents of the halo). This issue is certainly under active consideration from the point of view of the observational dependence of galaxy properties on tidal forces (e.g. Yan, Fan & White 2013). It is worth noting that our analysis is focused on overall halo abundances and not intrinsic halo properties, and hence it does not directly address the problem of assembly bias.

4.2 The four mass functions

Using the key result of the previous subsection, any halo statistic conditional to a given type of environment can be calculated within the excursion set formalism as the average of that statistic conditional to a given background density in that environment. For example, for the conditional collapsed fraction

$$F(>M|\alpha, \nu_{th}) = \int_0^\infty d\rho \int_{-\rho}^{\rho} d\theta \int_{\nu_{th}+f_1^\alpha(\rho,\theta)}^{\nu_{th}+f_2^\alpha(\rho,\theta)} d\nu_e \times \frac{p(\rho,\theta,\nu)\,F(>M|\delta_e \equiv \sigma_{ee}\nu_e)}{F_V(\alpha,\nu_{th})}, \tag{25}$$

where $F(>M|\delta_e)$ is obtained through either of the methods outlined in Section 2.

5 COMPARISON WITH SIMULATIONS

The analytical predictions explained in the previous sections have been compared with numerical data from the MultiDark Run 1 (MDR1) dark-matter N-body simulation (Prada et al. 2012). MDR1 simulates a $1\,(h^{-1}\,\mathrm{Gpc})^3$ cubic volume with a mass resolution of $m_p = 8.721 \times 10^9\,h^{-1}\,M_\odot$ in a Λ cold dark matter cosmology with $(\Omega_m = \Omega_{dm} + \Omega_b,\ \Omega_\Lambda,\ \Omega_b,\ h,\ n,\ \sigma_8) = (0.27,\ 0.73,\ 0.0469,\ 0.7,\ 0.95,\ 0.82)$. The corresponding halo catalogue was compiled using a friends-of-friends algorithm on the $z = 0.1$ snapshot, yielding a minimum mass cut of $M_{min} = 10^{11.5}\,h^{-1}\,M_\odot$.

The density field used to classify the different environments was computed by interpolating the dark-matter particle content on to a grid of spacing $a = 3.9\,h^{-1}$ Mpc by cloud in cell interpolation. This field was then smoothed using a Gaussian filter of variable width

$$W(k) = \exp(-S^2k^2/2); \tag{26}$$

we denote the width of the Gaussian by S, since the common symbol σ is used elsewhere. We choose to use Gaussian filtering for numerical stability in preference to spatial top-hat filtering. The results are closely equivalent to using a top-hat of radius $R \simeq \sqrt{5}S$. We show results for two different filtering scales, $S = 4$ and $10\,h^{-1}$ Mpc. The deformation tensor was calculated at each point in the grid by solving Poisson's equation in Fourier space and then transforming back to configuration space (making extensive use of fast Fourier transforms throughout the process). Fig. 1 shows an example of the density field and environment classification obtained for MDR1.

5.1 The eigenvalue threshold

The value of the eigenvalue threshold to be used in this type of analyses is somewhat arbitrary. One may choose $\lambda_{th} = 0$ on the basis that this value discriminates between purely compressing or stretching tidal forces. This choice, however, yields a low abundance of voids, compared to what one would expect from a visual examination of a large-scale structure map, and previous studies (e.g. Forero-Romero et al. 2009; Metuki et al. 2015) have chosen a threshold in order to yield a better match to the visually expected volume fractions. Ultimately we would like to be able to extract the maximum amount of information from the abundance of haloes in these four environments. While it is possible to improve the halo statistics by resampling techniques such as the one proposed by Aragon-Calvo (2012), for a limited amount of data it is possible to use the value of the eigenvalue threshold to make sure that we obtain sufficient statistics for all of the different environments. In order to do so we have computed the fraction of the total halo mass contained in each environment for different values of S and λ_{th}, $FM_\alpha(S, \lambda_{th})$, and the root mean square deviation of these fractions $FM_{rms}(S, \lambda_{th}) = \left[\sum_\alpha (FM_\alpha - \overline{FM})^2/4\right]^{1/2}$. We have then chosen the optimal value of λ_{th} for each smoothing scale S as the one that minimizes FM_{rms}. Besides this consideration, it is also important to make sure that the values of λ_{th} used yield physically sensible definitions for the different environments. One can then try to combine the aforementioned method with extra requirements, such as limiting the fraction of cells classified as voids that are overdense.

Fig. 2 shows the values of FM_{rms} for different smoothing scales and eigenvalue thresholds computed from the simulation (left-hand panel). In view of this figure we have chosen two eigenvalue thresholds for each of the two smoothing scales:

$$S = 4\,h^{-1}\,\mathrm{Mpc} \longrightarrow \lambda_{th} = 0.4,\ 0.6, \tag{27}$$

$$S = 10\,h^{-1}\,\mathrm{Mpc} \longrightarrow \lambda_{th} = 0.1,\ 0.3. \tag{28}$$

In both cases the first value was chosen by restricting the fraction of overdense void cells to be smaller than 10 per cent, while we have chosen a second slightly larger threshold to illustrate how

