The nature of voids – II. Tracing underdensities with biased galaxies

S. Nadathur[1]★ and S. Hotchkiss[2]★

[1]*Department of Physics, University of Helsinki and Helsinki Institute of Physics, PO Box 64, FIN-00014 Helsinki, Finland*
[2]*Department of Physics and Astronomy, University of Sussex, Falmer, Brighton BN1 9QH, UK*

ABSTRACT

We study how the properties of cosmic voids depend on those of the tracer galaxy populations in which they are identified. We use a suite of halo occupation distribution mocks in a simulation, identify voids in these populations using the ZOBOV void finder and measure their abundances, sizes, tracer densities and dark matter content. To separate the effects of bias from those of sampling density, we do the same for voids traced by randomly downsampled subsets of the simulation dark matter particles. At the same sampling density, galaxy bias reduces the total number of voids by ∼50 per cent and can dramatically change their size distribution. The matter content of voids in biased and unbiased tracers also differs. Deducing void properties from simulation therefore requires the use of realistic galaxy mocks. We discuss how the void observables can be related to their matter content. In particular we consider the compensation of the total mass deficit in voids and find that the distinction between over- and undercompensated voids is not a function of void size alone, as has previously been suggested. However, we find a simple linear relationship between the average density of tracers in the void and the total mass compensation on much larger scales. The existence of this linear relationship holds independent of the bias and sampling density of the tracers. This provides a universal tool to classify void environments and will be important for the use of voids in observational cosmology.

Key words: methods: data analysis – methods: numerical – cosmology: observations – large-scale structure of Universe.

1 INTRODUCTION

Several recent studies have advocated the use of large underdensities in the matter distribution, known as cosmic voids, which can be important tools for cosmology in a number of different contexts. These include using voids to test modified gravity scenarios (Li, Zhao & Koyama 2012; Clampitt, Cai & Li 2013; Cai, Padilla & Li 2015; Zivick et al. 2015), to probe dark energy (Ryden 1995; Lee & Park 2009; Bos et al. 2012; Pisani et al. 2015; Sutter et al. 2015a) or to detect the effects of warm dark matter (Tikhonov et al. 2009; Yang et al. 2015). These tests make use of observables such as the abundances, size distributions, shapes and radial density profiles of voids. This information will be available for very large numbers of voids in future large-scale survey data such as *Euclid* (Laureijs et al. 2011).

Weak lensing by voids (Krause et al. 2013; Melchior et al. 2014) and their possible gravitational effects on the cosmic microwave background (CMB) via the integrated Sachs–Wolfe (ISW)

effect have also been studied (Granett, Neyrinck & Szapudi 2008; Cai et al. 2014; Hotchkiss et al. 2015; Planck Collaboration XXI 2015). In principle for rare voids such studies probe the extreme tails of the distribution of fluctuations of the gravitational potential so may provide unique insights. Indeed the case of possible ISW imprints of voids is particularly intriguing, since theoretical considerations and simulation results show that any such detection would be in strong tension with the standard Λ cold dark matter (ΛCDM) cosmology (Nadathur, Hotchkiss & Sarkar 2012; Flender, Hotchkiss & Nadathur 2013; Cai et al. 2014; Aiola, Kosowsky & Wang 2015; Hotchkiss et al. 2015).

Several catalogues of voids found in galaxy redshift surveys have already been published (Pan et al. 2012; Sutter et al. 2012; Nadathur & Hotchkiss 2014), and more will be available from future data releases. These catalogues can be used for the various studies described above. However, since the dark matter (DM) distribution is not directly observable, the comparison of observations with theoretical expectation necessary for precision cosmology requires a detailed understanding of how biased galaxies actually trace matter underdensities.

A common theoretical model of voids is that based on the excursion set formalism (Sheth & van de Weygaert 2004), which

★ E-mail: seshadri.nadathur@helsinki.fi (SN); s.a.hotchkiss@sussex.ac.uk (SH)

can be modified to account for the bias of galaxy tracers (e.g. Furlanetto & Piran 2006). However, this model assumes that voids can be described as spherical underdensities surrounded by higher density walls that have reached shell-crossing. Recent results (Falck & Neyrinck 2015; Nadathur & Hotchkiss 2015) show that this assumption does not hold for the most commonly used watershed void finders, and more generally may not hold for any practical void-finding technique. Instead voids in simulation and observation correspond to a much more general set of density minima of varying depths.

This means that the development of cosmological predictions for void observables must be guided by and calibrated on simulation results. To properly account for the effects of bias in the galaxy distribution, these simulations should be populated with realistic mock galaxy populations, for example by using a halo occupation distribution (HOD; e.g. Berlind & Weinberg 2002) model. However, a number of recent studies of voids (e.g. Chan, Hamaus & Desjacques 2014; Zivick et al. 2015; Pisani et al. 2015; Sutter et al. 2015a) do not follow this approach. Instead these authors randomly downsample the DM particle output of simulations to the same mean number density as that of galaxies in various surveys, and use these particles as tracers of voids. This takes into account effects relating to the sampling density of tracers. However, the clustering properties of subsampled DM and galaxies are in general expected to differ, which should also affect voids. The justification for such a procedure has therefore relied on the results of Sutter et al. (2014), who suggest that tracer bias has a smaller effect on void properties than sampling density and can therefore be ignored.

This result is however hard to reconcile with the expected dependence of the halo mass function on the background matter density on larger scales (Sheth & Tormen 2002). Gottlöber et al. (2003) find that the numbers of high- and intermediate-mass haloes in voids is sharply suppressed, and Alonso, Eardley & Peacock (2015) also find a strong dependence of the halo mass function on the density environment. This means that properties such as the sizes, abundances and density profiles of voids traced by galaxies in such haloes should be expected to be very different to those traced by unbiased DM. The picture is further complicated by the fact that the halo abundance within voids is also expected to be cosmology dependent (Lee 2012; Neyrinck et al. 2014).

We therefore revisit this question in this paper. We make use of a suite of different HOD models to generate biased galaxy tracer populations in an *N*-body simulation, and also use sets of subsampled DM particles with the same number densities as controls. We use these sets of tracers to identify voids using the ZOBOV watershed void-finding algorithm, and compare their properties. We show that whereas the mean sampling density determines the minimum resolvable void size, at the same sampling density differences in the tracer bias produce large differences in the void size distribution, especially at large void sizes, as well as affecting the total number of detected voids and the densities within voids. This underscores the importance of using appropriate HOD models in simulations when obtaining predictions for void observables.

We then investigate the distribution of DM within galaxy voids. The minimum tracer density within each void, n_{min}, varies significantly over the population and we find that it can be used to predict the minimum DM density within the void. The nature of this relationship however depends strongly both on the tracer bias and the sampling density. The stacked average density profile of voids is also of interest, as it could potentially be a sensitive test of cosmology (Cai et al. 2015; Barreira et al. 2015). Some studies (Hamaus, Sutter & Wandelt 2014; Nadathur et al. 2015) have also suggested

the possibility of a universal density profile, focusing on the variation of the average profile with the mean void size. We show that the average profile also depends strongly on n_{min}.

An important related characteristic of voids is whether the total mass deficit within them is compensated or not. This is a key element in determining the environment within voids (Hamaus et al. 2014). Voids in which the central density deficit is not compensated out to large distances will correspond to gravitational potential hills, while overcompensated voids are smaller local density minima in larger scale overdense regions, and will therefore correspond to regions of negative gravitational potential. Distinguishing between these two cases is therefore crucial for understanding void environments, and in particular for the use of voids in the cosmological tests described above.

It has previously been suggested (Ceccarelli et al. 2013; Cai et al. 2014; Hamaus et al. 2014) that such a distinction can be made on the basis of the void size. Instead we find that the transition from under- to overcompensation is dependent on *both* the void size and minimum tracer density within the void, and is not well predicted by either property alone. However, we are able to identify a very simple linear relationship between the enclosed mass density contrast Δ at scales much larger than the void and the integrated average tracer number density within the void, n_{avg}. The existence of this linear relationship is found to be remarkably independent of the bias or sampling density $\bar{n}$ of the tracer population, with $n_{avg} = \bar{n}$ marking the transition between $\Delta < 0$ (undercompensated) and $\Delta > 0$ (overcompensated) for all tracer types. In contrast, the void size does not provide such a universal diagnostic of void compensation.

In Section 2 we describe the properties of the simulation, the construction of HOD galaxy samples and the void-finding algorithm. In Section 3 we investigate the properties of these voids and lay out our results. We discuss their significance and conclude in Section 4. Some details of the stacked density profiles around voids and comparisons with previous results are presented in Appendix A.

2 METHODS

2.1 Simulation

We make use of *N*-body simulation data from the MultiDark simulation project (Prada et al. 2012), in particular the MDR1 data release.[1] MDR1 simulates a ΛCDM cosmological model with parameters $(\Omega_m, \Omega_\Lambda, \Omega_b, h, n_s, \sigma_8) = (0.27, 0.73, 0.0469, 0.7, 0.95, 0.82)$ using 2048^3 DM particles in a cubic box of side $1\,h^{-1}$ Gpc. The mass of each DM particle in the simulation is $8.7 \times 10^9\,h^{-1}\,M_\odot$, and it has a force resolution of $7\,h^{-1}$ kpc. The initial conditions for the density evolution are set at redshift $z = 65$ using the Zeldovich approximation.

