



# Voids in the SDSS DR9: observations, simulations, and the impact of the survey mask

P. M. Sutter, Guilhem Lavaux, Benjamin D. Wandelt, David H. Weinberg,  
Michael S. Warren, Alice Pisani

## ► To cite this version:

P. M. Sutter, Guilhem Lavaux, Benjamin D. Wandelt, David H. Weinberg, Michael S. Warren, et al.. Voids in the SDSS DR9: observations, simulations, and the impact of the survey mask. Monthly Notices of the Royal Astronomical Society, 2014, 442, pp.3127-3137. 10.1093/mnras/stu1094 . insu-03645337

**HAL Id: insu-03645337**

**<https://insu.hal.science/insu-03645337>**

Submitted on 25 Apr 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Voids in the SDSS DR9: observations, simulations, and the impact of the survey mask

P. M. Sutter,<sup>1,2,3,4★</sup> Guilhem Lavaux,<sup>1,2,5,6,7</sup> Benjamin D. Wandelt,<sup>1,2,4,8</sup>  
David H. Weinberg,<sup>3,9</sup> Michael S. Warren<sup>10</sup> and Alice Pisani<sup>1,2</sup>

<sup>1</sup>*Sorbonne Universités, UPMC Univ Paris 06, UMR7095, Institut d'Astrophysique de Paris, F-75014 Paris, France*

<sup>2</sup>*CNRS, UMR7095, Institut d'Astrophysique de Paris, F-75014 Paris, France*

<sup>3</sup>*Center for Cosmology and Astro-Particle Physics, Ohio State University, Columbus, OH 43210, USA*

<sup>4</sup>*Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*

<sup>5</sup>*Department of Physics & Astronomy, University of Waterloo, Waterloo, ON N2L 3G1, Canada*

<sup>6</sup>*Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada*

<sup>7</sup>*Canadian Institute for Theoretical Astrophysics, 60 St George St, Toronto, ON M5S 3H8, Canada*

<sup>8</sup>*Department of Astronomy, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*

<sup>9</sup>*Department of Astronomy, Ohio State University, Columbus, OH 43210, USA*

<sup>10</sup>*Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

Accepted 2014 June 1. Received 2014 May 31; in original form 2013 October 26

## ABSTRACT

We present and study cosmic voids identified using the watershed void finder `VIDE` in the Sloan Digital Sky Survey Data Release 9, compare these voids to ones identified in mock catalogues, and assess the impact of the survey mask on void statistics such as number functions, ellipticity distributions, and radial density profiles. The nearly 1000 identified voids span three nearly volume-limited samples from redshift  $z = 0.43$  to  $0.7$ . For comparison, we use 98 of the publicly available second-order Lagrangian perturbation theory-based mock galaxy catalogues of Manera et al., and also generate our own mock catalogues by applying a Halo Occupation Distribution model to an  $N$ -body simulation. We find that the mask reduces the number density of voids at all scales by a factor of 3 and slightly skews the relative size distributions. This engenders an increase in the mean ellipticity by roughly 30 per cent. However, we find that radial density profiles are largely robust to the effects of the mask. We see excellent agreement between the data and both mock catalogues, and find no tension between the observed void properties and the properties derived from  $\Lambda$  cold dark matter simulations. We have added the void catalogues from both data and mock galaxy populations discussed in this work to the Public Cosmic Void Catalog at <http://www.cosmicvoids.net>.

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

## 1 INTRODUCTION

With the recent advent of large-scale comprehensive void catalogues (Pan et al. 2012; Sutter et al. 2012a), cosmological analysis is beginning to fan out from probes solely focused on overdensities such as galaxy correlations (Sánchez et al. 2012; Marín et al. 2013) and baryon acoustic oscillations (Bassett & Hlozek 2010) to more general studies based on alternative information sources available in the cosmic web. Since the primary target of cosmological analysis is often quantifying and understanding dark energy (Weinberg et al. 2013), exploiting the underdense voids in the matter distribution of

the universe is a natural choice: the interiors of voids are dominated by dark energy (Goldberg & Vogeley 2004), so their shapes, sizes, and growth histories are intimately tied to the global properties of the large-scale universe (Thompson & Gregory 2011).

Already researchers have begun to exploit the public void catalogues. Ilic, Langer & Douspis (2013) correlated void positions with *Wilkinson Microwave Anisotropy Probe* (WMAP) measurements of cosmic microwave background temperature anisotropies (Komatsu et al. 2011) to obtain a weak measurement of the integrated Sachs–Wolfe effect (Thompson & Vishniac 1987). The Planck Collaboration followed up on this study to confirm a detection (Planck Collaboration 2013a). Melchior et al. (2014) have performed a measurement of gravitational weak lensing around voids in data (theoretically predicted by Krause et al. 2013; Higuchi, Oguri &

★E-mail: [sutter@iap.fr](mailto:sutter@iap.fr)

Hamana 2013) to directly measure the underdensities in the dark matter. Pisani et al. (2013) used a novel method to measure the real-space radial density profiles within voids without assumptions about cosmological or redshift-space distortion models. Finally, Sutter et al. (2012b) began to measure cosmological parameters by leveraging the statistical isotropy of stacked voids (Lavaux & Wandelt 2012) to perform an Alcock–Paczynski test (Alcock & Paczynski 1979; Ryden 1995).

Looking ahead, there are many more promising applications of voids to cosmology and astrophysics. At the most simple level, the size distribution of voids is sensitive to cosmological parameters (Jennings, Li & Hu 2013) and modified gravity (Clampitt, Cai & Li 2013), though these effects can be confused by uncertainties in galaxy formation physics (Little & Weinberg 1994; Muller et al. 2000; Tinker & Conroy 2009). A measurement of the shapes of voids as encoded by the mean ellipticity would shed light on dark energy (Biswas, Alizadeh & Wandelt 2010; Bos et al. 2012) as well as the two-point correlation of the void positions (Padilla, Ceccarelli & Lambas 2005; Paranjape, Lam & Sheth 2012; Hamaus et al. 2014a). The radial density profile, reconstructed in real space using techniques such as those described in Pisani et al. (2013), can also be used to constrain exotic dark energy models (Shoji & Lee 2012; Spolyar, Sahlén & Silk 2013). Astrophysically, voids can also be used to measure primordial magnetic fields (Taylor, Vovk & Neronov 2011; Beck et al. 2013) and the effects of environment on galaxy formation (Gottlober et al. 2003; Rojas et al. 2004; Hoyle et al. 2005; Rojas et al. 2005; Ferreras & Pasquali 2011; Ceccarelli et al. 2012; Hoyle, Vogeley & Pan 2012).

To support current and future void-based science efforts, we must continue to identify voids in the latest galaxy surveys. This way we can take advantage of deeper and wider surveys for a greater redshift lever arm for cosmological parameter estimation and for more volume for increasing the signal to noise of statistical void properties. Also, even though current surveys such as the Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013) may not be optimal for void analysis due to their relatively low sampling density, we can use void catalogues from data to test and calibrate results against theoretical expectations in preparation for larger volume surveys in the future such as *Euclid* (Laureijs et al. 2011), BigBOSS (Schlegel et al. 2011), and WFIRST (Spergel et al. 2013).

Ever since Peebles (2001) pointed out a potential discrepancy between the interior contents of voids in  $\Lambda$  cold dark matter ( $\Lambda$ CDM) predictions and observations, there has been intense interest in comparing voids between simulations and observations. This has been done for earlier void catalogues in the 2-Degree Field Galaxy Redshift Survey (2dFGRS; Benson et al. 2003; Hoyle & Vogeley 2004; Ceccarelli et al. 2006) and the Sloan Digital Sky Survey (SDSS; Strauss et al. 2002) Data Release 7 (Pan et al. 2012), and in earlier surveys (Einasto et al. 1991; Weinberg & Cole 1992; Little & Weinberg 1994; Vogeley et al. 1994) but in a very restricted context: it is difficult to build simulations with high enough resolution to capture all the survey galaxies and sufficient size to enclose the entire survey volume. Rather than attempt to reproduce complex observational details such as survey geometry, typically authors take a limited volume within the survey and compare the statistical properties of the remaining voids to voids identified in a galaxy population generated with semi-analytic modelling (e.g. Tavasoli, Vasei & Mohayaee 2013). This common approach has several shortcomings: it is difficult to precisely tune semi-analytic models to a given survey (Baugh et al. 2003) and it does not take advantage of the full

survey volume. We can address any potential discrepancies in a more robust way by building nearly identical survey-like populations in our simulations.

An examination of the impacts of the survey mask is especially important, since only selecting voids far away from the survey area discards much useful information, and without rigorous void selection there may still be residual systematics. Also, since theoretical work with voids is done in simulations with cubic volumes, understanding the role of the mask is essential for building the links between theory and data. Since survey masks usually have complicated shapes, their impact is highly non-trivial, non-obvious, and different for each survey. von Benda-Beckmann & Mueller (2007) noted differences between masked and unmasked void populations in the 2dFGRS, although Pan et al. (2012) did not find significant differences when examining the properties of voids with their void finding algorithm in a low-redshift volume-limited sample of SDSS galaxies. However, there has been no such examination in higher redshifts of the SDSS with the *VIDE* algorithm (Sutter et al. 2014c), which is the source of the current large void catalogues.