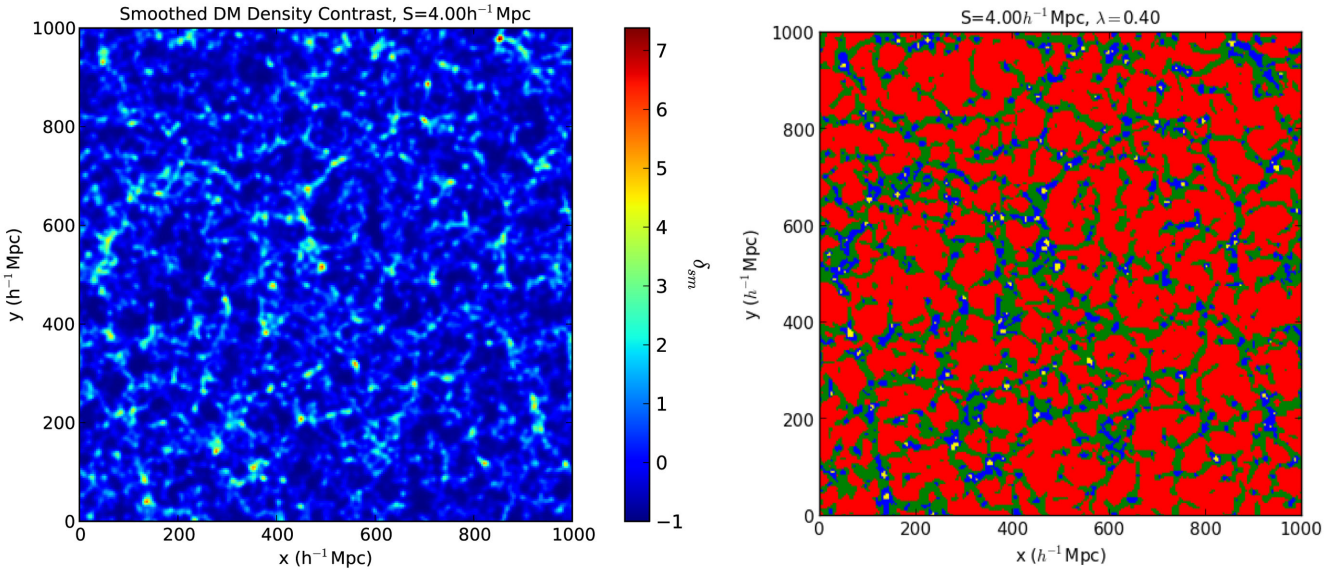

Figure 1. Density field (left) and environment classification (right) of a slice of the MDR1 simulation. The colour code is voids (red); sheets (green); filaments (blue); knots (yellow). A Gaussian filter with $S = 4\,h^{-1}$ Mpc and a threshold $\lambda_{\rm th} = 0.4$ were used.

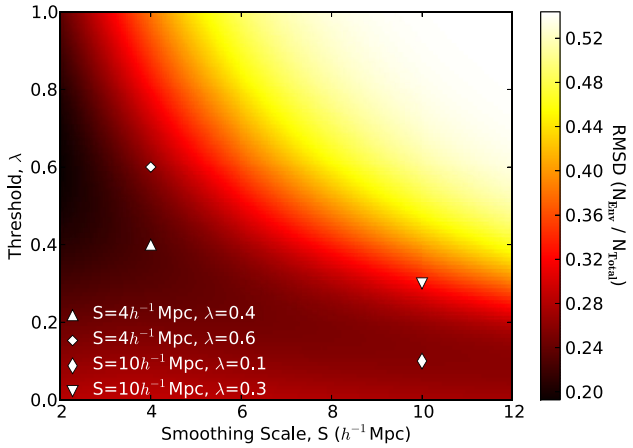

Figure 2. Illustrating how the partition of the cosmic mass density between different geometrical environments varies with smoothing scale and threshold. A practically useful partition will place approximately equal quantities of mass in the four environments, and the colour scale shows the dispersion in these mass fractions. The optimum is approximately $\lambda = 0.4$ for $S = 4\,h^{-1}$ Mpc and $\lambda = 0.1$ for $S = 10\,h^{-1}$ Mpc (shown as points); but we also consider other thresholds, 0.6 and 0.3 respectively, to illustrate how our results depend on the choice of these parameters.

our results depend on this choice. Note also that the combinations $(S,\ \lambda_{\rm th}) = (10\,h^{-1}\,\text{Mpc},\ 0.1)$ and $(4\,h^{-1}\,\text{Mpc},\ 0.4)$ are similar to the values used in the analysis of the Galaxy And Mass Assembly (GAMA) data by Eardley et al. (2014).

It is useful to verify the meaning of this choice by looking at the volume and mass fractions in each environment as a function of $\lambda_{\rm th}$. This is shown in Fig. 3. As can be seen, our choice of eigenvalue threshold tries to maximize statistics of the four environments simultaneously. The figure also shows the Gaussian-field prediction for the volume and mass fractions within the formalism described in Section 4. The match is rather good for the larger value of the smoothing scale.

5.2 Conditional density distributions

According to the theoretical framework described in Section 4, the halo mass function should be the same in environments with the same density, depending only implicitly on the environment classification due to the different distribution of densities for each environment type. Therefore, our ability to predict the abundance of haloes in each element of the cosmic web depends, on the one hand, on reproducing these distributions correctly and, on the other, on the accuracy of our model for the conditional mass function. These distributions are given, according to our formalism, by

$$P(\delta_{\rm e}|\alpha,\,\nu_{\rm th})\mathrm{d}\delta_{\rm e} \equiv \frac{P(\nu_{\rm e},\,\alpha,\,\nu_{\rm th})}{F_{\rm V}(\alpha,\,\nu_{\rm th})}\mathrm{d}\nu_{\rm e}, \tag{29}$$

where $P(\nu_{\rm e},\,\alpha,\,\nu_{\rm th})$ is given by equation (23).

This quantity is shown in Fig. 4 as measured from the MDR1 Simulation for different values of S and $\lambda_{\rm th}$. The two top panels in this figure show how the density distributions measured directly from the simulation (continuous histogram) are very poorly described by the Gaussian theory, especially for high-density environments. This is not so surprising, since it is well known that the matter density is significantly non-linear on our filtering scales. In order to understand these differences better, we have compared these results with the predictions of the lognormal distribution (Coles & Jones 1991), which has often been used as a convenient approximate model for the non-linear density field. Specifically, we used the real density field to perform the environment classification, and then we studied the distribution within each environment of the field that results from undoing the lognormal transformation. This is given in terms of the real overdensity $\delta_{\rm r}$ and its variance $\sigma_{\rm r}^2$ as

$$\delta_G = \ln\left[(1+\delta_{\rm r})\sqrt{1+\sigma_{\rm r}^2}\right]. \tag{30}$$

The distribution of this 'de-lognormalized' overdensity is shown in the bottom panels of Fig. 4. We can see that this field follows the Gaussian theoretical distribution much better for the largest smoothing scale $S = 10\,h^{-1}$ Mpc. For smaller scales, however, the lognormal transformation is no longer a good description of the