Haloes are identified in the simulation using the Bound Density Maximum (BDM) algorithm first described by Klypin & Holtzman (1997), and as updated by Riebe et al. (2013). This algorithm detects as haloes spherical regions around density maxima exceeding the virialization density threshold in a top-hat model of the growth of density fluctuations and removes gravitationally unbound particles from the halo. The virial radius of the halo, R_{vir}, is defined as the radius within which the density contrast $\Delta_{vir}(z)$ defined with respect to the background density $\bar{\rho}(z)$ exceeds the collapse threshold given by the approximation of Bryan & Norman (1998) – at the redshift

---

[1] Publicly available at www.cosmosim.org

$z = 0$ snapshot we use, this corresponds to $\Delta_{\rm vir} = 360$. The minimum resolved halo mass in MDR1 is $1.7 \times 10^{11}\, h^{-1}\,{\rm M}_{\odot}$.

The creation of HOD mocks and other tracer samples for void identification is explained below; however, to measure the true density at void locations we make use of the underlying DM density field from the full resolution simulation output at redshift 0. This density field is determined on a 1024^3 grid using a cloud-in-cell interpolation, and is then smoothed using a Gaussian kernel of width equal to one grid cell. The resulting density field has a sub-Mpc resolution, which is sufficient given the void sizes in our catalogues.

2.2 HOD mocks and tracer samples

We populate the halo catalogue with mock galaxies using the HOD model of Zheng, Coil & Zehavi (2007), which assigns galaxies to a DM halo according to a distribution based on the halo mass M. Central and satellite galaxies are treated separately. According to this model, the mean occupation function of central galaxies is parametrized as

$$\langle N_{\rm cen}(M) \rangle = \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log M - \log M_{\rm min}}{\sigma_{\log M}}\right)\right], \qquad (1)$$

and the number of central galaxies in each mass bin follows a nearest integer distribution. On the other hand, the number of satellite galaxies follows a Poisson distribution with

$$\langle N_{\rm sat}(M) \rangle = \langle N_{\rm cen}(M) \rangle \left(\frac{M - M_0}{M_1'}\right)^{\alpha}. \qquad (2)$$

Central galaxies are placed at the centre of their respective haloes, while satellite galaxies are distributed through the halo with radial distances from the centre drawn from a random distribution based on a fiducial Navarro–Frenk–White (NFW) mass profile (Navarro, Frenk & White 1996, 1997).

The HOD distribution is thus characterized by five parameters: $M_{\rm min}$, M_0, M_1', $\sigma_{\log M}$ and α. We choose these parameters to match those determined in the literature from fits to the number density and clustering properties of observed galaxy samples.

Two of our mock samples, designated Main1 and Main2, are chosen to match volume-limited luminosity threshold samples from the Sloan Digital Sky Survey (SDSS) Data Release 7 (DR7) main galaxy sample, with r-band absolute magnitude $M_r^{\rm max}$ of -20.5 and -21, respectively (Zehavi et al. 2011), with median redshift $z \sim 0.1$. As these authors use a definition of the halo overdensity $\Delta_{\rm c} = 200$ with respect to the critical density which differs from the MultiDark definition above, we rescale the halo masses and virial radii from the simulation results using the concentration values provided and assuming an NFW profile. The large-scale linear bias b, defined by $\delta_{\rm g} = b\delta$, where $\delta_{\rm g}$ is the galaxy density contrast and δ that of matter is 1.29 for Main1 and 1.40 for Main2.

Our third sample is chosen to match the properties of the low redshift (LOWZ) sample of galaxies from the SDSS Baryon Oscillation Spectroscopic Survey (BOSS) Data Releases 10 and 11, with $0.15 \lesssim z \lesssim 0.43$, and is based on the HOD parameters obtained by Manera et al. (2015). Although the LOWZ sample in reality has a redshift-dependent mean sample density, we choose the parameter values for $\overline{n} = 2.98 \times 10^{-4}\,h^3\,{\rm Mpc}^{-3}$ and uniformly subsample our mocks down to achieve this mean density. Manera et al. (2015) use haloes within a simulation generated using the PTHalos method within second-order Lagrangian perturbation theory (2LPT) rather than full N-body simulations; their minimum halo mass is also too large to be able to resolve all populated haloes, so they assign approximately 7 per cent of galaxies to DM particle

Table 1. Properties of the mock HOD samples used in this work. Bias values are approximate and intended as a guide only.

Sample name	$\log M_{\rm min}$	$\sigma_{\log M}$	$\log M_0$	$\log M_1'$	α	$\overline{n}$ ($h^3\,{\rm Mpc}^{-3}$)	b
Main1	12.14	0.17	11.62	13.43	1.15	3.18×10^{-3}	1.3
Main2	12.78	0.68	12.71	13.76	1.15	1.16×10^{-3}	1.4
LOWZ	13.24	0.62	14.32	13.20	0.93	2.98×10^{-4}	2.0

positions rather than haloes. These differences mean that our mock sample is unlikely to precisely match the clustering properties of the observed LOWZ galaxies; nevertheless it provides a sufficiently realistic realization of a distribution of galaxies with bias $b \sim 2$.

Table 1 summarizes the parameter values and resulting average number density $\overline{n}$ and large-scale bias b of our mocks.

To complement these mock samples and to serve as controls, we also use three sets of randomly subsampled DM particles from the simulation output as tracers with which to identify voids. These samples are referred to as DM Main1, DM Main2 and DM LOWZ and have the same average tracer densities as the Main1, Main2 and LOWZ HOD samples, respectively. However, as the random subsampling does not change the fundamental clustering properties of the DM field, these three samples are all unbiased ($b = 1$). Finally, we also use three Poisson point samples, consisting of uniformly distributed random points with the same mean density as those of the mock galaxy tracers.

2.3 Void finding and merging

We make use of the ZOBOV watershed void-finding algorithm (Neyrinck 2008) to identify voids in each of the tracer samples described above. ZOBOV uses a Voronoi tessellation field estimator (VTFE) technique to reconstruct the tracer density field from a discrete distribution of particles. It then identifies local minima in this field and the watershed basins around them. These basins form a non-overlapping set of density depressions or voids.

At this stage, ZOBOV can also merge neighbouring voids together according to the watershed principle to form a final hierarchy of voids and subvoids. Depending on the choice of input parameters to control void merging, the resulting final void hierarchy and distribution of void sizes can vary widely: at one extreme one obtains percolation, with the entire simulation box or Universe being included in the largest void; at the other extreme there is no merging at all and the original voids are retained.

Unfortunately, the choice of how to control void merging is largely subjective and different options have been applied in the literature. Neyrinck (2008) suggests using the ratio of the lowest (VTFE reconstructed) tracer number density along the void wall to the minimum VTFE density at its centre, relative to the same ratio for spurious voids in Poisson noise, to determine the 'most probable extent'. An alternative is to stop the growth of a void through merging if the minimum VTFE density along the watershed ridge separating it from a shallower potential subvoid, $n_{\rm link}$, is greater than some pre-defined threshold. The VIDE void-finding toolkit (Sutter et al. 2015b), which is based on ZOBOV, sets this threshold at 0.2 times the mean tracer density $\overline{n}$. However, this value of 0.2 lacks any theoretical justification even when the tracers are DM particles in a simulation, because the recovered voids do not in any case correspond to the model objects described by Sheth & van de Weygaert (2004) from which the canonical value for the

Table 2. Strategies for controlling merging in the void hierarchy.

Label	Criteria for merging	
	Link density	Density ratio
VIDE	$n_{\text{link}} < 0.2\overline{n}$	Unconstrained
Minimal	$n_{\text{link}} < \overline{n}$	$r < 2$
Isolated	No merging	No merging

enclosed density contrast $\Delta = -0.8$ derives (Nadathur & Hotchkiss 2015).[2]

A further difficulty with the VIDE threshold for merging is that it applies to the minimum tracer *number* density along void walls. Therefore, when applied to differently biased tracer samples, the same numerical value for the threshold corresponds to rather different DM densities, and thus to different physical criteria for merging. Only for the special case where the threshold value is set equal to the mean sample density, i.e. neighbouring voids are merged together if $n_{\text{link}} < \overline{n}$, does this strategy result in the same physical consequences independent of the tracer bias.

In this paper we will compare three different strategies to control the merging of voids in the formation of the final hierarchy. The first strategy is that applied in the VIDE algorithm, which has been used in several recent works (Chan et al. 2014; Hamaus et al. 2014; Zivick et al. 2015; Pisani et al. 2015; Pollina et al. 2015; Sutter et al. 2015a) studying void properties. The second strategy is simply to use the original output of density minima from ZOBOV without any merging at all, which provides a set of non-overlapping voids defined purely by the topology of the tracer density field. We shall refer to the voids thus obtained as 'Isolated' voids.

The third strategy we consider allows merging of voids if the minimum tracer density on the separating ridge is below the mean. However, as this may still allow genuinely distinct structures to be classified as a single void, we implement an additional condition on the density ratio described above, i.e. we prevent the merging of two voids if the minimum density on the watershed ridge is more than twice the minimum density of the shallower void. This choice is based on the Poisson noise characteristics discussed by Neyrinck (2008): it is designed to count neighbouring voids as distinct if the ridge separating them could not be an artefact of the discrete sampling of the density field. We refer to this as the 'Minimal' choice of void merging criteria; it has also previously been used by Nadathur & Hotchkiss (2014), Hotchkiss et al. (2015) and Nadathur et al. (2015). The three strategies for void merging are summarized in Table 2.

Most qualitative conclusions regarding the properties of voids in the simulation are independent of the particulars of the merging strategy chosen; some exceptions to this statement are highlighted below. However, quantitative results will inevitably depend on this choice. Since the 'Isolated' option of preventing merging altogether provides a void definition that depends only on the topology of the tracer density field without further arbitrary parameters, it may be the one most easily described by a future theoretical model. Therefore, unless otherwise stated, the results presented in this paper relate to this case.