We explore another important link, the impacts of sparsity and galaxy bias, in another work (Sutter et al. 2014a), while earlier works such as Ryden & Melott (1996) have connected redshift-space voids to those in real space.

In this work, we present voids in the SDSS Data Release 9 CMASS sample (Ahn et al. 2012), a survey covering 3000 deg<sup>2</sup> from redshift 0.43 to 0.7, for a total volume of nearly 1.5 cubic  $h^{-1}$  Gpc. We compare these voids to voids found in two sets of mock catalogues: the published mocks of Manera et al. (2013) and our own derived from a large-volume high-resolution  $N$ -body simulation. In the simulation, we are able to capture the entire survey without overlapping or stitching simulation volumes, allowing us to examine the systematic impacts of the survey mask on void statistical properties such as number functions, radial profiles, and ellipticities. While the low galaxy density of this survey is not ideal for void identification, the analysis of Sutter et al. (2014b) indicates that voids found in surveys of this type still correspond to physical underdensities in the dark matter, and thus are still useful probes of cosmology and astrophysics. In addition, Sutter et al. (2014a) shows that the universal density profile of Hamaus, Sutter & Wandelt (2014b) fits voids identified in all densities of samples, and that there exist simple scaling relations between voids in different samples, which means that the objects identified even in low-density surveys correspond to voids.

In the next section, we establish our coordinate system and briefly discuss our void-finding method and strategies for handling masks in the survey data. Section 3 introduces our galaxy survey samples and the properties of the voids identified in them. In Section 4, we present our mock galaxy populations and compare voids in these masked and unmasked populations to the voids in the data. Finally, Section 5 offers concluding remarks regarding implications for future surveys and void-based science.

## 2 VOID FINDING

For each galaxy in the survey, we transform its sky latitude  $\theta$ , sky longitude  $\phi$ , and redshift  $z$ , to a comoving coordinate system

$$\begin{aligned} x' &= D_c(z) \cos \phi \cos \theta, \\ y' &= D_c(z) \sin \phi \cos \theta, \\ z' &= D_c(z) \sin \theta, \end{aligned}$$

where  $D_c(z)$  is the comoving distance to the galaxy at redshift  $z$ . We assume a  $\Lambda$ CDM cosmology consistent with WMAP 7-year results (Komatsu et al. 2011):  $\Omega_M = 0.27$ ,  $\Omega_\Lambda = 0.73$ , and  $h = 0.71$ .

We identify voids with a modified and extended version of ZOBOV (Neyrinck 2008; Lavaux & Wandelt 2012; Sutter et al. 2012a), called VIDE (Sutter et al. 2014c). VIDE creates a Voronoi tessellation of the tracer particle population and uses the watershed transform to group Voronoi cells into zones and voids (Platen, van de Weygaert & Jones 2007). The watershed transform identifies catchment basins as the cores of voids and ridgelines, which separate the flow of water, as the boundaries of voids. The watershed transform builds a nested hierarchy of voids (Bos et al. 2012; Lavaux & Wandelt 2012), and for the purposes of this work we only examine *root* voids, which are voids at the base of the tree hierarchy and hence have no parents. We also impose two density-based criteria on our void catalogue. The first is a threshold cut within VIDE itself where voids only include as additional members Voronoi zones with density less than 0.2 the mean particle density. If a void consists of only a single zone (as they often do in sparse populations), then this restriction does not apply. We apply the second density criterion as a post-processing step: we only include voids with mean central densities below 0.2 the mean particle density. We measure this central density within a sphere with radius  $R = 0.25R_{\text{eff}}$ , where

$$R_{\text{eff}} \equiv \left( \frac{3}{4\pi} V \right)^{1/3}. \quad (1)$$

In the expression above,  $V$  is the sum of the Voronoi volumes of the particles which comprise the void. We also ignore voids with  $R_{\text{eff}}$  below the mean particle spacing of the tracer population.

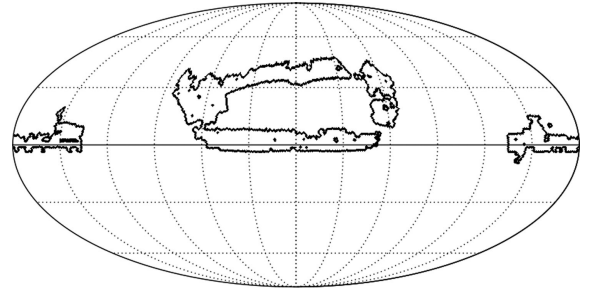
Additionally, for the analysis below we need to define a centre for the void. For our work, we take the barycentre, or volume-weighted centre of all the Voronoi cells in the void:

$$\mathbf{X}_v = \frac{1}{\sum_i V_i} \sum_i \mathbf{x}_i V_i, \quad (2)$$

where  $\mathbf{x}_i$  and  $V_i$  are the positions and Voronoi volumes of each tracer  $i$ , respectively.

As presented in Sutter et al. (2012a) and Sutter et al. (2014c), VIDE includes modifications to ZOBOV to account for survey boundaries, internal masks, and redshift limits. To handle line-of-sight boundaries and internal masks, we pixelize the survey region using HEALPIX (Gorski et al. 2005)<sup>1</sup> and identify boundary pixels (i.e. pixels with at least one non-survey region neighbour). We inject particles along the line of sight within each boundary pixel with a spatial density of  $10^{-3}(h^{-1} \text{ Mpc})^{-3}$ . By giving these boundary particles essentially infinite density and breaking their degeneracies in the Voronoi graph, we prevent the watershed algorithm from growing voids outside the survey region. Also, these boundary particles serve as a marker for identifying voids near the edge. Fig. 1 shows our identification of the SDSS DR9 (Ahn et al. 2012) survey boundary pixels. To accurately capture the shape of the mask, we required a resolution of  $N_{\text{side}} = 512$  ( $\sim 2-3 h^{-1} \text{ Mpc}$  at  $z = 0.7$  in a  $\Lambda$ CDM universe).

Following the procedure of Sutter et al. (2012a), we generate two void catalogues for each sample, which we call *all* and *central* catalogues. Naturally, the *all* sample contains every identified void which satisfied the density cutoff criteria, even voids near the survey edge. On the other hand, *central* voids do not touch any boundary particle (i.e. the most distant void member particle is closer than any



**Figure 1.** HEALPIX map in a Mollweide projection of identified boundary zones (black) around and within the SDSS DR9 survey area where we inject boundary particles.

boundary particle), and thus are not near any survey boundary or internal mask. The *central* sample is designed to ensure that we have a fair distribution of void shapes and alignments within the survey volume. To handle survey high-redshift caps, we simply exclude from all catalogues any void which extends beyond the redshift limits of a given sample. This is a more restrictive approach than the procedure discussed in Sutter et al. (2012a), since it is difficult to construct stable tessellations on cospherical points. To evaluate this, if the distance from the void barycentre to the redshift edge is closer than the distance to any void member particle, we reject the void.

### 3 VOIDS IN DATA

#### 3.1 Galaxy populations

We take our galaxy sample from the CMASS selection of SDSS DR9 (BOSS) spectroscopic targets (Ahn et al. 2012). This is the same sample of galaxies used in the analysis of Reid et al. (2012). The 455 281 galaxies in this selection of the survey extend from  $z = 0.43$  to  $0.7$ . As before in our catalogue of voids in SDSS DR7 (Sutter et al. 2012a), we take volume-limited samples to ensure statistical uniformity and constant shot noise and galaxy bias around our identified voids. Also, we require volume-limited samples in order to compare to our mocks based on Halo Occupation Distribution (HOD) models, which are constructed around volume-limited surveys (Berlind & Weinberg 2002).

We apply simple evolution and  $K$ -corrections of the form

$$z_e = -2(1 - z + 0.1)(z - 0.1) \quad (3)$$

$$z_k = -0.242659 + 1.38731z, \quad (4)$$

and compute absolute magnitudes  $M_r$  assuming the cosmological parameters noted above. We choose three redshift bins. Each redshift range is characterized by a typical galaxy luminosity, which we differentiate by the labels *dim*, *mid*, and *bright*. Our redshift bins are:  $0.43 < z < 0.5$ , which we label *CMASS Dim*,  $0.5 < z < 0.6$ , labelled as *CMASS Mid*, and  $0.6 < z < 0.7$ , called *CMASS Bright*. Table 1 lists the sample name, limiting absolute magnitude, redshift bound, comoving volume, number of galaxies, and the mean galaxy separation in that sample. The mean galaxy spacing is  $(n_g/V)^{-1/3}$ , where  $n_g$  is the number of galaxies within each sample and  $V$  is the sample volume.