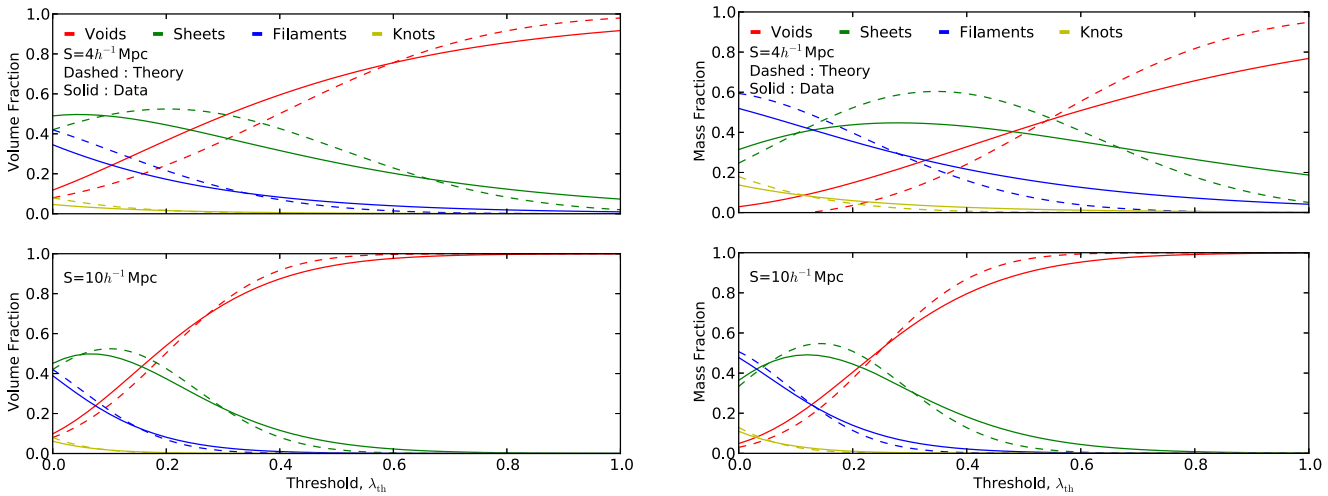

Figure 3. Volume (left) and mass (right) fractions in the four environments for $S = 4, 10\,h^{-1}$ Mpc as a function of the eigenvalue threshold, measured from the MDR1 simulation (dashed lines). The results are shown for voids (red), sheets (green), filaments (blue) and knots (yellow). The Gaussian-field prediction for the volume fractions are also shown as solid lines for comparison.

non-Gaussianity of the density field, a result that has been reported by other authors (e.g. Kitaura, Jasche & Metcalf 2010).

5.3 Halo abundances

The overall multiplicity function in voids, sheets, filaments and knots is calculated by averaging $f(M|\delta_e)$ over δ_e using the overdensity distribution of each environment. As we showed in the previous section, the Gaussian prediction for the density distribution $P(\delta_e|\alpha, \nu_{th})$ is not a good description of the real distribution in most cases, even after attempting to account for non-Gaussianity using the log-normal approximation. Hence, even if we had an accurate model for the conditional mass function, we would still not be able to predict $f(M|\alpha, \nu_{th})$ correctly. For this reason, in order to isolate the inaccuracies due to the incorrect modelling of the conditional mass function from those due to the Gaussian approximation, we have integrated over the actual density distributions measured from the MDR1 simulation (solid lines in Fig. 4) in order to obtain a theoretical prediction for the four mass functions. These are shown in Fig. 5 for the cases quoted in equation (28) together with the theoretical predictions for the excursion set and the effective-universe approaches. These predictions are based on rescaling the universal collapsed fraction, which was estimated using the fitting formula proposed by Peacock (2007). Our results are in qualitative agreement with Hahn et al. (2007), who used an eigenvalue threshold $\lambda_{th} = 0$. As is discussed in the next section, the excursion set formalism is only able to make reasonable predictions for environments involving small overdensities (i.e. voids and sheets) and for large filter scales, while the effective-universe approach shows an overall better agreement with the data. Nevertheless, neither model is able to describe the data accurately. For the present, we have to be content that we have an approximate understanding of the trends in halo properties with environment; accurate work will require calibration from numerical simulation, just as with the original PS mass function.

5.4 Universality of density dependence

The above results show that the excursion set approach is able to make relatively good predictions for large smoothing scales and mild environmental densities, but that it fails to do so for smaller values of S and high δ_e. This is a reasonable result: the excursion set model is based on following trajectories in the δ–R plane that, starting at some (R_e, δ_e), cross the threshold δ_c at some scale $R(M) < R_e$. In this regime the excursion set predicts that too many small-mass haloes have already merged into larger ones, due to the fact that the large environmental density makes gravitational collapse more efficient. At the large-mass end, on the other hand, too few haloes have formed, since the total halo mass is limited by the mass that can be found within the smoothing radius. Also, for small R_e the correlation between adjacent steps, which is generally ignored, may play a significant role, since the scale of the halo mass can be close to the scale of the environment. In these limits, the effective-universe approach outperforms the excursion set, providing a better description for the conditional mass function – although far from being per cent level precise.

Nevertheless, one of the most important predictions of our formalism is that the abundance of haloes should be the same for all environments with the same background density. Regardless of whether or not the theoretical prediction for the conditional mass function is quantitatively precise, it is interesting to test the qualitative validity of this result according to the simulated data.

Fig. 6 shows the multiplicity function ($f(M) \equiv |\mathrm{d}\,F(<M)/\mathrm{d}\ln M|$) for haloes residing in the four different environments with restricted local overdensities, together with the environment-independent theoretical predictions derived from the excursion set (dashed lines) and the effective-universe approach (solid lines) as detailed in Section 2. The values of the local overdensity were chosen to guarantee the simultaneous presence of as many different environments as possible. For the range of masses, smoothing scales and densities explored, we find that the prediction that the abundance of haloes should depend only on the environmental density and not on the environment classification holds very well, with little or no deviation within statistical errors. We have quantified this agreement as follows: for each overdensity bin we take all pairs of multiplicity functions that we have been able to calculate in it. Assuming Gaussian statistics, we estimate the probability that both multiplicity functions are compatible in each mass bin given their statistical uncertainties. We then quantify the agreement between multiplicity functions by computing the relative difference between them in the