---

[2] Even if such a correspondence existed between the theoretical model and practical void-finding algorithm, there is no clear relationship between n_{link} and Δ – note that a restriction on $n_{\text{link}} < 0.2\overline{n}$ does *not* restrict the average density of the void to be less than 0.2 – so this choice of threshold value would still be unmotivated.

2.4 Defining void properties

Based on the operational choices described above, we obtain a catalogue of voids within each of our tracer samples. Each void contains a number of tracer particles, and the total void volume is the sum of the volumes of the Voronoi cells of each of its member particles in the tessellation. We define the void effective radius, $R_{\text{v}} = (3V/4\pi)^{1/3}$, to be the radius of a sphere with the same volume as the void, although note that individual voids may be highly aspherical. The void member particle associated with the minimum VTFE-reconstructed tracer density (i.e. the largest Voronoi cell) is labelled the core particle, and this density is denoted by n_{min}. We previously showed (Nadathur & Hotchkiss 2015) that for subsampled DM tracers n_{min} is tightly correlated with (but not the same as) the minimum DM density within the void, and inversely correlated with R_{v}, so that larger voids contain deeper density minima.

The centre of each void is defined as the point of intersection of the Voronoi cells of the void core particle and its three lowest density mutually adjacent neighbours in the tessellation, which is the point within the void maximally distant from any tracer particles. Compared to the alternative volume-weighted barycentre definition often used in the literature, this definition is much better at locating the true DM density minimum within voids (Nadathur & Hotchkiss 2015). Unlike the barycentre definition, it is also independent of the details of void merging and is less affected by shot noise.

Given the total number of member tracer particles of a void and its total volume, it is also simple to define an integrated average tracer density for the void, n_{avg}. This density is obviously always greater than n_{min}; it is also well known (e.g. Nadathur & Hotchkiss 2014; Nadathur et al. 2015) that the operation of the watershed algorithm means that for some voids n_{avg} can be much larger than n_{min} and even $> \overline{n}$ due to the inclusion of higher density filaments, walls and haloes within the watershed basin constituting a void.

The location of the centre together with n_{min}, n_{avg} and R_{v} constitute the core observable properties of a void, which can be defined from the tracer population alone.[3] For the same set of minima in the tracer density field, n_{avg} and R_{v} will differ depending on the details of the void merging criteria described above, but the location of the void centre and n_{min} are independent of this choice.

Using the gridded DM density field from the simulation output we can also measure the true density contrast $\delta = \rho/\overline{\rho} - 1$ within the void and define the total enclosed density contrast within a sphere of radius r, centred at the void centre:

$$\Delta(r) = \frac{3}{r^3} \int_0^r \left[\frac{\rho(y)}{\overline{\rho}} - 1 \right] y^2 \mathrm{d}y. \tag{3}$$

Within the void, at $r < R_{\text{v}}$, values of Δ vary widely for voids of different radius and n_{min} (Nadathur & Hotchkiss 2015). In this paper, a property we consider is the compensation of the void underdensity on scales much larger than the void, for which we will use the value of Δ within a sphere three times the void radius, $\Delta(r = 3R_{\text{v}})$, as a diagnostic.

---

[3] To this list one could also add some measure of the asphericity and direction of alignment of the axes of aspherical voids, but we will not consider these properties in this work. The ratio of tracer densities at the void centre and walls and the full tracer density profile around the void are also in principle measurable, but can be complicated by survey boundary effects in observational data (Nadathur & Hotchkiss 2014; Nadathur et al. 2015).
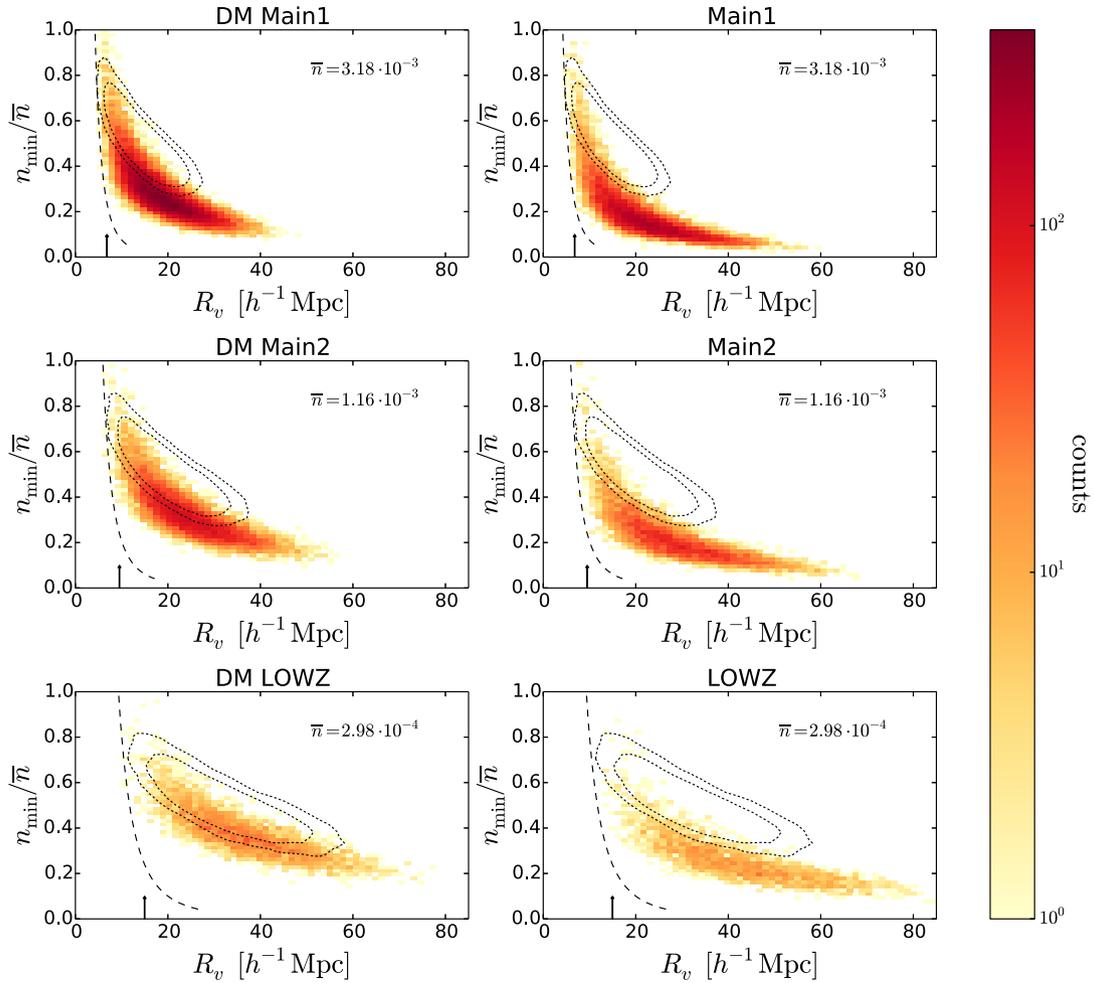
Figure 1. Binned distribution of void minimum tracer densities $n_{\min}$ and radii R_v, for all tracer populations as indicated. Plots in the left-hand column are for unbiased subsampled DM particle tracers, and those in the right-hand column are for the corresponding galaxy tracers with the same average number density $\bar{n}$. Dotted contours in each plot enclose 95 and 99 per cent of all spurious voids found in uniform random distributions of points with the same $\bar{n}$ but no clustering. The dashed line indicates the minimum resolvable void size as a function of density, and the arrow shows the approximate mean interparticle separation $\bar{n}^{-1/3}$. Note that the number of bins remains the same in each plot, so bins cover larger ranges in the right-hand column.

3 PROPERTIES OF VOIDS

3.1 Effect of tracer bias on void abundances and sizes

Fig. 1 shows the distribution of void sizes R_v and core tracer densities $n_{\min}$ for voids each of the tracer samples described in Section 2.2. The right-hand plot in each row shows the distribution for one of the HOD mock samples (Main1, Main2 or LOWZ); on the left is the corresponding distribution for the voids found using the DM particles subsampled down to the same average tracer density $\bar{n}$. Therefore within each row, the sampling density remains the same but the bias increases. In the left-hand column, the bias remains the same (i.e. $b = 1$) but the sampling density decreases. The dotted contours on each plot show the 95 and 99 per cent confidence level contours for the equivalent distributions for voids found by the same algorithm in the random Poisson point samples: these contours therefore remain the same from left to right in a row. The dashed curve in each plot indicates the minimum resolvable void size at the given $\bar{n}$ (Nadathur & Hotchkiss 2015). To clarify the discussion below, for the populations shown in this figure no void mergers have been allowed (corresponding to 'Isolated' voids

described in Section 2.3). However, the qualitative features of the figure do not depend on this choice.

Several effects of tracer bias can be discerned from this figure. As discussed in Nadathur & Hotchkiss (2015), the characteristic 'banana' shape of the distribution is naturally common to voids in all samples. However, at a quantitative level void sizes and core tracer densities depend strongly on the tracer bias as well as the mean tracer density. Although the minimum void size is determined by the resolution limit set by the sampling density $\bar{n}$, voids in mock galaxy samples span a larger range of sizes than those in corresponding subsampled DM populations. Even more importantly, galaxy voids are emptier of tracers than DM voids, as evidenced by the distribution of minimum tracer densities $n_{\min}$ extending to lower values in the right-hand column.