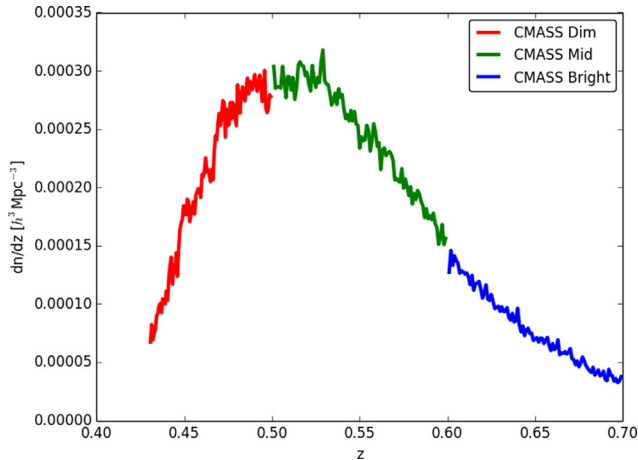
Fig. 2 shows the galaxy number density as a function of redshift for each of our samples. These plots show that a simple luminosity cut does not produce a truly volume-limited sample due to the

<sup>1</sup> <http://healpix.jpl.nasa.gov>



**Table 1.** Volume-limited galaxy samples used in this work.

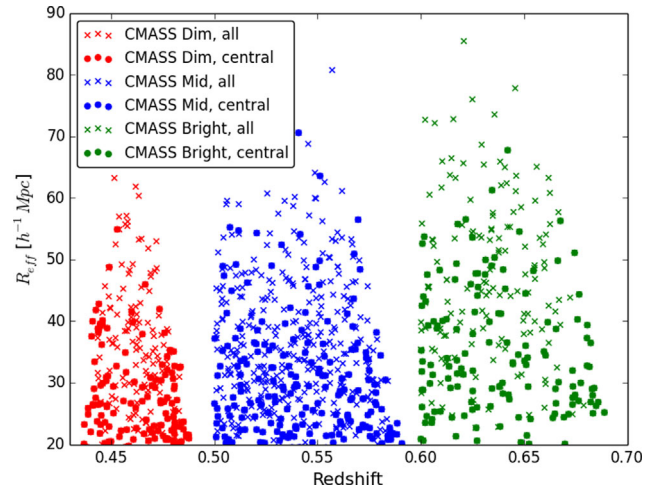
Sample name	$M_{r, \max}$	$z_{\min}$	$z_{\max}$	Volume ( $h^{-3} \text{ Gpc}^3$ )	$N_{\text{gal}}$	$\bar{n}^{-1/3}$ ( $h^{-1} \text{ Mpc}$ )
CMASS Dim	−20.1	0.43	0.5	0.29	61 249	16.76
CMASS Mid	−20.1	0.43	0.6	0.82	188 300	16.29
CMASS Bright	−20.8	0.43	0.7	1.48	205 732	19.29

**Figure 2.** Density of galaxies as a function of redshift in our three nearly volume-limited samples.**Table 2.** Summary of voids in data.

Sample name	$z_{\min}$	$z_{\max}$	Volume ( $h^{-3} \text{ Gpc}^3$ )	$N_{\text{voids}}$
CMASS Dim, all	0.43	0.5	0.29	283
CMASS Dim, central	0.43	0.5	0.29	151
CMASS Mid, all	0.5	0.6	0.53	570
CMASS Mid, central	0.5	0.6	0.53	242
CMASS Bright, all	0.6	0.7	0.66	283
CMASS Bright, central	0.6	0.7	0.66	137

complex target selection procedure in CMASS. The density in the *Dim* sample varies by a factor of  $\sim 4$ , in the *Mid* sample by a factor of  $\sim 2$ , and in the *Bright* sample by a factor of  $\sim 3$ . While we can account for redshift dependence in the densities by weighting individual galaxies in the `VIDE` code (Neyrinck 2008), our testing has shown that this does not strongly impact void properties, for two reasons. First, only the largest voids will span a deep enough redshift range to be sensitive to changes in the underlying density, and since the number of large voids are exponentially suppressed (and they are more likely to be removed from consideration because they intersect an internal boundary), varying galaxy number density will only affect a small percentage of voids. Secondly, the nature of the watershed algorithm naturally guards against effects of varying density: since each particle has on average 17 adjacent particles (Neyrinck 2008), we must reduce the local density of a wall by a large factor before it is no longer identified as a void boundary.

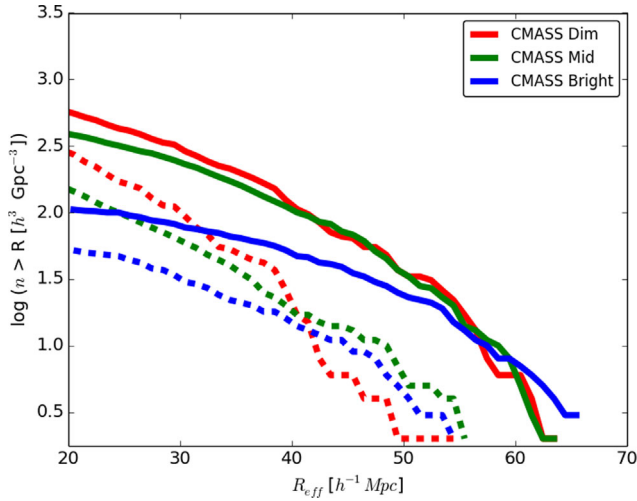
Table 2 summarizes the data samples used in this work, the redshift ranges used to produce the void samples, their respective volumes, and the total number of voids identified in each sample. In total, we identify nearly 1000 voids in the *all* sample, while the *central* sample produced  $\sim 480$  voids.

**Figure 3.** Distribution of voids within the CMASS samples. We plot void effective radius  $R_{\text{eff}}$  (equation 1) versus redshift. We show the voids in the volume-limited samples *CMASS Dim* (red), *CMASS Mid* (blue), and *CMASS Bright* (green). All voids are marked with a cross, and voids in the *central* catalogue are marked by filled circles.

### 3.2 Void properties

Fig. 3 shows the distribution of void sizes as a function of redshift for all the galaxy samples. We show both *all* and *central* voids. We see that though a few voids in *CMASS Mid* and *CMASS Bright* reach an effective radius of  $\sim 80 h^{-1} \text{ Mpc}$ , most – especially in the *central* catalogue – are below  $50 h^{-1} \text{ Mpc}$ . For *Mid* and *Bright* samples, voids pervade the low-redshift boundary, because here we keep the population of galaxies below that boundary and only reject voids whose centres fall below the redshift cutoff. On the other hand, we see a tapering in the distribution at the high-redshift caps: here, we reject any void that might intersect the cap, and we are more likely to cut progressively larger voids. We remove even more of the largest voids when creating the *central* samples, since these voids are more likely to lie nearby the mask line-of-sight boundaries. We observe a distinct lack of small voids in the *Bright* sample, which is a consequence of its slightly lower mean galaxy spacing. While we expect to see smaller voids at higher redshift, the effects of sparsity and biasing lead to larger observed voids (D’Aloisio & Furlanetto 2007).

We show another way of expressing the size distribution in Fig. 4. This is a plot of the cumulative number function: the total number of voids in each sample above a given effective radius. The number function is a potentially powerful probe of cosmology with voids (Sheth & van de Weygaert 2004). All samples have roughly the same number density of the largest voids, but the lack of smaller voids in the *CMASS Bright* sample manifests as a uniformly reduced number function for small- and medium-scale voids. In all samples, the *central* catalogues host roughly half as many voids as the *all* catalogues. This is a consequence of the relatively large surface-to-volume ratio of the current CMASS survey. The relatively narrow



**Figure 4.** Cumulative void number function for *CMASS Dim* (red), *CMASS Mid* (blue), and *CMASS Bright* (green). The number functions of all voids, including those near survey boundaries and internal masks, are shown as solid lines, while *central* catalogue voids are shown as dotted lines.

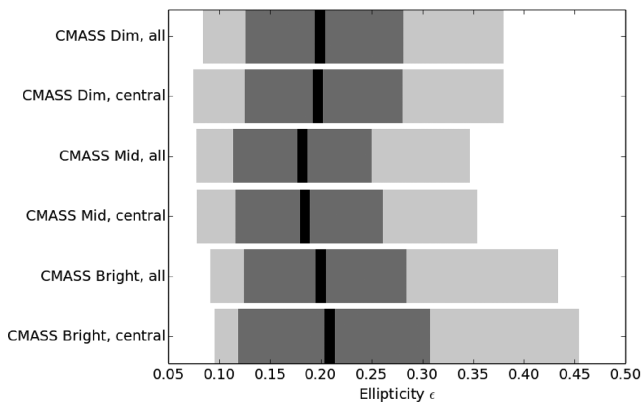
angular extent especially impacts our void populations; the small internal holes do little disruption. As future data releases fill in the central regions of the expected coverage area, we should approach the higher fraction of *central* voids seen in earlier surveys (Sutter et al. 2012a).

Though voids have complex shapes, we can assign them a unique ellipticity. This simple scalar captures most of the shape information of the void, and its distribution is a sensitive cosmological probe (Park & Lee 2007; Biswas et al. 2010; Lavaux & Wandelt 2010; Bos et al. 2012). To compute the ellipticity, for a given set of galaxies within a void, we first construct the inertia tensor

$$M_{xx} = \sum_{i=1}^{N_p} (y_i^2 + z_i^2)$$

$$M_{xy} = -\sum_{i=1}^{N_p} x_i y_i, \quad (5)$$

where  $N_p$  is the number of galaxies in the void, and  $x_i$ ,  $y_i$ , and  $z_i$  are coordinates of the particle  $i$  relative to the void barycentre. The other components of the tensor are obtained by cyclic permutations.