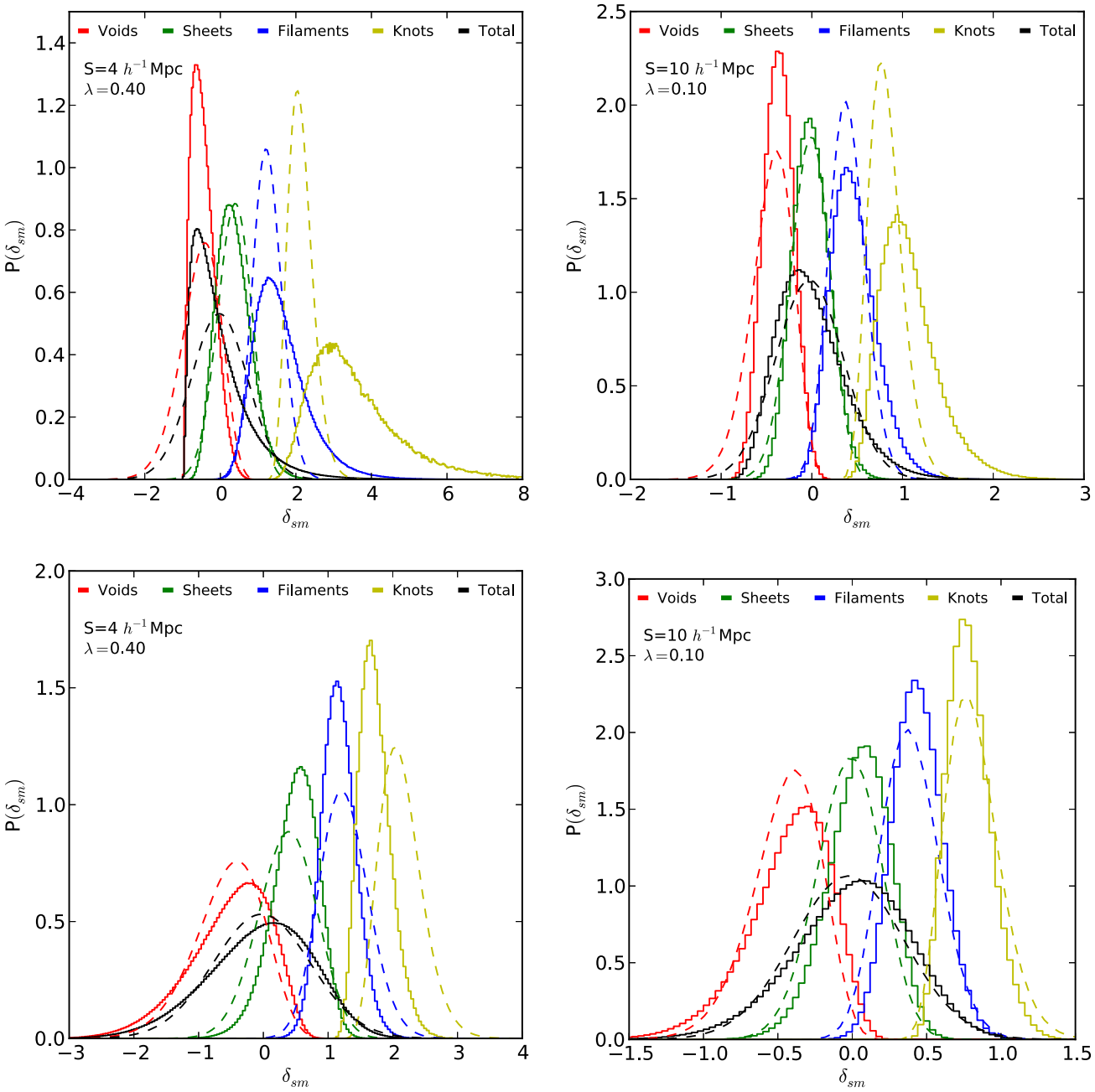

Figure 4. Overdensity distributions in each of the four environments and the overall overdensity distribution for $(S, \lambda_{\rm th}) = (4\,h^{-1}\,{\rm Mpc}, 0.4)$ – left-hand panels – and $(S, \lambda_{\rm th}) = (10\,h^{-1}\,{\rm Mpc}, 0.1)$ – right-hand panels. The dashed lines show the Gaussian theoretical prediction, while the solid histograms show the distributions extracted from the MDR1 simulation. In the top panels these histograms correspond to the distribution of the real density field, while the bottom panels show the distribution of the 'de-lognormalized' overdensity (see equation 30). The colour code is voids (red); sheets (green); filaments (blue); knots (yellow); overall distribution (black). In the same order, the density distribution for the four environments peaks on increasing values of $\delta_{\rm sm}$.

mass bin with the smallest p-value. The results are shown in Table 1 for the bins of overdensity explored above. In the worst case, the largest deviation is about 13 per cent. However it is worth noting that, even in this case, both multiplicity functions are fully compatible, with a minimum p-value of 0.32, and in all other cases the values of the different multiplicity functions are compatible within 2σ.

Even though we have verified the prediction that the halo abundances in different environments depend only on the environmental density, the exact dependence of these abundances on halo mass is not reproduced accurately by either the effective-universe approach or the excursion set formalism, although they both qualitatively follow the same trend as the data.

6 SUMMARY AND DISCUSSION

We have considered the statistics of dark-matter haloes within the cosmic web, using the eigenvalues of the Hessian of the potential to