This difference in $n_{\min}$ distributions has important consequences for void merging if the criterion for allowing merging is based – as in the VIDE case discussed in Section 2.3 – on an arbitrary number density threshold. To see this, note that according to the VIDE choice of merging criteria, a necessary but not sufficient condition for two neighbouring voids to be merged together as part of a larger void is that both must have $n_{\min} < 0.2\bar{n}$. As can be seen by comparing
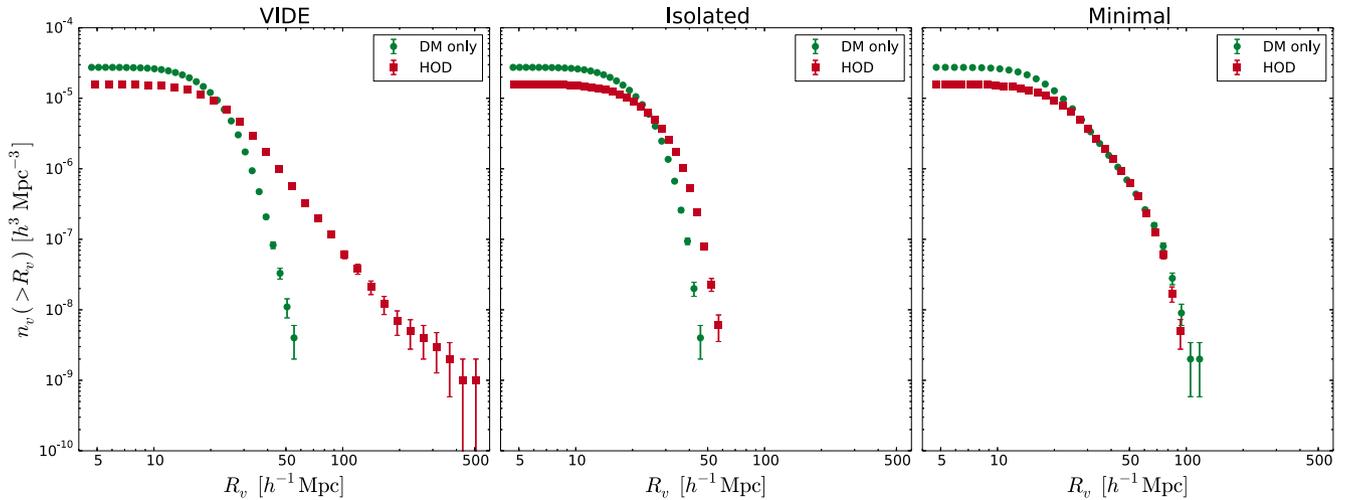
Figure 2. The cumulative number density of voids with radius $>R_v$, for all voids in the unbiased DM Main1 sample, and for those in the Main1 HOD galaxy mocks. The mean tracer density $\bar{n}$ is the same for these two samples, so differences in the void sizes are purely due to galaxy bias. The three panels correspond to three choices of the criteria to handle void merging described in Section 2.3 and Table 2. The bias of the HOD mocks decreases the total void abundance by ~ 50 per cent in all cases, and has a large effect on the shape of the distribution, especially for the merging criteria implemented in the VIDE toolkit.

columns of Fig. 1, a much larger fraction of galaxy voids satisfy this condition than DM voids. This means that, if merging is controlled by such a simple threshold alone, void mergers will be much more common for voids in galaxy distributions than in subsampled DM. Similarly, mergers would be more common between Main1 voids than those in LOWZ. This will necessarily increase the discrepancy between the observable properties of voids in the two populations, as we show later.

It is also noticeable that at each sampling density $\bar{n}$ the distribution of voids in subsampled DM populations shows a greater overlap with that of spurious Poisson voids than that of the corresponding voids in biased galaxy tracers. An intuitive explanation for this is that random subsampling increases shot noise, obscuring the true clustering properties of DM and increasing spurious void detections. For galaxy samples, the decrease in $\bar{n}$ is accompanied by increasing bias and therefore enhanced clustering, so they should be less affected by shot noise. However, we caution that just because a void identified in the simulation has (R_v, n_{min}) values compatible with those found for spurious Poisson voids it does not necessarily follow that this void is itself an artefact of noise! While results for individual voids vary, we find that statistically speaking voids within the overlap region still tend to correspond to true matter underdensities. For this reason we do not apply any of the conservative cuts to the void population previously used in the literature (see Nadathur & Hotchkiss 2015 for a discussion).

Finally, we also find that the overall abundance of voids decreases significantly when the tracer bias is increased at fixed sampling density $\bar{n}$. In fact for each of the sampling densities considered, the total number of voids in unbiased subsampled DM particle populations exceeds that in biased galaxy populations by ~ 50 per cent. This is significantly larger than the variation in void abundance due to changes in cosmological model described by e.g. Zivick et al. (2015), Sutter et al. (2015a) and Pisani et al. (2015). We conclude that predictions for this key observable that are obtained from calibration with simulations that do not include realistic mock galaxy populations are not observationally relevant.

To demonstrate this difference in void abundances more clearly, in Fig. 2 we plot the cumulative number function of voids as a function

of their size for the DM Main1 and Main1 HOD samples, which both have the same sampling density $\bar{n} = 3.18 \times 10^{-3} \, h^3 \, \text{Mpc}^{-3}$. Since the three choices of void merging strategy produce different results, they are shown separately, as indicated by the titles of each panel. The overall deficit of galaxy voids is clearly visible and is due to a lack of small-to-intermediate sized voids, in the range $5 \lesssim R_v \lesssim 20 \, h^{-1} \, \text{Mpc}$.

At large R_v, the number density of galaxy voids is strongly dependent on the void merging criteria. The biggest difference between the DM-only and HOD distributions at large R_v is obtained for the VIDE choice of merging threshold, for the reasons explained above. Note that the VIDE toolkit has been used by a large number of recent void studies which describe the properties of voids using downsampled DM tracers as a substitute for galaxies (e.g. Chan et al. 2014; Sutter et al. 2014, 2015a; Zivick et al. 2015; Pisani et al. 2015). In fact, for the Main1 HOD sample the VIDE criterion leads to percolation through the simulation box, since the largest void has $R_v > 500 \, h^{-1} \, \text{Mpc}$. Such percolation only appears when void merging is controlled using a simple density threshold alone, as is the case for VIDE. It is strongly dependent on the choice of the arbitrary density threshold, as well as the sampling density and bias of the tracers. Such percolation may generally be considered an undesirable feature in choosing a void definition.

Note that since the topology of the tracer density field is bias dependent, even for the case of 'Isolated' voids with no merging the void size distribution differs between HOD mocks and DM tracers at the same sampling density. For the 'Minimal' merging criteria $n_v(>R_v)$ is independent of tracer bias for large R_v but significant differences remain in the distribution at small and intermediate scales, and in the total void abundance.

Fig. 3 shows the differences in the void size distribution over the three mock galaxy samples. Intuitively, increasing tracer bias should shift the distribution to larger void sizes as well as decreasing the total number of voids: more highly biased galaxies live in higher mass haloes and are therefore confined to higher density environments, resulting in more large voids, while the lower sampling density $\bar{n}$ inhibits the ability to resolve small voids. Indeed this is what is seen in the middle and right-hand panels of Fig. 3, for the 'Isolated'
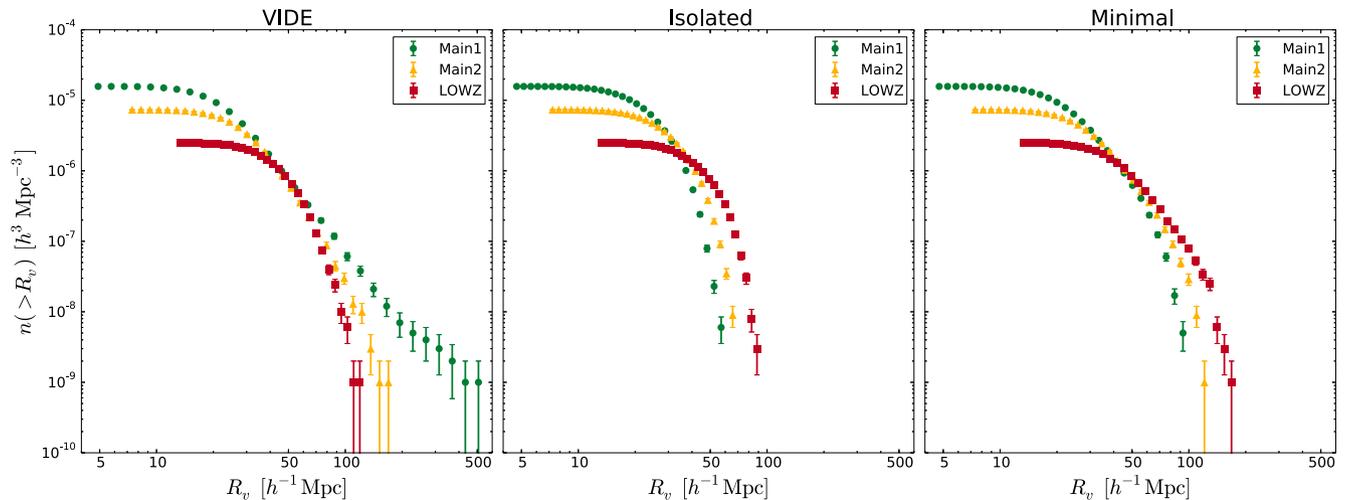
Figure 3. The cumulative number density of voids with radius $>R_v$, for all voids in the three different HOD galaxy mocks. The panels correspond to the different choices of the criteria to handle void merging described in Section 2.3. For the 'Isolated' and 'Minimal' merging criteria sparser and more highly biased galaxies consistently trace larger voids, as expected, but the situation is reversed for the VIDE choice of void merging.

and 'Minimal' merging choices. However, this intuitive picture is inverted for the VIDE case, for which the maximum void size in each sample instead decreases with increasing galaxy bias. This is because of the fixed threshold on tracer number density for merging, which has different physical consequences in each galaxy sample. It provides another reason to disfavour this choice of merging.