Given the inertia tensor, we compute the eigenvalues and form the ellipticity

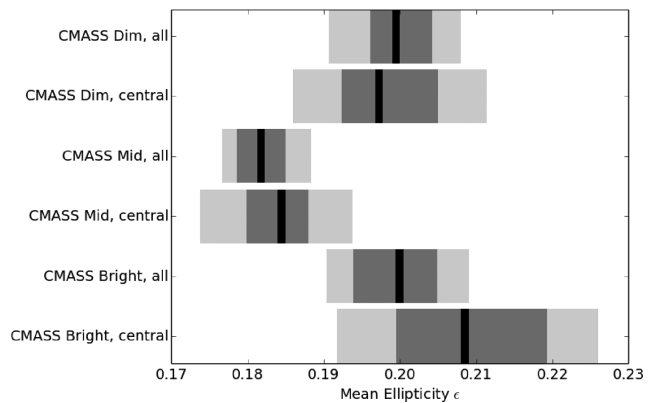
$$\epsilon = 1 - \left( \frac{J_1}{J_3} \right)^{1/4}, \quad (6)$$

where  $J_1$  and  $J_3$  are the smallest and largest eigenvalues, respectively. Note that this definition differs from that of Bos et al. (2012).

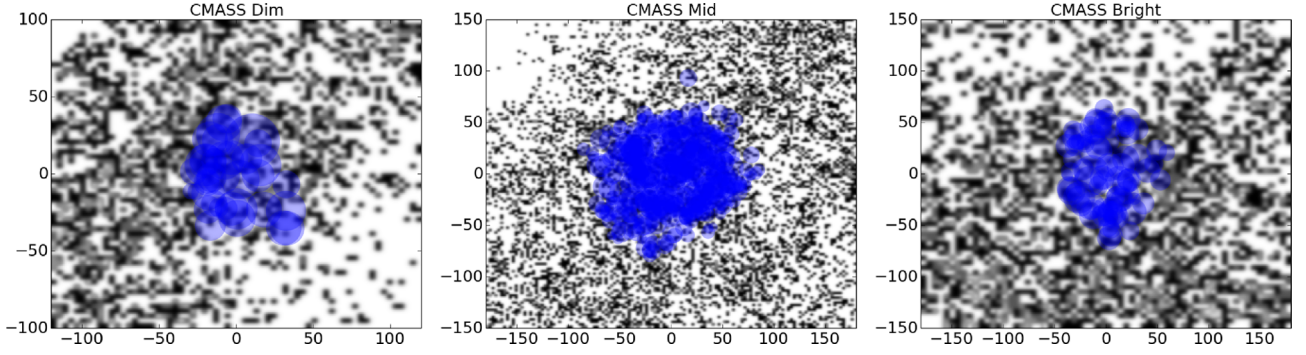
Fig. 5 shows the distribution of ellipticities for each of our samples. The ellipticities for all the samples are remarkably consistent, with means  $\sim 0.2$  and a slight skew in the distribution favouring slightly more elliptical voids. The *central* catalogue of the *CMASS Bright* voids contains a few highly elliptical voids. While similar voids exist in all samples, the limited number of total voids in this sample leads to a highly lopsided distribution. In the same figure, we also show the mean ellipticity and the standard error on the mean (i.e.  $\sigma/N_v$ , where  $\sigma$  is the standard deviation and  $N_v$  is the number of voids) for each sample. The means broadly agree, with the *CMASS Mid* sample favouring slightly more spherical voids. The ellipticities in the *central* catalogues are different than those in the *all* catalogues. We will see below that the mean ellipticity is sensitive to the distribution of void sizes in a particular sample. Since the mask affects the void size distribution in a non-trivial way for each sample, depending on the relative surface-to-volume ratio in the sample, we should not be surprised to measure slightly different mean void ellipticities. Fortunately, this does not appear to be a large effect from sample to sample, and we conclude that we only need to understand mask effects for a particular survey geometry, not for individual volume-limited samples within that survey.

In Fig. 6, we give a visual impression of some of the identified voids. We chose the particular slices randomly but selected a representative sample from the range of scales in the void catalogue. We represent the selected void as collections of overlapping circles, where each circle is a void member galaxy with radius equal to the effective radius of each Voronoi cell. We overplot these circles on slices from the galaxy distribution. We see that voids at all scales indeed sit within underdensities in the galaxy distribution, though in some samples the sparsity makes it difficult to clearly distinguish the surrounding walls and filaments. However, our analysis in Sutter et al. (2014a) and Hamaus et al. (2014b) show that these objects share common features with voids in high-density surveys.

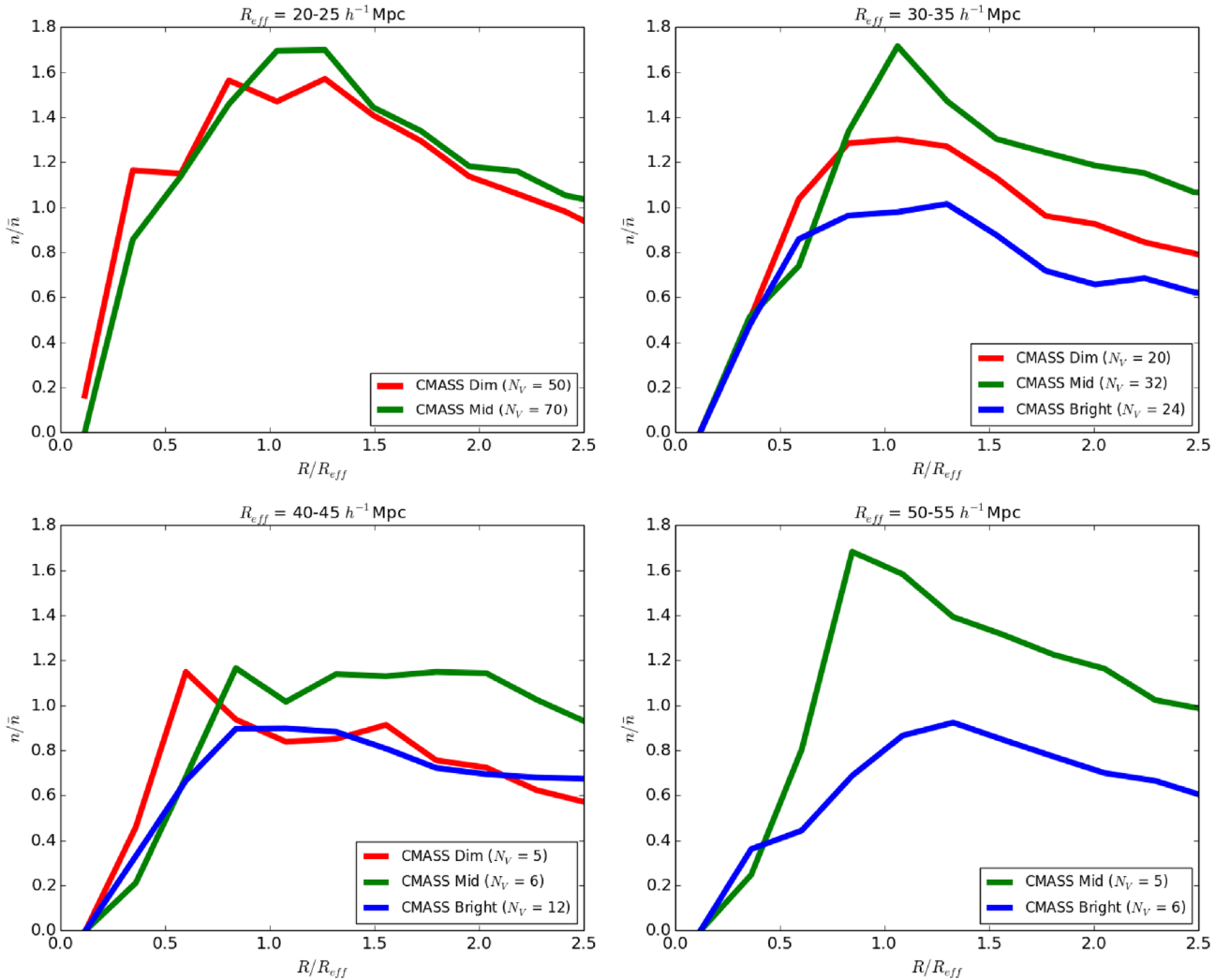
Fig. 7 shows one-dimensional radial density profiles of stacked voids in each sample. To compute the profiles, we take all voids in a sample of a given size range (e.g. 20–25  $h^{-1}$  Mpc), align all their



**Figure 5.** Ellipticity distributions  $\epsilon$  (equation 6) with 68 per cent (dark grey) and 95 per cent (light grey) ranges for each sample (left), and mean ellipticities with  $1\sigma$  and  $2\sigma$  uncertainties on the mean (right). For the mean ellipticities (right), error bars are calculated with a bootstrap method.