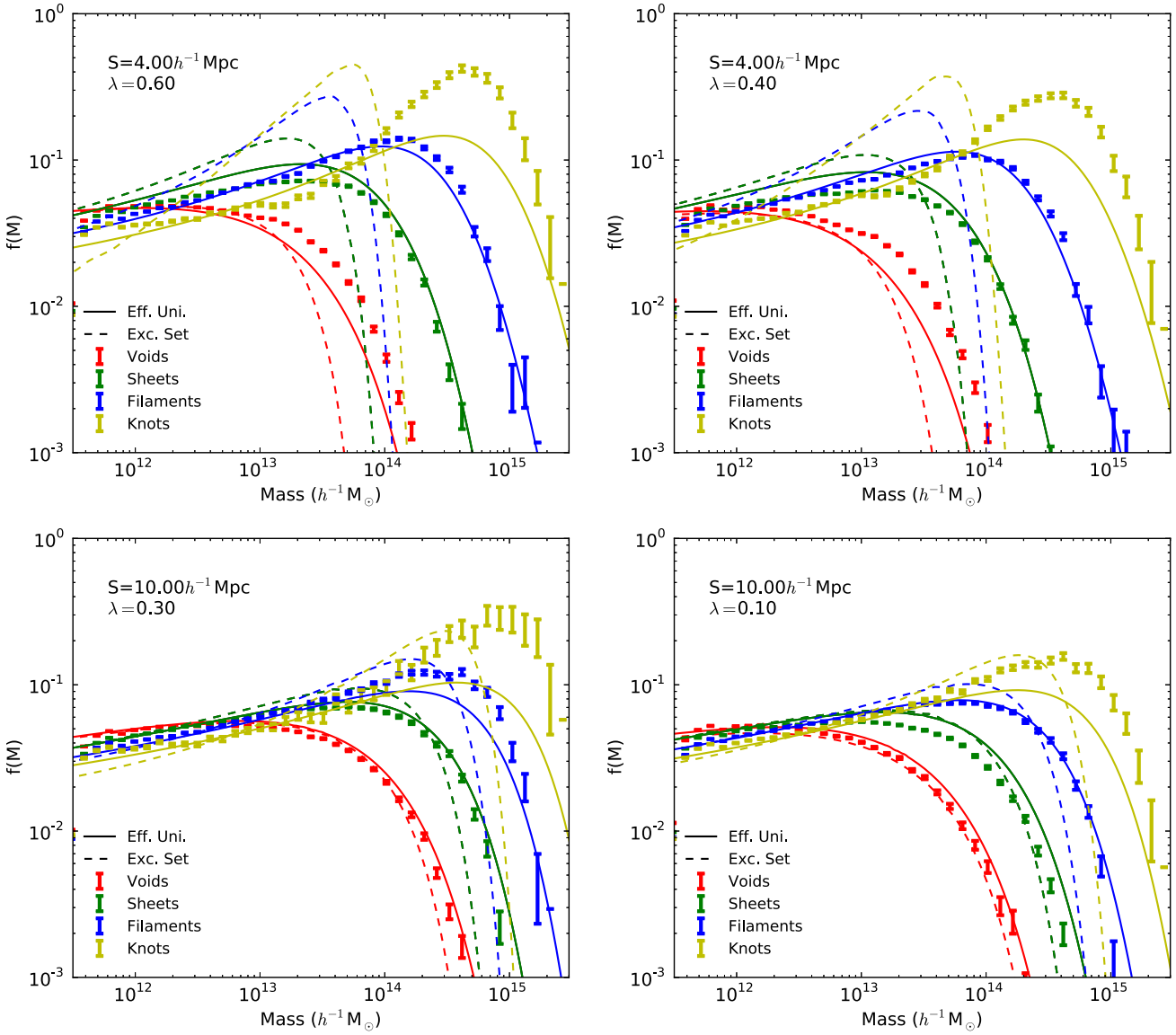

Figure 5. Multiplicity function for the four environments for our fiducial smoothing scales and eigenvalue thresholds. In increasing order of amplitude on high masses, the different multiplicity functions correspond to voids, sheets, filaments and knots. Reasonable agreement is only obtained for large smoothing radii and small $\lambda_{\rm th}$. This can be attributed to the fact that the standard prediction for the conditional mass function is only valid for densities below the collapse threshold and masses below the filter scale. The theoretical prediction was made by rescaling the empirical formula in Peacock (2007) for the mass function.

classify regions of space into one of four geometrical environments. The main results from this work are:

(1) Assuming the density contrast field to be Gaussian, clear predictions can be made regarding the abundance (i.e. volume and mass fractions) of the different environments classified according to the tidal tensor prescription. These are reasonable approximations for large smoothing scales and can be used to select eigenvalue thresholds that are useful for practical comparisons – partitioning the Universe nearly equally between the four environments.

(2) We have compared the simulated halo abundances with the predictions within the excursion set formalism and the effective-universe approach. Neither of these approaches are able to yield quantitatively precise results. However, the effective-universe picture provides an overall better description, especially for small smoothing scales and large environmental densities.

(3) The Gaussian approach predicts that the only local property of the environment on which the conditional mass function depends is the density contrast $\delta_{\rm e}$. Thus a prediction of the mass function in the different geometrical environments should be possible if we know the overall dependence of the mass function on overdensity, and if we can predict the overdensity distributions for the different environments.

(4) A detailed test of this prediction does not succeed very well, since the overdensity distributions are not well predicted by the Gaussian theory. This is improved in part by considering a lognormal model for the evolved density field, but discrepancies with numerical data remain.

(5) Nevertheless, we have been able to test directly the fundamental prediction of this work, which is that different geometrical environments with the same overdensity should have identical halo mass functions. This is verified in the MDR1 simulation to a good