3.2 Dark matter within voids

Although it is the DM content of voids that is of primary interest for cosmology, in practice it is only the tracer number densities at void locations that can be directly observed. These two quantities are obviously distinct for galaxy tracers, but Nadathur & Hotchkiss (2015) showed that the tracer number density is not the same as the true matter density *even when the tracers are themselves a random subset of all DM particles*. This is especially true when the number densities are measured by simply counting tracer particles in a volume around special locations within the void, such as the void centre. In fact, since as described in Section 2.4 the void centre is by construction far from all tracers, naively reconstructing the tracer number density within a small enough volume around the centre returns a tracer number density contrast $\delta_n = -1$ for practically all voids. As the true matter densities within voids vary quite widely, this means that this naive measure of the tracer number density is not useful.

However, the VTFE-reconstructed tracer density is a better indicator of the true matter density (Nadathur & Hotchkiss 2015). This is because it is less subject to sparse sampling effects (indeed this is the main reason why the VTFE density estimator is used in the first place). We find that the minimum or core VTFE-reconstructed tracer density contrast within the void, $\delta_{n,\min}^{\mathrm{VTFE}} \equiv n_{\min}/\bar{n} - 1$, is a good predictor of the underlying matter density contrast $\delta_{\min}$ at the void centre.[4] Fig. 4 shows measurements of $\delta_{\min}$ for voids binned by

their values of $\delta_{\min}^{\mathrm{VTFE}} \equiv n_{\min}/\bar{n} - 1$, for each of our tracer samples. There is a simple linear relationship between the two, as evidenced by the goodness of the straight-line fits to the data. However, note that $\delta_{\min}^{\mathrm{VTFE}} \neq \delta_{\min}$ even when using DM tracers, and $\delta_{\min}^{\mathrm{VTFE}} \neq b\delta_{\min}$ for voids in the galaxy mocks.

Some qualitative aspects of Fig. 4 can be understood in terms of the role of bias and subsampling. At low values of $\delta_{\min}^{\mathrm{VTFE}}$, voids in galaxy mocks on average correspond to shallower DM underdensities than do voids in the subsampled DM tracers with the same $\delta_{\min}^{\mathrm{VTFE}}$. This is a result of the overall bias of the galaxies and the suppression of higher mass haloes in low-density regions – such regions therefore look even emptier in the galaxy number density. However, as $\delta_{\min}^{\mathrm{VTFE}}$ increases towards zero, the situation reverses: this is because subsampling of the DM particles increases the effect of shot noise and means that shallow depressions in the tracer density are less likely to correspond to true matter underdensities in this case than for the biased galaxy mocks. This is also consistent with the picture from Fig. 1, where voids in subsampled DM are much more likely to overlap with those in equivalent Poisson point sets.

These qualitative features are common to voids in each of the Main1, Main2 and LOWZ samples. However, at a more quantitative level a cursory search did not reveal a simple way to relate the slope and intercept of the straight-line fits to the tracer bias and mean number density. This suggests that in order to assess the depth of the matter underdensity in voids in observational data, calibration using realistic HOD mocks is required on a case-by-case basis.

A quantity of significant cosmological interest is the distribution of mass within galaxy voids, which can be determined by measuring the spherically averaged stacked matter density profile about void locations. Previous studies have generally attempted to describe the variation of the stacked profile with the mean size of the voids included in the stack. Barreira et al. (2015) have proposed a functional form with five free parameters:

$$\frac{\rho(r')}{\bar{\rho}} = 1 + \delta_c \left(\frac{1 - (r'/s_1)^\alpha}{1 + (r'/s_2)^\beta} \right), \tag{4}$$

where $r' = r/R_v$, to describe the resulting profiles. In Appendix A, we show that this function can provide a satisfactory fit to the

---

[4] Strictly speaking the correspondence between the void centre and the DM density minimum is statistical and may not hold for an individual void. However, any other location within the void that can be deduced on the basis of the tracer positions alone will on average always have a higher matter density.
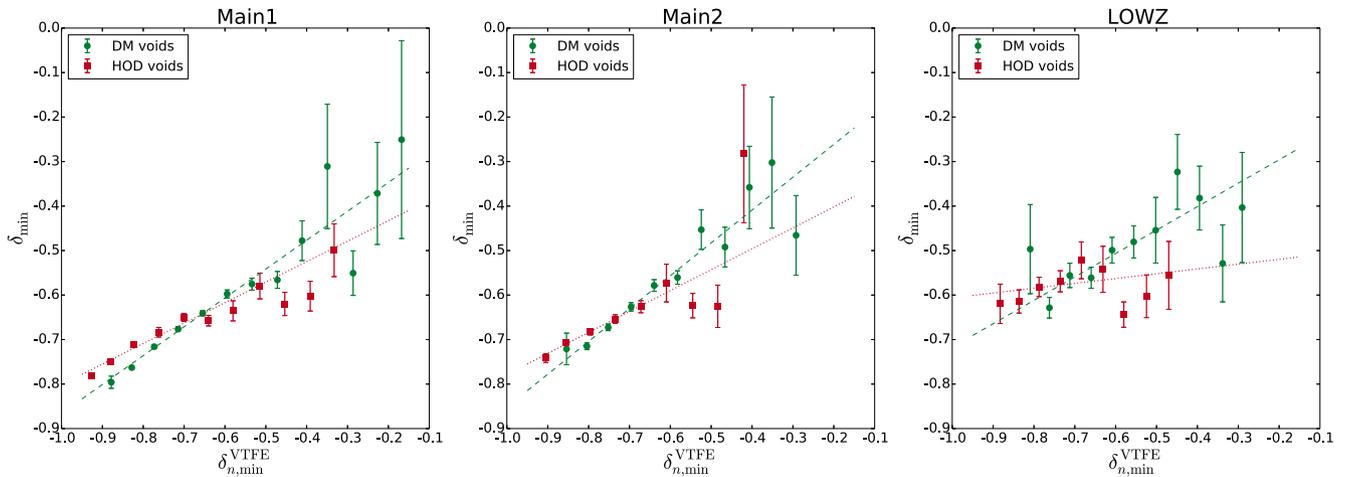
Figure 4. The binned average DM density at the void centre as a function of the minimum tracer number density determined from the VTFE reconstruction, for voids in different tracer populations. In each plot, green circles are for voids traced by subsampled DM particles and red squares are for voids in the corresponding biased galaxy mocks. Error bars indicate the 1σ errors in the mean. The best straight-line fits to the unbinned data are shown with dashed (dotted) lines for DM (HOD) voids in each case. At low densities, voids in HOD galaxy mocks show smaller values of $\delta_{\min}^{\mathrm{VTFE}}$ due to biasing effects. Shallow voids in the galaxy distribution are less affected by shot noise due to subsampling and so trace deeper matter underdensities.

stacked profile data for all void sizes R_{v}, and examine similar fitting forms proposed in other studies (Hamaus et al. 2014; Sutter et al. 2014; Nadathur et al. 2015). Unfortunately the behaviour of the fitted parameters with R_{v} differs across the different mock galaxy samples, meaning that equation (4) is not predictive, i.e. the best-fitting parameters must be determined on a case-by-case basis for voids in each sample.

In any case, describing the best-fitting profile parameters as a function of R_{v} alone is not a useful strategy, since it obscures the wide variation in tracer densities within voids of the same size that can be seen in Fig. 1. Indeed Nadathur & Hotchkiss (2015) showed explicitly that the mass density profile of voids is also strongly dependent on the tracer $n_{\min}$ values even for fixed R_{v}. To demonstrate the full variety of watershed void properties, in the next section we consider the dependence on $n_{\min}$ and R_{v} simultaneously.

3.3 Compensation of mass around voids

Describing the full variation of the density distribution around void centres in terms of a fitting formula such as equation (4) and the dependence of its free parameters on void properties is a complicated task. Instead we choose to describe the density environment around the void location by a single number, namely the total enclosed mass density contrast Δ within a sphere of radius three times the effective void radius R_{v}. This radius is large enough that Δ undergoes no further zero-crossings at larger r. Therefore the sign of $\Delta(r = 3R_{\mathrm{v}})$ distinguishes between two important classes of void environments. Overcompensated voids, with $\Delta > 0$, are local density minima embedded within larger scale contracting overdensities, reminiscent of the void-in-cloud scenario described by Sheth & van de Weygaert (2004). On the other hand, for $\Delta < 0$ the interior mass deficit of the void is not compensated by surrounding high-density regions. Such voids will correspond to regions of positive gravitational potential, $\Phi > 0$. Being able to distinguish between these two cases on the basis of primary void observables is important for studies of the gravitational effects of voids, particularly measurements of lensing and secondary CMB anisotropies due to voids.

Fig. 5 shows the dependence of $\Delta(r = 3R_{\mathrm{v}})$, measured from the full resolution simulation output, on $n_{\min}$ and R_{v} for voids in each of the three mock galaxy samples. The same trends are seen across all samples, suggesting the possibility of a universal description of the void environment. However, contrary to previous suggestions (Ceccarelli et al. 2013; Cai et al. 2014; Hamaus et al. 2014) the transition from overcompensation to undercompensation is not a function of the void size alone: both the smallest and largest values of Δ are found at small void radii. Shallower voids are on average more likely to be overcompensated, so a cut based on the minimum tracer density at roughly $n_{\min}/\bar{n} \lesssim 0.3$ (as applied by Hotchkiss et al. 2015) would better distinguish undercompensated voids, but is still not completely satisfactory.