**Figure 6.** Void and galaxy density slices. We select three locations in each sample centred on a representative void. We show voids from *CMASS Dim* (left), *CMASS Mid* (middle), and *CMASS Bright* (right). We represent the void member galaxies as small circles with radii equal to the effective radii of their corresponding Voronoi cells. The width of each galaxy slice along the  $z$ -axis is the entire sample for *CMASS Dim* and  $300 h^{-1}$  Mpc for *CMASS Mid* and *CMASS Bright*. To avoid unnecessary overlap from projection, we take void particles from a thinner slice:  $10 h^{-1}$  Mpc for *CMASS Dim*,  $75 h^{-1}$  Mpc for *CMASS Mid*, and  $150 h^{-1}$  Mpc for *CMASS Bright*. The galaxies are binned into pixels with the number of pixels varied to best highlight the surrounding structure: 64 bins for *CMASS Dim*, 128 for *CMASS Mid*, and 64 for *CMASS Bright*. The projected galaxy density is coloured from 0.0 (white) to 1.5 (black) and is shown on a logarithmic scale. Axes are marked in units of  $h^{-1}$  Mpc.



**Figure 7.** One-dimensional radial density profiles of stacked voids. Each profile is normalized to the mean number density  $\bar{n}$  of that sample and  $R_{\text{eff}}$  corresponds to the median void size in the stack. We only show profiles from stacked *central* voids. Void profiles do not necessarily reach the mean density because of the influence of boundary particles and empty regions outside the survey volume. The caption lists the number of voids stacked in each profile.

barycentres, and measure the density in thin spherical shells. We normalize each density profile to the mean number density of the sample and show all profiles as a function of relative radius,  $R/R_{\text{eff}}$ . We only show profiles from the *central* voids of the samples. We show four stacks: 20–25, 30–35, 40–45, and 50–55  $h^{-1}$  Mpc.

The smallest stack, 20–25  $h^{-1}$  Mpc, is very close to the mean galaxy separation for the samples, and this manifests in an extremely steep profile. While these small voids may be unreliable due to Poisson shot noise, we do not assess statistical significance in this work. We believe that the common technique used with watershed algorithms (comparing to voids in a Poisson distribution of equal number density) is inadequate because small voids tend to appear in higher density regions, and thus are more likely to be real voids than one would expect. We are currently developing a more robust significance criteria based on Bayesian analysis of constrained realizations of a given sample.

As we progress to larger stacks, the profiles become more shallow and the overdense region surrounding the voids becomes less pronounced, as also seen in Ceccarelli et al. (2013). However, as before in the voids of SDSS DR7 (Sutter et al. 2012a), we see a qualitatively universal profile across all void sizes and redshift ranges: an underdense centre, a steep power-law slope at the wall of the void, a slightly overdense ‘compensation’, and a steady declining to the mean density. We see here that there is a larger difference among the samples than in the DR7 voids. A significant reason for this is the smaller survey area: while we guarantee *central* voids to sit completely within the survey, these profiles extend beyond the void effective radius. Since the survey area is relative small, the profile quickly reaches into volumes beyond the survey mask. There are also boundary particles at the edges. Depending on the relative density of the boundary particles and the location of stacked voids relative to the boundary, we see different profiles at larger radii. Within  $R_{\text{eff}}$ , we see strong consistency among the samples, as expected. The number of voids in the stack strongly affects the smoothness of the profile. In particular, the 50–55  $h^{-1}$  Mpc stack in *CMASS Mid* contains only five voids, which leads to a highly irregular profile shape.

## 4 VOIDS IN MOCKS

### 4.1 Mock galaxy populations

We take several avenues for comparison to the voids in the survey data. For the first set of mocks, we compute a single  $\Lambda$ CDM dark matter  $N$ -body simulation, extract haloes from the simulation, and use the haloes positions and masses as inputs for an HOD model. For the simulation we use the `2HOT` code, an adaptive treecode  $N$ -body method whose operation count scales as  $N \log N$  in the number of particles (Warren 2013). Accuracy and error behaviour have been improved significantly for cosmological volumes through the use of a technique to subtract the uniform background density, as well as using a compensating smoothing kernel for small-scale force softening (Dehnen 2001). We use a standard symplectic integrator (Quinn et al. 1997) and an efficient implementation of periodic boundary conditions using a high-order ( $p = 8$ ) multipole local expansion. We adjust the error tolerance parameter to limit absolute errors to 0.1 per cent of the rms peculiar acceleration. As an example, a complete 4096<sup>3</sup> particle simulation requires about 120 wall-clock hours using 12 000 CPU cores. Initial conditions were generated using a power spectrum calculated with `CLASS` (Blas, Lesgourgues & Tram 2011) and realized with a modified version of `2LPTIC` (Crocce, Pueblas & Scoccimarro 2006).

This particular simulation assumed Planck first-year cosmological parameters (Planck Collaboration 2013b). The box size was  $4 h^{-1}$  Gpc on a side and contained 4096<sup>3</sup> particles, giving a particle mass resolution of  $7.36 \times 10^{10} h^{-1} M_{\odot}$ . All analysis in this work used a single snapshot at  $z = 0.53$ . We identified haloes in the simulation volume using the `ROCKSTAR` halo finder (Behroozi, Wechsler & Wu 2013), a six-dimensional phase-space plus time halo finder, to identify spherical overdensity (SO) haloes at 200 times the background density. We use the default `ROCKSTAR` parameters, except for requiring strict SO masses which includes unbound particles and particles which may exist outside of the FoF group for the halo.

We produce galaxy catalogues from the halo population using the code described in Tinker, Weinberg & Zheng (2006) and the HOD model described in Zheng, Coil & Zehavi (2007). HOD modelling assigns central and satellite galaxies to a dark matter halo of mass  $M$  according to a parametrized distribution. In the case of the Zheng et al. (2007) parametrization, the mean number of central galaxies is given by

$$\langle N_{\text{cen}}(M) \rangle = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log M - \log M_{\min}}{\sigma_{\log M}} \right) \right] \quad (7)$$

and the mean number of satellites is given by

$$\langle N_{\text{sat}}(M) \rangle = \langle N_{\text{cen}}(M) \rangle \left( \frac{M - M_0}{M'_1} \right)^{\alpha}, \quad (8)$$

where  $M_{\min}$ ,  $\sigma_{\log M}$ ,  $M_0$ ,  $M'_1$ , and  $\alpha$  are free parameters that must be fitted to a given survey. The probability distribution of central galaxies  $P(N_{\text{cen}} | \langle N_{\text{cen}} \rangle)$  is a nearest-integer distribution, and satellites follow a Poisson  $P(N_{\text{sat}} | \langle N_{\text{sat}} \rangle)$ . Central galaxies are given the peculiar velocities of the host halo, and satellite galaxies are given an additional random velocity drawn from a Maxwellian distribution with the halo velocity dispersion. Using the HOD parameters found in Manera et al. (2013,  $\sigma_{\log M} = 0.596$ ,  $M_0 = 1.2 \times 10^{13} h^{-1} M_{\odot}$ ,  $M'_1 = 10^{14} h^{-1} M_{\odot}$ ,  $\alpha = 1.0127$ , and  $M_{\min}$  chosen to fit the mean number density of our sample), we generate the mock galaxy population. Although the HOD fitting of Manera et al. (2013) assumed a slightly different cosmology, we found that this did not affect our results. To simplify comparison, we will only target the *CMASS Mid* sample.

We note that we only fit the mean density of the galaxy sample; we make no attempt to model the variation of the number density as a function of redshift. However, even with this restriction, we will see below that we find excellent agreement between the mocks and data for all void statistics. While ignoring the density variation may not be adequate for precise modelling, at the current level of statistical uncertainty we only wish to make an initial comparison, and save a more detailed treatment for future work.

For the full-volume simulation analysis, we use the entire three-dimensional volume of the simulation box and perturb each galaxy according to its peculiar velocity. We call this full-volume set of mock galaxies the *N-body Mock* sample. Even though we analyse only a single realization, the large volume produces over 90 000 voids.

To provide a more direct comparison to the data and to understand the effects of the mask on void properties, we apply the same survey geometry to mock galaxies as is used in the SDSS DR9 samples. Instead of placing galaxies in redshift space along the  $z$ -axis of the simulation box as we do above, we place an observer at the centre of the volume and measure each galaxy position as its radial distance from that observer. We also perturb the galaxies according to the peculiar velocities in all directions. We project all galaxies on to the sky and apply the mask in Fig. 1. Since we wish to compare



**Table 3.** Summary of voids in mocks.

Sample name	Volume ( $h^{-3} \text{ Gpc}^3$ )	# Mocks	$N_{\text{voids}}$
<i>N</i> -body Mock	64.00	1	91 711
Masked <i>N</i> -body Mock, all	0.53	5	3336
Masked <i>N</i> -body Mock, central	0.53	5	1313
Masked 2LPT Mock, all	0.53	98	56 096
Masked 2LPT Mock, central	0.53	98	25 299

solely to the *CMASS Mid* sample, we only take mock galaxies within  $0.43 < z < 0.5$ . For void finding, we treat this sample exactly as data: we deploy boundary particles and we build *all* and *central* catalogues of voids. While it is very expensive to compute multiple realizations with this volume and resolution, we can take advantage of the relative narrowness of the current DR9 survey by rotating our mask orientation within the same simulation volume. We take five separate orientations. Even though these samples are not quite independent, and we are likely to not capture enough of the very largest voids due to finite-volume effects, this technique still provides a good proxy for multiple realizations and allows us to gauge the range of void statistics predicted from simulations. We call these samples *Masked N-body Mock*. Taken together, we find  $\sim 3\,300$  voids in the *all* samples of all orientations, and  $\sim 1\,300$  voids in the *central* samples.