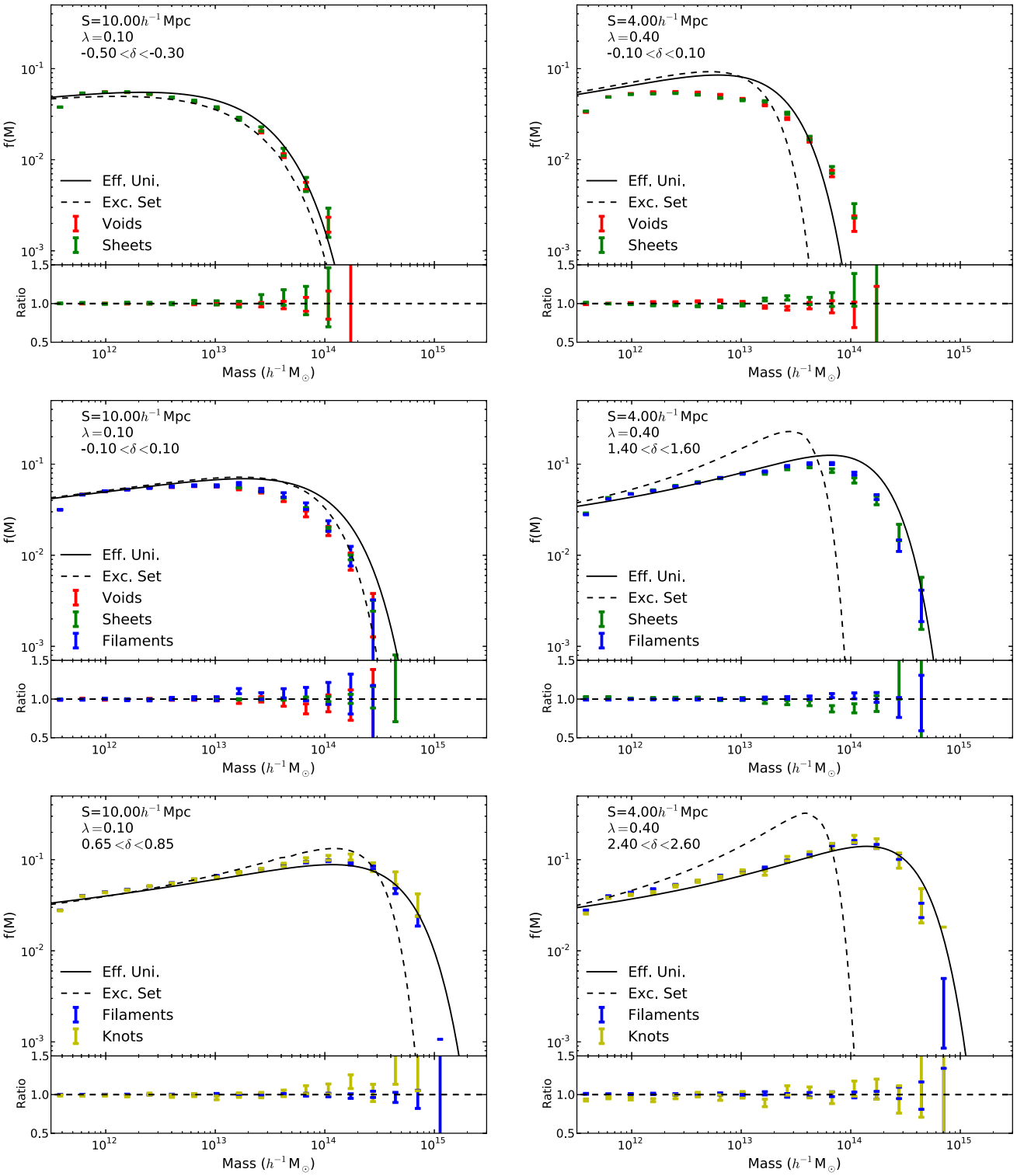

Figure 6. Multiplicity function for the four different environments with their local densities restricted to a given range. In each plot, the lower panel shows the ratio between the different multiplicity functions and their mean value. Our theoretical treatment predicts the same function for all environments in this case, which is realized to a good approximation in all cases. Neither the excursion set prediction nor the effective-universe approach agree quantitatively with the simulated data in all cases, but an overall better agreement is obtained for the effective-universe formalism, especially for smaller scales and larger environmental overdensities.

Table 1. Compatibility of the mass functions for different environments with restricted environmental densities for the two combinations of (S, $\lambda_{\rm th}$) explored in Fig. 6: ($10\,h^{-1}$ Mpc, 0.1), left-hand column, and ($4\,h^{-1}$ Mpc, 0.4), right-hand column.

$S = 10\,h^{-1}$ Mpc, $\lambda_{\rm th} = 0.1$		$S = 4\,h^{-1}$ Mpc, $\lambda_{\rm th} = 0.4$	
δ	$\Delta f(M)$ (per cent)	δ	$\Delta f(M)$ (per cent)
$(-0.5, -0.3)$	1.1	$(-0.1, 0.1)$	4.6
$(-0.1, 0.1)$	13.2	$(1.4, 1.6)$	3.5
$(0.65, 0.85)$	5.2	$(2.4, 2.6)$	8.6

approximation for a wide range of masses, filters and eigenvalue thresholds. We find a maximum relative deviation of about 13 per cent between mass functions in different environments, which are, nevertheless, fully compatible in a statistical sense. In this regard, we see no evidence for any effect of tidal forces on halo abundances in addition to the impact of local overdensity. This could be consistent with the claim by Yan et al. (2013) that galaxy properties in the SDSS lacked an explicit dependence on environmental ellipticity, as well as with the results found by Metuki et al. (2015) in N-body simulations; it will be interesting to repeat such an analysis using the explicit decomposition by geometrical environment that we have studied here.

(6) This result suggests that scalar halo properties are not heavily influenced by the tidal field beyond the local overdensity. This is not at odds with the results found by Forero-Romero, Contreras & Padilla (2014), Libeskind et al. (2014), who find that directional quantities, such as orientations or angular momenta, show a strong correlation with the directions defined by the tidal tensor.

ACKNOWLEDGEMENTS

We would like to thank Catherine Heymans and Juan García-Bellido for useful comments and discussions, as well as the journal editor and referee for their thorough report and very helpful comments. DA is supported by ERC grant 259505 and acknowledges the hospitality of the Royal Observatory, Edinburgh, during a research stay supported by a JAE-Predoc scholarship. EE is supported by a STFC postgraduate studentship. The MultiDark Database used in this paper and the web application providing online access to it were constructed as part of the activities of the German Astrophysical Virtual Observatory as result of a collaboration between the Leibniz-Institute for Astrophysics Potsdam (AIP) and the Spanish MultiDark Consolider Project CSD2009-00064. The Bolshoi and MultiDark simulations were run on the NASA's Pleiades supercomputer at the NASA Ames Research Center. The MultiDark-Planck (MDPL) and the BigMD simulation suite have been performed in the Supermuc supercomputer at LRZ using time granted by PRACE.