The dependence of void properties on variables other than the void radius is significant for their theoretical modelling. Thus far, such models have followed Sheth & van de Weygaert (2004), who develop an excursion set model roughly analogous to that for haloes. In this scenario, the void radius R_{v} is analogous to the halo mass and so all void properties, in particular the variation of the average density profile, are described purely with reference to R_{v}. Fig. 5 shows that this analogy does not work well, because it fails to account for the strong systematic dependence on $n_{\min}$. In a companion paper (Nadathur & Hotchkiss 2015), we have also shown that watershed voids fail to satisfy several other important assumptions of the Sheth & van de Weygaert (2004) model.

It is however possible to ignore the dependence on $n_{\min}$ and describe the average behaviour of Δ as a function of R_{v} alone, as has been done in previous works. The resulting behaviour is shown in Fig. 6. For all samples, small voids are on average overcompensated ($\Delta > 0$), but approximate compensation is reached at large R_{v}. Note that all information of the undercompensated voids clearly visible in Fig. 5 is lost in marginalizing over $n_{\min}$. The radius scale above which voids are on average compensated is dependent on both the tracer density and bias.

In Fig. 7 we show the dependence of Δ instead on the average tracer number density within a void, n_{avg}, for all six tracer populations used in this study. Using n_{avg} as the independent variable preserves information on both under- and overcompensated voids.
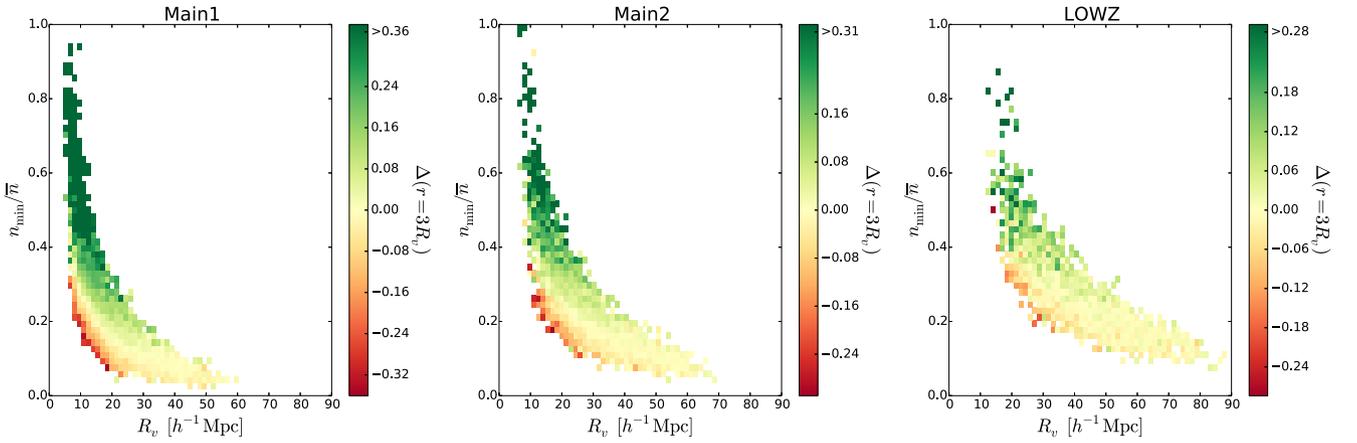
Figure 5. The variation of void compensation as a function of minimum tracer density and void size. The colour bar indicates the binned average values of the total matter density contrast enclosed within a sphere of radius $3R_v$ centred at the void locations. The same trends are seen for voids in all three galaxy mocks: small and shallow voids are on average overcompensated, while deep and large voids are undercompensated, but the variation depends on a combination of $n_{\min}$ and R_v and is not satisfactorily accounted for by either variable alone.
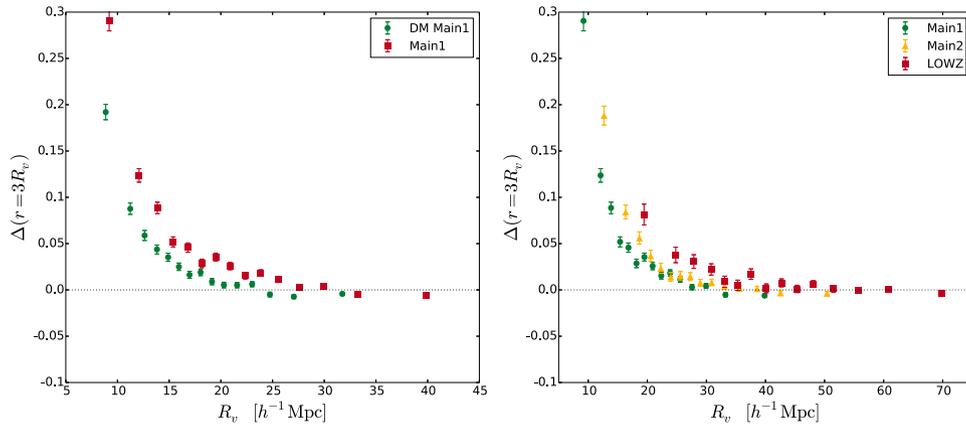
Figure 6. The mean enclosed matter density contrast $\Delta(r = 3R_v)$ for voids as a function of the void radius R_v alone. Left: comparison between voids in the Main1 HOD sample and those in the unbiased DM Main1 tracers, which have the same mean tracer density $\overline{n}$. Similar results hold for other pairs of samples with the same tracer density. Right: voids in the three HOD mock galaxy samples. In all cases, small voids are on average overcompensated ($\Delta > 0$) and large voids on average exactly compensated ($\Delta \simeq 0$). The radius scale for the approach to compensation depends on both $\overline{n}$ and the tracer bias. Information on undercompensated voids seen in Fig. 5 is lost when through averaging over $n_{\min}$.

In all cases a simple linear relationship exists between n_{avg} and Δ. Straight-line fits to the data show that $n_{\mathrm{avg}} = \overline{n}$ is the universal dividing line between undercompensated and overcompensated voids in every case. We conclude that, in contrast to R_v, n_{avg} is a universal indicator of the compensation of the mass deficit within a void, independent of the bias properties or sampling density of the tracer population in which the voids are identified.

At one level, this result appears intuitively simple: regions containing a higher than average number density of tracers, whether DM particles or galaxies, also correspond to a higher than average matter density. However, note that n_{avg} is defined for voids of arbitrary – and generally aspherical – shape determined by the topology of the tracer density field, whereas Δ refers to the mass density within idealized spheres. There is also a large difference in scales, since Δ is measured within a volume 27 times larger than that of the void.

It is also worth noting that the relationship between *local* tracer and matter densities at the void centres is not independent of the

properties of the tracer population, as shown in Fig. 4. Voids in different samples trace minima of very different depths. Similarly we were also unable to find any comparable relationship between the locally measured tracer and matter densities at void locations, i.e. the profiles of $\delta(r)$ are not proportional to those of $\delta_n(r)$ for any tracer type. This simple linear relationship exists only for integrated quantities n_{avg} and Δ.

4 CONCLUSIONS

The existing theoretical model of voids based on the excursion set formalism of Sheth & van de Weygaert (2004) has previously been shown to be inadequate for describing the properties of voids that can be practically identified in simulations by means of the commonly used watershed algorithm (Nadathur & Hotchkiss 2015). This means that meaningful predictions for void properties and the cosmological constraints that may be obtained from the use of voids can at present only be obtained through a simulation-led approach,
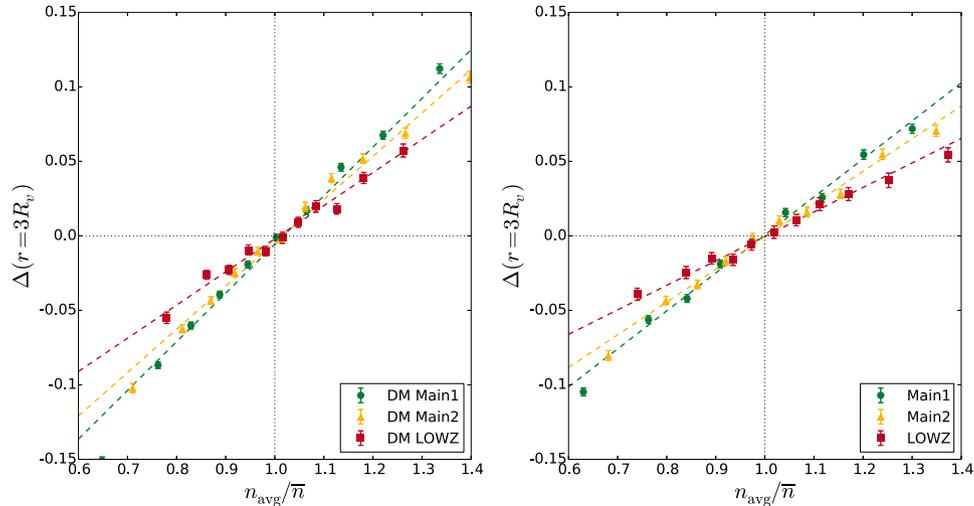
Figure 7. Void compensation as measured by the enclosed matter density contrast $\Delta(r = 3R_v)$ from equation (3), as a function of the overall density of tracers within the void, n_{avg}. The left-hand panel shows the trends for voids in subsampled DM, and the right-hand panel for voids in galaxy tracers. The dashed lines show the best fits to the data; in all cases a simple linear relationship provides an excellent description. The value $n_{avg} = \overline{n}$ provides a universal distinction between undercompensated and overcompensated voids, irrespective of the properties of the tracer population in which voids are identified.

but to be observationally relevant such simulations must aim to replicate the types of voids that would be seen in actual galaxy survey data.