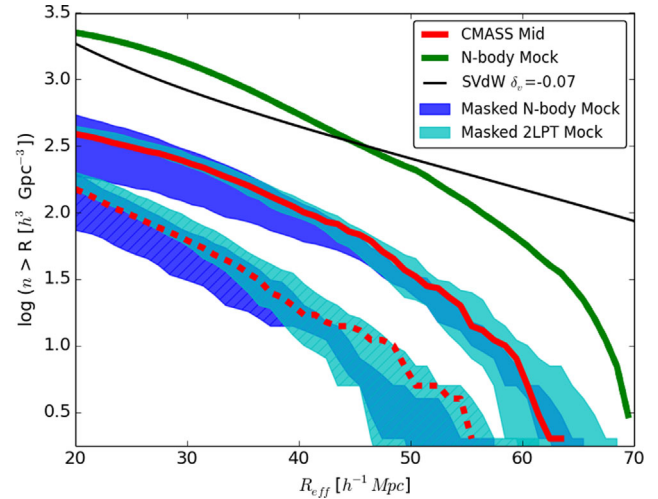
Since high-resolution *N*-body simulations are very expensive, and the galaxy correlation analysis of Reid et al. (2012) required many mocks for estimating shot noise, the BOSS team produced many mock catalogues based on second-order Lagrangian perturbation theory (2LPT) with a *WMAP* 7-year cosmology and the same HOD prescription described above (Manera et al. 2013). We used 98 of these publicly available mock catalogues to validate our *N*-body simulation results and to further estimate the uncertainties in the predicted void statistics. We denote these mock samples as *Masked 2LPT Mocks* and process them identically to real data. With the 98 mocks we find over 56 000 voids in the *all* sample and over 25 000 in the *central* sample. However, these mocks were only made available with the survey mask already in place; thus we will only use our *N*-body Mocks for interpreting the effects of the mask.

Finally, to evaluate the significance of our ellipticity measurements, we identify voids in a single random realization provided by the *CMASS* team. In this realization, galaxies are randomly distributed with Poissonian statistics within the survey volume.

Table 3 summarizes the mock samples used in this work, their respective volumes, the number of independent mock samples, and the total number of voids identified in each sample.

## 4.2 Comparison to data

Our first point of comparison is the cumulative number function. In Fig. 8, we compare the number function of voids in the *CMASS Mid* data sample to all our mocks. First, the unmasked *N*-body Mock simulation hosts roughly three times as many voids per unit volume than the data, even though they have similar galaxy populations. This occurs at all scales, though there are  $\sim 4$  times as many small voids in the unmasked mock as in the data. We can understand this disruption by considering the effects of the mask on a particular void: it will slice the void, making it appear as a smaller void. So after the mask is applied large voids become medium voids, medium voids become small voids, and so on. Since the number function falls steeply with void size, if we consider a given void size range,



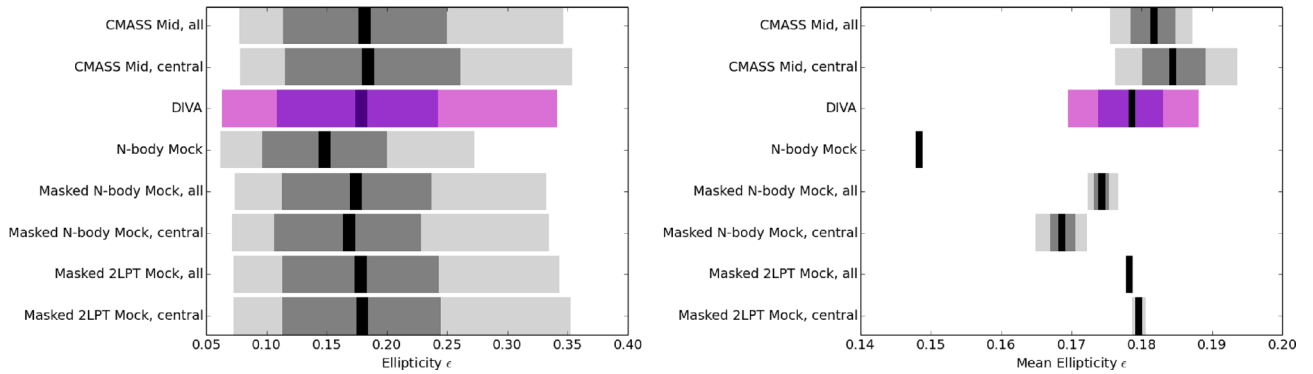
**Figure 8.** Comparison of cumulative void number function between mocks and data. The solid (dotted) line corresponds to the *all* (*central*) void population of the *CMASS Mid* galaxy sample. The green solid line shows the number function of the full-volume *N*-body Mock sample. The light blue shaded region shows the full range of number functions from the 98 *Masked 2LPT Mock* runs, while the dark blue shaded region shows the full range of the five *Masked N-body Mock* runs. For the shaded regions, solid indicates *all* voids and hatched indicates *central* voids. The solid black line is the theoretical expectation from the Sheth & van de Weygaert (2004) number function with  $\delta_v = -0.07$ .

then there are far fewer voids being added to that range (by being sliced and becoming smaller) than voids that are lost to smaller ranges. This leads to a systematic reduction in the number density of voids. For larger survey areas and more complete coverage, we expect less drastic impacts.

In the same figure, we plot the theoretical number function of Sheth & van de Weygaert (2004), which was derived from an excursion set formalism. As found in Sutter et al., (2014a), the best match to voids in low-density galaxy surveys comes from adjusting the ‘void parameter’  $\delta_v$  to  $-0.07$ . While the number function roughly agrees with the order of magnitude of the full *N*-body Mock void population, it does not fall off as steeply as in the mocks, though this might be influenced by finite-volume effects. Still, the correspondence of these curves shows that theoretical modelling can qualitatively match unmasked void populations, but further adjustments must be made to match void statistics from masked volumes.

When we apply the DR9 survey geometry to our mock catalogues, we see excellent agreement in the number functions for both *all* and *central* catalogues. The 2LPT mocks show much less variance at small void sizes and reach larger void sizes than the *N*-body mocks, since there are many more 2LPT simulations and they are each drawn from independent realizations. The different orientations of the *N*-body mocks are limited by the cosmic variance of a single simulation, so it is not surprising that we are not able to match the void number function at the largest void sizes. However, with the 2LPT mocks, we are able to capture the very largest voids. Tavasoli et al. (2013) first pointed out a potential tension between  $\Lambda$ CDM model and data, but we see no indication of this tension in our results.

Fig. 9 shows the ellipticity distributions and mean ellipticities, identically to Fig. 5. In this plot, we compare *CMASS Mid* ellipticities to those found in mock samples. Voids in all mock and data samples tend to be more elliptical than voids drawn from



**Figure 9.** Comparing the ellipticity distributions (left) and mean ellipticities with their uncertainties (right) of the *CMASS Mid* sample to the various mocks. Error bars are the same as indicated for Fig. 5. The violet-coloured distribution is the theoretical expectation from DIVA (Lavaux & Wandelt 2010) with the rescaling parameter  $\alpha$  set to 0.25. DIVA predicts the ellipticity distribution from the size distribution of voids in the *CMASS Mid central* sample.

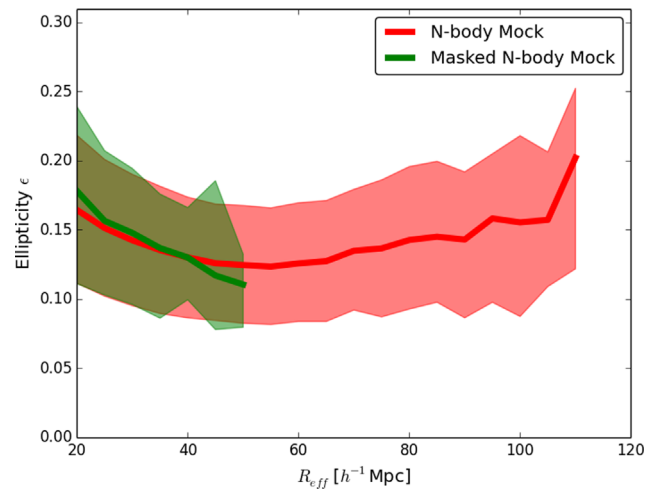
Poissonian distributions (which give  $\epsilon \sim 0.12$ ), confirming the analysis of Sutter et al. (2014b), which showed that ellipticity is a robust measurement even in low-density samples such as SDSS DR9.

As we saw earlier for the cumulative number functions, we find agreement between data and theoretical expectations in the form of mock catalogues *only* when we apply the mask: then the distributions are nearly identical. The *all* void samples agree among all the masked mocks and data. The mean ellipticities of the *central Masked N-body Mock* are slightly more than  $2\sigma$  discrepant than the data, while the *Masked 2LPT Mocks* have much better agreement with the data. The differences between the *N-body* and *2LPT* mock populations are not surprising since they have slightly different cosmologies, and the ellipticity is a very sensitive probe of cosmological parameters (Biswas et al. 2010; Bos et al. 2012). Also, the *2LPT Mocks* cover multiple realizations, whereas the *N-body* mocks are restricted to a single simulation. With the exception of the *Masked N-body Mock*, we see the same relationship between *all* and *central* ellipticities as in the data: voids in *central* samples less spherical. Indeed, the 90 000 voids in all the *2LPT* mocks reduce the uncertainties to such a degree that this difference is easily distinguishable.