REFERENCES

Abell G. O., 1965, ARA&A, 3, 1
Alpaslan M. et al., 2014, MNRAS, 440, L106
Aragon-Calvo M. A., 2012, preprint (arXiv:1210.7871)
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
Berlind A. A. et al., 2003, ApJ, 593, 1
Bernardeau F., 1994, ApJ, 427, 51
Bond J. R., Myers S. T., 1996, ApJS, 103, 1
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Cautun M., van de Weygaert R., Jones B. J. T., Frenk C. S., 2014, MNRAS, 441, 2923
Cole S., Helly J., Frenk C. S., Parkinson H., 2008, MNRAS, 383, 546
Coles P., Jones B., 1991, MNRAS, 248, 1
Doroshkevich A. G., 1970, Astrophysics, 6, 320
Eardley E. et al., 2014, preprint (arXiv:1412.2141)
Falck B., Koyama K., Zhao G.-b., Li B., 2014, J. Cosmol. Astropart. Phys., 07, 58
Forero-Romero J. E., Hoffman Y., Gottlöber S., Klypin A., Yepes G., 2009, MNRAS, 396, 1815
Forero-Romero J. E., Contreras S., Padilla N., 2014, MNRAS, 443, 1090
Gao L., White S. D. M., 2007, MNRAS, 377, L5
Goldberg D. M., Vogeley M. S., 2004, ApJ, 605, 1
Gottlöber S., Łokas E. L., Klypin A., Hoffman Y., 2003, MNRAS, 344, 715
Hahn O., Porciani C., Carollo C. M., Dekel A., 2007, MNRAS, 375, 489
Hoffman Y., Metuki O., Yepes G., Gottlöber S., Forero-Romero J. E., Libeskind N. I., Knebe A., 2012, MNRAS, 425, 2049
Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372
Kitaura F.-S., Jasche J., Metcalf R. B., 2010, MNRAS, 403, 589
Lacey C., Cole S., 1993, MNRAS, 262, 627
Libeskind N. I., Hoffman Y., Forero-Romero J., Gottlöber S., Knebe A., Steinmetz M., Klypin A., 2013, MNRAS, 428, 2489
Libeskind N. I., Knebe A., Hoffman Y., Gottloeber S., 2014, MNRAS, 443, 1274
Maggiore M., Riotto A., 2010, ApJ, 711, 907
Martino M. C., Sheth R. K., 2009, MNRAS, 394, 2109
Metuki O., Libeskind N. I., Hoffman Y., Crain R. A., Theuns T., 2015, MNRAS, 446, 1458
Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
Muldrew S. I. et al., 2012, MNRAS, 419, 2670
Musso M., Paranjape A., Sheth R. K., 2012, MNRAS, 427, 3145
Nuza S. E., Kitaura F.-S., Hess S., Libeskind N. I., Mueller V., 2014, MNRAS, 445, 988
Paranjape A., Sheth R. K., Desjacques V., 2013, MNRAS, 431, 1503
Peacock J. A., 2007, MNRAS, 379, 1067
Peacock J. A., Heavens A. F., 1990, MNRAS, 243, 133
Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144
Peebles J. P. E., 1993, Principles of Physical Cosmology. Princeton Univ. Press, Princeton, NJ
Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018
Press W. H., Schechter P., 1974, ApJ, 187, 425 (PS)
Rossi G., 2013, MNRAS, 430, 1486
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Sousbie T., Colombi S., Pichon C., 2009, MNRAS, 393, 457
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, MNRAS, 433, 1230
Weinberg S., 1972, Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity. Wiley, New York
Yan H., Fan Z., White S. D. M., 2013, MNRAS, 430, 3432

APPENDIX A: THE EFFECTIVE-UNIVERSE APPROACH

Consider a spherical perturbation in an otherwise homogeneous universe. It is a well-known result in gravitational theory that at any distance from the centre of the perturbation, it must evolve as a parallel FRW cosmology with some effective cosmological parameters which can be entirely determined in terms of the amplitude of the density perturbation. This result allows us to interpret the

conditional mass function for an environment with overdensity δ_e as the mass function in the corresponding effective universe.

The 'environmental' cosmological parameters are related to the background ones and the perturbation's overdensity through

$$\Omega_m^e = \Omega_m^{BG}(1+\delta)/\eta^2 \quad \Omega_\Lambda^e = \Omega_\Lambda^{BG}/\eta^2 \quad H_0^e = \eta^2 H_0^{BG}, \tag{A1}$$

where the superscripts BG and e denote quantities in the background and in the effective universe, respectively. The ratio between the current expansion rates inside and outside the perturbation, η, can be fixed by imposing that the age of the Universe

$$t_{BB} = \frac{1}{H_0}\int_0^1 \frac{dx}{x\sqrt{\Omega_m x^{-3} + \Omega_\Lambda + \Omega_k x^{-2}}} \tag{A2}$$

must be the same as measured by any observer. This effectively implies that the perturbation must be a purely growing mode that disappears at early times.

Once the effective cosmological parameters are known, the scaling factor D_g in equation (10) is given by the ratio of the growth factors in the two cosmologies. Normalizing this ratio to be 1 at early times (where the perturbation gradually disappears), this quantity is given by

$$D_g = \left(\frac{\Omega_m^{BG} h_{BG}^2}{\Omega_m^e h_e^2}\right)^{1/3} \frac{\Omega_m^e}{\Omega_m^{BG}} \frac{g\left(\Omega_m^e, \Omega_\Lambda^e\right)}{g\left(\Omega_m^{BG}, \Omega_\Lambda^{BG}\right)}, \tag{A3}$$

where

$$g(\Omega_m, \Omega_\Lambda) = \int_0^1 dx \left(\frac{x}{\Omega_m + \Omega_\Lambda x^3 + \Omega_k x}\right)^{3/2}. \tag{A4}$$

Notice that at this point we have not taken into account the size of the environment. For large smoothing scales or comparatively small halo masses, this is not an important concern: we may treat the environment as an infinite effective universe in which haloes of any mass may form. In practice however, the mass of the largest haloes ($M \sim 10^{15}\,M_\odot \rightarrow R_h \sim 15$ Mpc) corresponds to scales of the order of the filter scale used to define the environment, and hence halo masses must be restricted by the amount of matter that is available in their environment. We have taken this effect into account by restricting the Fourier modes that can contribute to the variance of the overdensity field in a given environment, suppressing those corresponding to scales larger than R_e. In practice we have implemented this by weighting each mode by the 'inverse' of the window function used to define the environment, W_e:

$$\sigma_{eff}(M) = \int_0^\infty \frac{k^2 dk}{2\pi^2}\,[1 - W_e(kR_e)]^2\,|W(kR_h)|^2 P_k. \tag{A5}$$