Our aim in this paper has been to provide a comprehensive examination of the twin roles of galaxy bias and sparsity in determining the properties of voids, and in the process to improve existing treatments in the literature. To do so we made use of a suite of mock galaxy samples in an *N*-body simulation, which were used as tracer populations for void identification. These void catalogues were compared with equivalent ones obtained using unbiased subsets of the DM particles in the simulation as tracers. Although our results were obtained specifically using the popular ZOBOV algorithm, the general qualitative features of voids we describe will be applicable to any void finder based on the watershed principle.

Our first practical result has been to show that, contrary to previous claims (Sutter et al. 2014, 2015a; Zivick et al. 2015), tracer bias plays a crucially important role in determining the overall abundance of voids, their size distributions and matter content. The differences in such void properties introduced by tracer bias are comparable to or larger than those reported from differences in cosmology (Yang et al. 2015; Zivick et al. 2015; Pisani et al. 2015; Sutter et al. 2015a). This result is not entirely unexpected, since all of these quantities naturally depend strongly on the clustering properties of the tracer population, which are in turn related to the tracer bias. However, it means that the commonly used strategy of approximating galaxy voids by simply downsampling the density of DM particles in a simulation is incorrect and that predictions for void observables calibrated on such simulation results will unfortunately bear little or no relation to data obtained from real galaxy surveys. A similar conclusion has been independently reached by Pollina et al. (2015).

Our results have also clarified the general nature of watershed voids. Such voids correspond to density minima of widely varying depths typically in the linear regime, rather than the highly non-linear evolved objects postulated by some theoretical models. Most previous studies of voids, from both theoretical and simulation perspectives, have focused on describing void properties as functions

of the void size alone. In reality the properties of voids of the same size show a strong systematic dependence on other factors as well. Our work here and in the companion paper Nadathur & Hotchkiss (2015) has aimed to highlight the important role of tracer number densities in particular as predictors of the matter content of voids.

From this perspective, the most important results in this paper are shown in Figs 5 and 7, which show how the large-scale matter fluctuation around void locations can be predicted from easily measurable properties of the tracer galaxies within the void, irrespective of the detailed clustering properties of the galaxy population. In particular, the average galaxy number density within the void can serve as a universal diagnostic of the mass compensation on much larger scales.

The nature of this mass compensation is interesting in its own right and as a clue towards the formation and evolution history of voids. However, it is primarily important for the use of voids in cosmology because it serves as a signifier of the value and sign of the gravitational potential Φ at the void location. For instance, for any observational measurement of the gravitational effects of voids, such as through lensing distortions or secondary CMB anisotropies, one would at a minimum like to exclude voids corresponding with $\Phi < 0$. In the past, various authors have attempted to use other directly observable properties of watershed voids, such as the ratio of densities in the void wall and at its centre (Granett et al. 2008), the void size (Cai et al. 2014; Planck Collaboration XIX 2014) or the minimum galaxy number density within the void (Hotchkiss et al. 2015) as proxies for the value of Φ. Our result indicates instead that n_{avg} is a better indicator, and that this holds for voids seen in any galaxy sample. A generalization of this measure to a weighted average density for galaxies with a redshift-varying selection function should also be quite straightforward.

We should at this point introduce a note of caution. Strictly speaking, our results in this paper concern the compensation of mass in voids, and not the potential itself. We have used Δ as a single number with which to characterize the total large-scale mass deficit at void locations, but the gravitational potential depends also on the details of the density profile and the size of the void, as it is related

to an integral over r of $\Delta(r)$. Nevertheless, it remains the case that n_{avg} is an excellent predictor of Φ, as we will show in forthcoming work.

REFERENCES

Aiola S., Kosowsky A., Wang B., 2015, Phys. Rev. D, 91, 043510
Alonso D., Eardley E., Peacock J. A., 2015, MNRAS, 447, 2683
Barreira A., Cautun M., Li B., Baugh C., Pascoli S., 2015, J. Cosmol. Astropart. Phys., 1508, 28
Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
Bos E. G. P., van de Weygaert R., Dolag K., Pettorino V., 2012, MNRAS, 426, 440
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Cai Y.-C., Neyrinck M. C., Szapudi I., Cole S., Frenk C. S., 2014, ApJ, 786, 110
Cai Y.-C., Padilla N., Li B., 2015, MNRAS, 451, 1036
Ceccarelli L., Paz D., Lares M., Padilla N., Lambas D. G., 2013, MNRAS, 434, 1435
Chan K. C., Hamaus N., Desjacques V., 2014, Phys. Rev. D, 90, 103521
Clampitt J., Cai Y.-C., Li B., 2013, MNRAS, 431, 749
Falck B., Neyrinck M. C., 2015, MNRAS, 450, 3239
Flender S., Hotchkiss S., Nadathur S., 2013, J. Cosmol. Astropart. Phys., 1302, 013
Furlanetto S., Piran T., 2006, MNRAS, 366, 467
Gottlöber S., Łokas E. L., Klypin A., Hoffman Y., 2003, MNRAS, 344, 715
Granett B. R., Neyrinck M. C., Szapudi I., 2008, ApJ, 683, L99
Hamaus N., Sutter P. M., Wandelt B. D., 2014, Phys. Rev. Lett., 112, 251302
Hotchkiss S., Nadathur S., Gottlöber S., Iliev I. T., Knebe A., Watson W. A., Yepes G., 2015, MNRAS, 446, 1321
Klypin A., Holtzman J., 1997, preprint (arXiv:astro-ph/9712217)
Krause E., Chang T.-C., Doré O., Umetsu K., 2013, ApJ, 762, L20
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Lee J., 2012, ApJ, 752, 40
Lee J., Park D., 2009, ApJ, 696, L10
Li B., Zhao G.-B., Koyama K., 2012, MNRAS, 421, 3481
Manera M. et al., 2015, MNRAS, 447, 437
Melchior P., Sutter P. M., Sheldon E. S., Krause E., Wandelt B. D., 2014, MNRAS, 440, 2922
Nadathur S., Hotchkiss S., 2014, MNRAS, 440, 1248
Nadathur S., Hotchkiss S., 2015, MNRAS, preprint (arXiv:1504.06510)
Nadathur S., Hotchkiss S., Sarkar S., 2012, J. Cosmol. Astropart. Phys., 1206, 042
Nadathur S., Hotchkiss S., Diego J. M., Iliev I. T., Gottlöber S., Watson W. A., Yepes G., 2015, MNRAS, 449, 3997
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Neyrinck M. C., 2008, MNRAS, 386, 2101
Neyrinck M. C., Aragón-Calvo M. A., Jeong D., Wang X., 2014, MNRAS, 441, 646
Pan D. C., Vogeley M. S., Hoyle F., Choi Y.-Y., Park C., 2012, MNRAS, 421, 926
Pisani A., Sutter P. M., Hamaus N., Alizadeh E., Biswas R., Wandelt B. D., Hirata C. M., 2015, preprint (arXiv:1503.07690)
Planck Collaboration XIX, 2014, A&A, 571, A19
Planck Collaboration XXI 2015, preprint (arXiv:1502.01595)
Pollina G., Baldi M., Marulli F., Moscardini L., 2015, arXiv e-prints
Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018
Ricciardelli E., Quilis V., Varela J., 2014, MNRAS, 440, 601
Riebe K. et al., 2013, Astron. Nachr., 334, 691
Ryden B. S., 1995, ApJ, 452, 25
Sheth R. K., Tormen G., 2002, MNRAS, 329, 61
Sheth R. K., van de Weygaert R., 2004, MNRAS, 350, 517
Sutter P., Lavaux G., Wandelt B. D., Weinberg D. H., 2012, ApJ, 761, 44
Sutter P. M., Lavaux G., Hamaus N., Wandelt B. D., Weinberg D. H., Warren M. S., 2014, MNRAS, 442, 462
Sutter P. M., Carlesi E., Wandelt B. D., Knebe A., 2015a, MNRAS, 446, L1
Sutter P. M. et al., 2015b, Astron. Comput., 9, 1
Tikhonov A. V., Gottlöber S., Yepes G., Hoffman Y., 2009, MNRAS, 399, 1611
Yang L. F., Neyrinck M. C., Aragon-Calvo M. A., Falck B., Silk J., 2015, MNRAS, 451, 3606
Zehavi I. et al., 2011, ApJ, 736, 59
Zheng Z., Coil A. L., Zehavi I., 2007, ApJ, 667, 760
Zivick P., Sutter P. M., Wandelt B. D., Li B., Lam T. Y., 2015, MNRAS, 451, 4215

APPENDIX A: VOID DENSITY PROFILES AND FITTING FORMULAE

A1 Matter density profiles

Although stacked void density profiles have been studied by several authors, results in the literature are rather contradictory (e.g. Hamaus, Sutter & Wandelt 2014; Ricciardelli, Quilis & Varela 2014; Sutter et al. 2014; Barreira et al. 2015; Nadathur et al. 2015; Nadathur & Hotchkiss 2015). We start by considering the dependence of the average density profile on the void radius R_v. To do this, we create stacks of voids in bins of R_v and centred on the void centre, and determine the average profile $\rho(r)/\overline{\rho}$ for each stack using the gridded density information from the full resolution simulation output. From this data we determine the best-fitting parameters δ_c, s_1, s_2, α and β in equation (4) for each stack.