Voids in the data and masked mocks are more elliptical than in the full-volume simulation. We can understand the impact of the mask by examining the relationship between ellipticity and void size, as we show for the masked and unmasked *N-body* mocks in Fig. 10. The cause of the shift in ellipticities when masking data is readily apparent: larger voids tend to be more spherical, and their exclusion from the masked catalogues increases the overall mean ellipticity. The uncertainties in the *N-body* mocks are too large to distinguish any differences.

We derive the theoretical ellipticity distribution from DIVA (Lavaux & Wandelt 2010) with the rescaling parameter  $\alpha = 0.25$ , as discussed in the analysis of Sutter et al. (2014a). DIVA requires a void size distribution as input, and we take the actual distribution from the *CMASS Mid central* sample (i.e. Fig. 4). This choice of rescaling parameter  $\alpha$  provided good agreements with mock void populations, and again here we see agreement with data, indicating a relationship between the sizes of voids identified with the watershed transform and their dynamical cores.

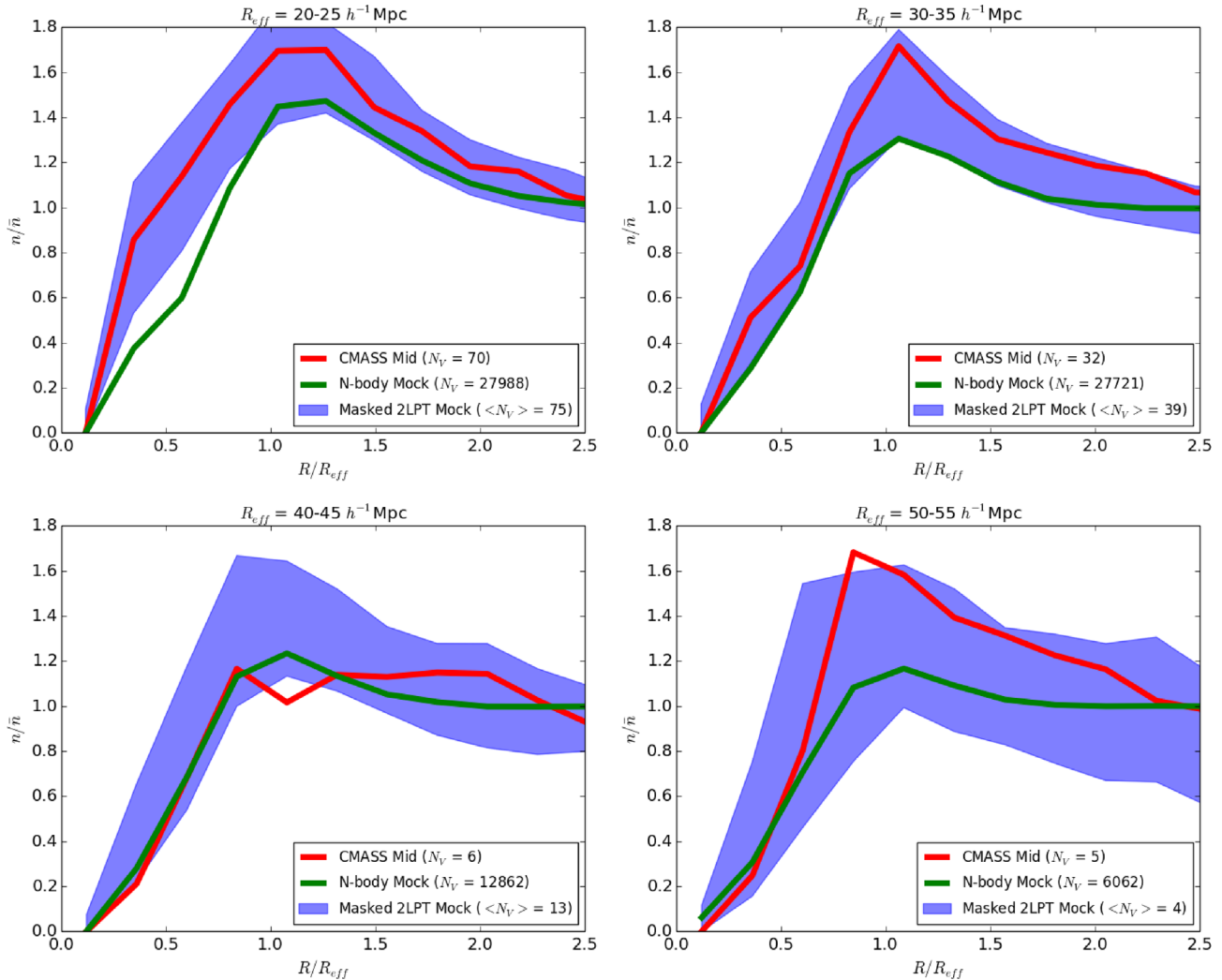
Interestingly, we see almost no significant differences between masked and unmasked radial density profiles, as we show in Fig. 11. Here, we compare the unmasked *N-body Mock* and the *Masked 2LPT Mocks* to the *CMASS Mid* data sample. We do not plot stacks



**Figure 10.** Distribution of ellipticities  $\epsilon$  in small bins of effective radius  $R_{\text{eff}}$  in the masked (green) and unmasked (red) *N-body Mock* simulations. The solid line shows the mean, and the shaded region is the  $1\sigma$  interval.

from the *Masked N-body Mock* since they simply overlap the existing curves but have larger scatter. The stacks shown here are the same as before in Fig. 7. With very few exceptions, profiles from both the data and the full-volume mock sit within the range of profiles constructed in the different realizations of the masked mock. This agreement persists at all void scales. The two visible excursions from the mocks are consistent with statistical expectations. Since radial density profiles by definition probe the interiors and immediate surroundings of voids, they naturally are more robust against distortions due to the mask. Also, the process of stacking turns sets of elliptical voids with different orientations into a roughly spherical shape, regardless of the distribution of the individual ellipticities (Lavaux & Wandelt 2012; Pisani et al. 2013). Thus any changes to the mean ellipticity due to the mask will not change the radial profiles significantly.

With the enormous number of voids in the *N-body Mock* sample, we begin to see the transition from over to undercompensated voids (Hamaus et al. 2014a). However, this transition is not clear in the data due to the lack of voids at these extreme scales.



**Figure 11.** Comparison of one-dimensional radial density profiles. The void stacks and normalizations are the same as in Fig. 7. The solid red lines are profiles from *CMASS Mid*, the solid green lines are from the full-volume unmasked *N-body Mock*, while the shaded blue region shows the full range of profiles from the 98 realizations of the *Masked 2LPT Mock* simulations. The caption lists the number of voids stacked in each profile, except for *Masked 2LPT Mock*, where the caption shows the mean number of voids in each realization for that stack. The number of excursions of the data from the mocks is consistent with expectations.

## 5 CONCLUSIONS

We have constructed a void catalogue from the SDSS DR9 CMASS spectroscopic galaxy survey. Combined with the voids from DR7, this is the largest void catalogue to date. We have used previously established methods for removing voids near survey boundaries to produce a catalogue of voids with a fair distribution of shapes. This catalogue also provides the most distant voids ever detected, extending our knowledge of galaxy underdensities to redshift  $z = 0.7$ . Our voids have effective radii between 20 and 70  $h^{-1}$  Mpc, and exhibit the same qualitatively universal radial density profile seen in earlier works (Sutter et al. 2012a; Ceccarelli et al. 2013). We used HOD modelling to produce mock catalogues for comparison purposes.

We find that the effects of the mask are highly non-trivial and can depend strongly on the relative surface area of the mask compared to the volume, the depth of the survey, the number of internal holes, and the detailed shape of the boundary. Cosmological statistics based on global void properties, such as number functions and ellipticity distributions, are especially vulnerable to the properties of the mask. However, we find that probes based on void

interiors, such as radial density or velocity profiles, are generally more robust. Also, cosmological applications which depend on the statistical isotropy of voids, such as the Alcock–Paczynski test (Sutter et al. 2012b), are resilient. In cases such as this, while the mask may change the average void shape or size, as long as the mask preserves a uniform sampling of their orientations then the methods are sound.

Furlanetto & Piran (2006) and Tinker & Conroy (2009) hypothesized that the so-called void phenomenon of Peebles (2001) can be explained by galaxy bias. This has led to a general discussion in the literature about potential discrepancies between the voids in theory and voids in data. For the void definition we adopt, abundances, ellipticity distributions, and radial profiles all indicate that voids in simulations have the same sizes, shapes, and interior contents as observed voids once galaxy bias, sparsity, and survey masks are accounted for. Overall, we find no significant discrepancies between observations of voids in the SDSS DR9 and  $\Lambda$ CDM mocks.

We have made all the SDSS DR9 voids, as well as all mock catalogues used in this work, publicly available online at the Public Cosmic Void Catalog at <http://www.cosmicvoids.net>.