APPENDIX B: CORRELATIONS BETWEEN ENVIRONMENT AND DENSITY FOR GAUSSIAN FIELDS

B1 The eigenvalue distribution

Consider a Gaussian potential field $\tilde{\phi}$ smoothed over a length-scale R. Since the tidal tensor $\hat{T}$ is a symmetric matrix only six of its components are independent. We will label them with a single index: $T_A = (T_{11}, T_{22}, T_{33}, T_{23}, T_{31}, T_{12})$. It is straightforward to calculate the covariance matrix of the T_As:

$$\langle T_A T_B\rangle = \frac{\sigma_R^2}{15}\begin{pmatrix} 3 & 1 & 1 & 0 & 0 & 0\\ 1 & 3 & 1 & 0 & 0 & 0\\ 1 & 1 & 3 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \tag{B1}$$

where σ_R^2 is given in equation (6) with $R_a = R_b = R$. This matrix can be diagonalized by changing to the variables

$$\begin{aligned} &\tau_1 \equiv \nu \equiv \frac{1}{\sigma_R}(T_1 + T_2 + T_3), \quad \tau_2 \equiv \rho \equiv \frac{1}{2\sigma_R}(T_1 - T_3),\\ &\tau_3 \equiv \theta \equiv \frac{1}{2\sigma_R}(T_1 - 2T_2 + T_3),\\ &\tau_4 \equiv \frac{T_4}{\sigma_R}, \quad \tau_5 \equiv \frac{T_5}{\sigma_R}, \quad \tau_6 \equiv \frac{T_6}{\sigma_R}. \end{aligned} \tag{B2}$$

Note that by definition, ν is proportional to the local density contrast, $\delta = \nu\sigma_R$, and that ρ and θ are trivially related to the ellipticity $e \equiv \rho/\nu$ and prolateness $p \equiv \theta/\nu$. In terms of these new variables the covariance matrix is diagonal:

$$\mu_{AB} \equiv \langle \tau_A \tau_B\rangle = \mathrm{diag}\left(1, \frac{1}{15}, \frac{1}{5}, \frac{1}{15}, \frac{1}{15}, \frac{1}{15}\right) \tag{B3}$$

and the joint distribution of the T_As is

$$\begin{aligned} &P(\{T_A\})\prod_A dT_A = \frac{e^{-Q/2}}{\sqrt{(2\pi)^6 \det(\hat{\mu})}}\prod_A d\tau_A,\\ &Q \equiv \nu^2 + 15\rho^2 + 5\theta^2 + 15\left(\tau_4^2 + \tau_5^2 + \tau_6^2\right). \end{aligned} \tag{B4}$$

This holds in any coordinate system, but it will be most useful in the one in which $\hat{T}$ is diagonal (i.e. $\hat{T} = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3, 0, 0, 0)$). As proved by Bardeen et al. (1986) the volume element of the space of 3×3 symmetric matrices can be written in terms of the matrix eigenvalues and the Euler angles of the rotation necessary to diagonalize it:

$$\prod_A dT_A = |(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_3)(\lambda_1 - \lambda_3)|\,d\lambda_1\,d\lambda_2\,d\lambda_3\frac{d\Omega_3}{6}, \tag{B5}$$

where $d\Omega_3$ is the volume element of $\mathcal{S}^3$ (the total volume of which is $2\pi^2$). Up to now we have not chosen any specific ordering for the eigenvalues. There are six possible orderings, and the probability density is symmetric with respect to these, therefore imposing a specific ordering would introduce a factor 6 in the probability density above. Choosing $\lambda_1 \geq \lambda_2 \geq \lambda_3$ (which implies $\theta \in [-\rho, \rho]$, $\rho \in [0, \infty)$) and integrating out the irrelevant angular part we obtain the probability distribution given in equation (15).

B2 Correlation with the local density contrast

Now let us consider the distribution of the density contrast smoothed over a scale R_h in different regions classified via the tidal tensor eigenvalues smoothed over a scale R_e. Here we will use the notation introduced in Section 2, with $\delta_e \equiv \mathrm{Tr}(\hat{T})$.

The covariance of δ_h with the T_As is easy to calculate:

$$\langle \delta_h\, T_A\rangle = \frac{\sigma_{eh}^2}{3}(\delta_{A1} + \delta_{A2} + \delta_{A3}). \tag{B6}$$

Defining $G_a = (\delta_{\rm h}, T_A)$, $(a = 0, \ldots, 6)$ the full covariance matrix is

$$\langle G_a G_b \rangle = \frac{1}{15} \times \begin{pmatrix} 15\sigma_{\rm hh}^2 & 5\sigma_{\rm eh}^2 & 5\sigma_{\rm eh}^2 & 5\sigma_{\rm eh}^2 & 0 & 0 & 0 \\ 5\sigma_{\rm eh}^2 & 3\sigma_{\rm ee}^2 & \sigma_{\rm ee}^2 & \sigma_{\rm ee}^2 & 0 & 0 & 0 \\ 5\sigma_{\rm eh}^2 & \sigma_{\rm ee}^2 & 3\sigma_{\rm ee}^2 & \sigma_{\rm ee}^2 & 0 & 0 & 0 \\ 5\sigma_{\rm eh}^2 & \sigma_{\rm ee}^2 & \sigma_{\rm ee}^2 & 3\sigma_{\rm ee}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\rm ee}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{\rm ee}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{\rm ee}^2 \end{pmatrix}. \quad \text{(B7)}$$

As before, in order to simplify this matrix we define

$$\gamma_0 \equiv \nu_{\rm h} \equiv \frac{\delta_{\rm h}}{\sigma_{\rm hh}}, \quad \gamma_1 \equiv \nu_{\rm e} \equiv \frac{T_1 + T_2 + T_3}{\sigma_{\rm ee}},$$
$$\gamma_2 \equiv \rho \equiv \frac{T_1 - T_3}{2\sigma_{\rm ee}}, \quad \gamma_3 \equiv \theta \equiv \frac{T_1 - 2T_2 + T_3}{2\sigma_{\rm ee}},$$
$$\gamma_4 \equiv \frac{T_4}{\sigma_{\rm ee}}, \quad \gamma_5 \equiv \frac{T_5}{\sigma_{\rm ee}}, \quad \gamma_6 \equiv \frac{T_6}{\sigma_{\rm ee}}, \quad \text{(B8)}$$

the covariance of which is

$$\langle \gamma_a \gamma_b \rangle = \begin{pmatrix} 1 & \varepsilon & 0 & 0 & 0 & 0 & 0 \\ \varepsilon & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/15 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/15 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/15 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/15 \end{pmatrix}. \quad \text{(B9)}$$

Now we can follow the same procedure as before: change the volume element of the T_As to the one in the space of eigenvalues and rotations, change to the coordinate system in which $\hat{T}$ is diagonal, transform everything to our variables $\{\gamma_a\}$, choose a specific ordering for the eigenvalues and integrate out the irrelevant angular part. At the end of the day we obtain the distribution given in equation (23).

This paper has been typeset from a TEX/LATEX file prepared by the author.