Fig. A1 shows the profile data for voids of different sizes in each galaxy sample, along with the best-fitting forms of equation (4) (solid curves) in each case. Hamaus et al. (2014) have proposed a 'universal' fitting form for the profile, which corresponds to a special case of equation (4) with $s_2 = 1$:

$$\frac{\rho(r)}{\overline{\rho}} = 1 + \delta_c \left(\frac{1 - (r/r_s)^{\alpha}}{1 + (r/R_v)^{\beta}} \right). \tag{A1}$$

The best-fitting versions of this constrained profile are shown with dashed lines in Fig. A1. It can be seen that they generally provide somewhat worse fits to the data, both inside the void and in the compensating shell.

A key feature of both fits is that, as shown in Fig. A2, within the sample of voids from each population the fitted parameter δ_c is a decreasing function of R_v. This is consistent with (and can be indirectly deduced from) the generic watershed behaviour shown in Figs 1 and 4, and also agrees with the results of Nadathur & Hotchkiss (2015) where we found that larger voids contain deeper density minima. However, it is opposite to the behaviour claimed in Hamaus et al. (2014), who find δ_c increases with R_v. For the other fitted parameters we find no consistent trends across the different galaxy samples, so we are also unable to confirm the claim of universality of the void density profile made by those authors. In
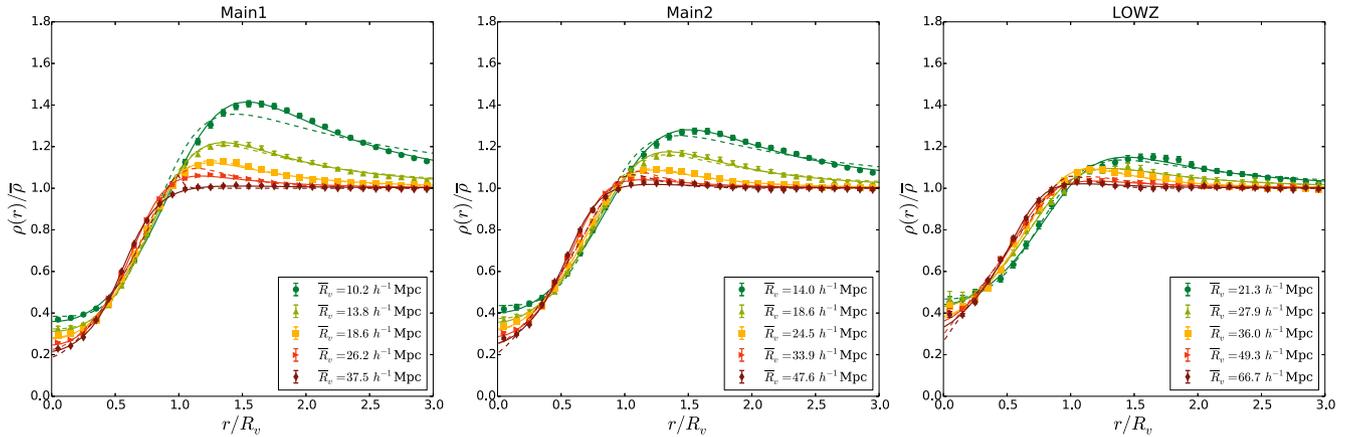
Figure A1. Average density profiles for voids of different sizes. Voids in each mock galaxy population are binned on the basis of their radius R_v and stacked so that their centres coincide. The average matter density distribution for the stack is then determined from the DM field of the simulation. Solid lines show the best-fitting forms of the profile in equation (4) in each case, and dashed lines show the best-fitting forms of equation (A1). Some radius bins are omitted for clarity.
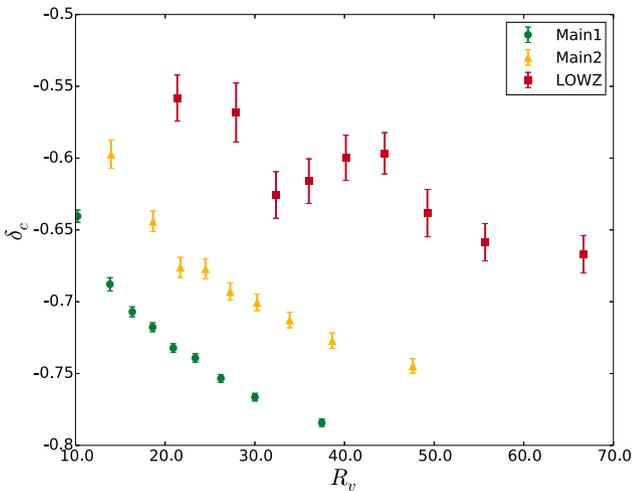

Figure A2. The dependence of the best-fitting values of the parameter δ_c in equation (4) as a function of the mean void radius of the stacks shown in Fig. A1. In each case δ_c decreases with R_v, and this trend is unchanged if the alternative fitting form in equation (A1) is used, or for stacks centred on the void barycentres.

particular, the dependence of α and β on scale radius r_s does not follow the pattern described by Hamaus et al. (2014) and Sutter et al. (2014) (see Nadathur & Hotchkiss 2015 for more details).

It should be noted that Hamaus et al. (2014) proposed equation (A1) as a description of the average density profile for stacks centred about void barycentres, whereas in this paper we define the void centre differently to better identify the true location of the underdensity. In Nadathur & Hotchkiss (2015) we showed that for some high-density tracer populations and for some choices of merging criteria, the average density at the barycentre can show a non-monotonic behaviour with increasing void size (see fig. 3 in particular). This occurs due to the displacement of the barycentre location from the true underdensity in voids composed of very large numbers of merged subvoids. Even in these situations, however, δ_c does not increase monotonically with R_v as claimed by Hamaus et al. (2014). This unusual behaviour also does not apply for any of the galaxy samples used in this work, which all show monotonically *decreasing* $\delta_c(R_v)$ irrespective of the choice of void centre.

However, it is worth stressing again that analysing the variation of the mean density profile as a function of R_v obscures the important systematic variation across the void population with the values of n_{min} shown in Fig. 5 and discussed in Nadathur & Hotchkiss (2015). The stacked profiles in R_v bins in this appendix are therefore included only to enable comparisons with previous results in the literature. For voids of the same size, the variation of the stacked density profile in stacks of the same n_{min} is large, as expected from Fig. 5. If voids are stacked according to their n_{min} or n_{avg} values rather than R_v, the form of equation (A1) is completely unable to fit the resulting profiles. However, the form of equation (4) continues to provide a good description in all cases.

A2 Galaxy number density profiles

In a previous work (Nadathur et al. 2015), we found that the stacked galaxy number density profile for a selected subset of voids in several different galaxy populations drawn from data from SDSS surveys at different redshifts were very similar, and argued that this might be a sign of universality. However, the size of the data samples used in this work was small, and the errors in determination of the profiles were correspondingly large. The number of voids in our simulation samples exceeds that used in the previous work by a factor of ~10: this is because the simulation volume is much larger than the survey volume for some of the observational samples, and boundary effects due to survey masks are not present. We are therefore able to re-examine the universality of galaxy density profiles with greater statistical precision.

Nadathur et al. (2015) analysed only those voids with $n_{min}/\overline{n} < 0.3$. In the absence of information of the true matter content of galaxy voids, this very conservative cut was intended to remove Poisson contamination of the void sample, based on calibration of the probability $P(n_{min}|\mathrm{Poisson})$. Ad hoc cuts to the void population on the basis of void radius R_v have been advocated elsewhere in order to achieve the same purpose. Our analysis of the matter density content of voids in simulation, both in this paper and in Nadathur & Hotchkiss (2015), indicates that on average all voids correspond to true matter underdensities, so such cuts are not necessary and lead to suboptimal use of available data. However, for the purposes of direct comparison with the results of Nadathur et al. (2015), we apply the same methodology here. We then stack the surviving voids in each
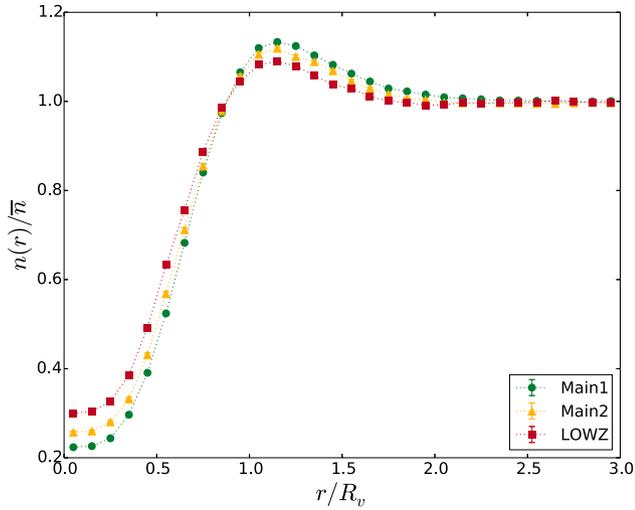
Figure A3. Galaxy number density profiles for voids in the three mock galaxy samples. To match the methodology used by Nadathur et al. (2015) a selection cut $n_{\min}/\overline{n} < 0.3$ is applied and stacks are centred on void barycentres. The resulting average profiles in the different galaxy populations are similar but statistically distinguishable given the large sample size. The dotted lines are linear interpolations and do not represent fits to the data.

of the Main1, Main2 and LOWZ mock galaxy samples separately, without binning by radius, to determine the average galaxy number density profiles. For consistency, we also use the 'VTFE' estimator for measuring galaxy density profiles discussed in that work and centre stacks on the void barycentres, though the results are not strongly dependent on this choice. The resulting profiles are shown in Fig. A3. The results in the different galaxy populations are similar to each other, but distinguishable with high statistical significance due to the small error bars. However they are also consistent with the profile seen in Nadathur et al. (2015) within the much larger statistical errors of that study.

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