## ACKNOWLEDGEMENTS

PMS and BDW acknowledge support from NSF Grant NSF AST 09-08693 ARRA. BDW acknowledges funding from an ANR Chaire d'Excellence (ANR-10-CEXC-004-01), the UPMC Chaire Internationale in Theoretical Cosmology, and NSF grants AST-0908 902 and AST-0708849. GL acknowledges support from CITA National Fellowship and financial support from the Government of Canada Post-Doctoral Research Fellowship. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation. DW acknowledges support from NSF Grant AST-1009505. This work made in the ILP LABEX (under reference ANR-10-LABX-63) was supported by French state funds managed by the ANR within the Investissements d'Avenir programme under reference ANR-11-IDEX-0004-02. We would like to thank Nico Hamaus for useful discussions.

## REFERENCES

- Ahn C. P. et al., 2012, *ApJS*, 203, 21
- Alcock C., Paczynski B., 1979, *Nature*, 281, 358
- Bassett B., Hlozek R., 2010, *Dark Energy: Observational and Theoretical Approaches*. Cambridge Univ. Press, Cambridge, p. 246
- Baugh C. M., Benson A. J., Cole S., Frenk C. S., Lacey C., 2003, in Bender R., Renzini A., eds, *Proc. Eur. South. Obs. Univ., The Mass of Galaxies at Low and High Redshift*. Springer-Verlag, Berlin, p. 91
- Beck A. M., Hanaasz M., Lesch H., Remus R.-S., Stasyszyn F. A., 2013, *MNRAS*, 429, L60
- Behroozi P. S., Wechsler R. H., Wu H., 2013, *ApJ*, 762, 109
- Benson A. J., Hoyle F., Torres F., Vogeley M. S., 2003, *MNRAS*, 340, 160
- Berlind A. A., Weinberg D. H., 2002, *ApJ*, 575, 587
- Biswas R., Alizadeh E., Wandelt B., 2010, *Phys. Rev. D*, 82
- Blas D., Lesgourgues J., Tram T., 2011 *J. Cosmol. Astropart. Phys.*, 2011, 034
- Bos E. G. P., van de Weygaert R., Dolag K., Pettorino V., 2012, *MNRAS*, 426, 440
- Ceccarelli L., Padilla N. D., Valotto C., Lambas D. G., 2006, *MNRAS*, 373, 1440
- Ceccarelli L., Herrera-Camus R., Lambas D. G., Galaz G., Padilla N. D., 2012, *MNRAS*, 426, L6
- Ceccarelli L., Paz D., Lares M., Padilla N., García Lambas D., 2013, *MNRAS*, 434, 1435
- Clampitt J., Cai Y.-C., Li B., 2013, *MNRAS*, 431, 749
- Crocce M., Pueblas S., Scoccimarro R., 2006, *MNRAS*, 373, 369
- D'Aloisio A., Furlanetto S. R., 2007, *MNRAS*, 382, 860
- Dawson K. S. et al., 2013, *AJ*, 145, 10
- Dehnen W., 2001, *MNRAS*, 324, 273
- Einasto J., Einasto M., Gramann M., Saar E., 1991, *MNRAS*, 248, 593
- Ferreras I., Pasquali A., eds, 2011, *Astrophysics and Space Science Proc. Environment and the Formation of Galaxies: 30 Years Later*. Springer-Verlag, Berlin
- Furlanetto S. R., Piran T., 2006, *MNRAS*, 366, 467
- Goldberg D. M., Vogeley M. S., 2004, *ApJ*, 605, 1
- Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, *ApJ*, 622, 759
- Gottlober S., Lokas E. L., Klypin A., Hoffman Y., 2003, *MNRAS*, 344, 715
- Hamaus N., Wandelt B. D., Sutter P. M., Lavaux G., Warren M. S., 2014a, *Phys. Rev. Lett.*, 112, 041304
- Hamaus N., Sutter P. M., Wandelt B. D., 2014b, preprint ([arXiv:e-prints](#))
- Higuchi Y., Oguri M., Hamana T., 2013, *MNRAS*, 432, 1021
- Hoyle F., Vogeley M. S., 2004, *ApJ*, 607, 751
- Hoyle F., Rojas R. R., Vogeley M. S., Brinkmann J., 2005, *ApJ*, 620, 618
- Hoyle F., Vogeley M. S., Pan D., 2012, *MNRAS*, 426, 3041
- Ilic S., Langer M., Douspis M., 2013, *A&A*, 556, A51
- Jennings E., Li Y., Hu W., 2013, *MNRAS*, 434, 2167
- Komatsu E. et al., 2011, *ApJS*, 192, 18
- Krause E., Chang T.-C., Doré O., Umetsu K., 2013, *ApJ*, 762, L20
- Laureijs R. et al., 2011, *Euclid Definition Study Report*, preprint ([arXiv:1110.3193](#))
- Lavaux G., Wandelt B. D., 2010, *MNRAS*, 403, 1392
- Lavaux G., Wandelt B. D., 2012, *ApJ*, 754, 109
- Little B., Weinberg D. H., 1994, *MNRAS*, 267, 605
- Manera M. et al., 2013, *MNRAS*, 428, 1036
- Marín F. A. et al., 2013, *MNRAS*, 432, 2654
- Melchior P., Sutter P. M., Sheldon E., Krause E., Wandelt B. D., 2014, *MNRAS*, 440, 2922
- Muller V., Arbabi-Bidgoli S., Einasto J., Tucker D., 2000, *MNRAS*, 318, 280
- Neyrinck M. C., 2008, *MNRAS*, 386, 2101
- Padilla N. D., Ceccarelli L., Lambas D. G., 2005, *MNRAS*, 363, 977
- Pan D. C., Vogeley M. S., Hoyle F., Choi Y.-Y., Park C., 2012, *MNRAS*, 421, 926
- Paranjape A., Lam T. Y., Sheth R. K., 2012, *MNRAS*, 420, 1648
- Park D., Lee J., 2007, *ApJ*, 665, 96
- Peebles P. J. E., 2001, *ApJ*, 557, 495
- Pisani A., Lavaux G., Sutter P. M., Lavaux G., Wandelt B. D., 2013, *MNRAS*, submitted
- Planck Collaboration 2013a, preprint ([arXiv:1303.5079](#))
- Planck Collaboration 2013b, preprint ([arXiv:1303.5076](#))
- Platen E., van de Weygaert R., Jones B. J. T., 2007, *MNRAS*, 380, 551
- Quinn T., Katz N., Stadel J., Lake G., 1997, preprint ([astro-ph/9710043](#))
- Reid B. A. et al., 2012, *MNRAS*, 426, 2719
- Rojas R. R., Vogeley M. S., Hoyle F., Brinkmann J., 2004, *ApJ*, 617, 50
- Rojas R. R., Vogeley M. S., Hoyle F., Brinkmann J., 2005, *ApJ*, 624, 571
- Ryden B. S., 1995, *ApJ*, 452, 25
- Ryden B. S., Melott A. L., 1996, *ApJ*, 470, 160
- Sánchez A. G. et al., 2012, *MNRAS*, 425, 415
- Schlegel D. et al., 2011, *The BigBOSS Experiment*, preprint ([arXiv:1106.1706](#))
- Sheth R. K., van de Weygaert R., 2004, *MNRAS*, 350, 517
- Shoji M., Lee J., 2012, preprint ([arXiv:1203.0869](#))
- Spergel D. et al., 2013, preprint ([arXiv:1305.5422](#))
- Spolyar D., Sahlén M., Silk J., 2013, *Phys. Rev. Lett.*, 111, 241103
- Strauss M. A. et al., 2002, *AJ*, 124, 1810
- Sutter P. M., Lavaux G., Wandelt B. D., Weinberg D. H., 2012a, *ApJ*, 761, 44
- Sutter P. M., Lavaux G., Wandelt B. D., Weinberg D. H., 2012b, *ApJ*, 761, 187
- Sutter P. M., Lavaux G., Hamaus N., Wandelt B. D., Weinberg D. H., Warren M. S., 2014a, *MNRAS*, 442, 462
- Sutter P. M., Lavaux G., Wandelt B. D., Weinberg D. H., Warren M. S., 2014b, *MNRAS*, 438, 3177
- Sutter P. M. et al., 2014c, preprint ([arXiv:1406.1191](#))
- Tavasoli S., Vasei K., Mohayaee R., 2013, *A&A*, 553, A15
- Taylor A. M., Vovk I., Neronov A., 2011, *A&A*, 529, A144
- Thompson L. A., Gregory S. A., 2011, preprint ([arXiv:1109.1268](#))
- Thompson K. L., Vishniac E. T., 1987, *ApJ*, 313, 517
- Tinker J. L., Conroy C., 2009, *ApJ*, 691, 633
- Tinker J. L., Weinberg D. H., Zheng Z., 2006, *MNRAS*, 368, 85
- Vogele M. S., Geller M. J., Park C., Huchra J. P., 1994, *AJ*, 108, 745
- von Benda-Beckmann A. M., Mueller V., 2007, preprint ([arXiv:0710.2783](#))
- Warren M. S., 2013, preprint ([arXiv:e-prints](#))
- Weinberg D. H., Cole S., 1992, *MNRAS*, 259, 652
- Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, *Phys. Rep.*, 530, 87
- Zheng Z., Coil A. L., Zehavi I., 2007, *ApJ*, 667, 760

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.