

This article has been accepted for publication in Monthly Notices of the Royal Astronomical Society ©: 2019 The Authors. Published by Oxford University Press on behalf of the Royal Astronomical Society. All rights reserved.

Cosmological exploitation of the size function of cosmic voids identified in the distribution of biased tracers

Sofia Contarini,^{1,2★} Tommaso Ronconi,^{3,4} Federico Marulli,^{1,2,5} Lauro Moscardini,^{1,2,5} Alfonso Veropalumbo⁶ and Marco Baldi^{1,2,5}

¹Dipartimento di Fisica e Astronomia – Alma Mater Studiorum Università di Bologna, via Piero Gobetti 93/2, I-40129 Bologna, Italy

²INAF – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, via Piero Gobetti 93/3, I-40129 Bologna, Italy

³SISSA – International School for Advanced Studies, via Bonomea 265, I-34136 Trieste, Italy

⁴INFN – Sezione di Trieste, via Valerio 2, I-34127 Trieste, Italy

⁵INFN – Sezione di Bologna, viale Berti Pichat 6/2, I-40127 Bologna, Italy

⁶Dipartimento di Fisica, Università degli Studi Roma Tre, via della Vasca Navale 84, I-00146 Roma, Italy

Accepted 2019 July 12. Received 2019 July 11; in original form 2019 April 1

ABSTRACT

Cosmic voids are large underdense regions that, together with galaxy clusters, filaments and walls, build-up the large-scale structure of the Universe. The void size function provides a powerful probe to test the cosmological framework. However, to fully exploit this statistics, the void sample has to be properly cleaned from spurious objects. Furthermore, the bias of the mass tracers used to detect these regions has to be taken into account in the size function model. In our work, we test a cleaning algorithm and a new void size function model on a set of simulated dark matter halo catalogues, with different mass and redshift selections, to investigate the statistics of voids identified in a biased mass density field. We then investigate how the density field tracers' bias affects the detected size of voids. The main result of this analysis is a new model of the size function, parametrized in terms of the linear effective bias of the tracers used, which is straightforwardly inferred from the large-scale two-point correlation function. This method is a crucial step in exploiting real surveys. The proposed size function model has been accurately calibrated on halo catalogues, and used to validate the possibility to provide forecasts on the cosmological constraints, namely on the matter density contrast, Ω_M , and on the normalization of the linear matter power spectrum, σ_8 , at different redshifts.

Key words: methods: statistical – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

Cosmic voids are large underdense regions from which matter is evacuated as a result of the collapse of the matter in between their boundaries and the repulsive action of dark energy (DE). They originate from the evolution of underdensities in the primordial density field. Voids constitute a major component of the Universe, while galaxy clusters enclose most of the mass, voids are the dominant spatial elements, accounting for about 90 per cent of the entire volume of the Universe (Platen, van de Weygaert & Jones 2007). Their sizes span over a wide range of scales, from diameters of few Mpc (*minivoids*) to about 200 Mpc (*supervoids*) (Tikhonov & Karachentsev 2006; Szapudi et al. 2015). Voids are only mildly non-linear objects, and tend to become more spherical as they evolve (Icke 1984; van de Weygaert & van Kampen 1993; Sheth & van

de Weygaert 2004), which suggests that their isolated evolution should be easier to reconstruct than that of positive perturbations, despite their sphericity can be compromised during their growth and merging.

Thanks to their relatively simple structure and shape, voids represent the ideal environment to test a variety of cosmological parameters. They represent a population of statistically ideal spheres with a uniform distribution in a homogeneous and isotropic universe, so that their observed shape can be used to probe the assumed cosmological model by means of the Alcock–Paczynski test (Alcock & Paczynski 1979, see e.g. Lavaux & Wandelt 2012; Sutter et al. 2012, 2014; Hamaus et al. 2014, 2016). Moreover, being almost completely devoid of matter by definition, voids are extremely sensitive to diffuse components and have indeed been shown to possess great potential for constraining the DE properties (Lee & Park 2009; Pisani et al. 2015), in particular for scalar field DE models (Bos et al. 2012; Adermann et al. 2017, 2018), and the mass of neutrinos (Massara et al. 2015; Kreisch et al. 2018; Sahlén

* E-mail: sofia.contarini3@unibo.it

2019). Thanks to their intrinsic low-density environment, cosmic voids have also proved to be promising objects to study modified gravity theories, since the effects of these scenarios, alternative to the general relativity (GR), are expected to be more prominent in voids (Clampitt, Cai & Li 2013; Cai, Padilla & Li 2015; Barreira et al. 2015; Zivick et al. 2015; Falck et al. 2018; Sahlén & Silk 2018). Deviations from GR can be observed also measuring the matter density profile of cosmic voids, which can be reconstructed exploiting voids as weak gravitational (anti-)lenses to infer their projected surface mass density (see e.g. Melchior et al. 2014; Clampitt & Jain 2015; Sánchez et al. 2017; Davies, Cautun & Li 2018). Modified gravity also causes a faster expansion around these objects, that can be revealed measuring the redshift-space distortions (RSD) in the cross-correlation of galaxies and void centres (Hamaus et al. 2015; Cai et al. 2016; Hamaus et al. 2017; Achitouv 2017; Hawken et al. 2017). One of the possible advantages of studying RSD around cosmic voids is that, in these regions, galaxy velocities are dominated by coherent bulk flows. Therefore the non-linear contributions can be in principle neglected and the RSD can be modelled using linear theory only (Nadathur, Carter & Percival 2019; Nadathur & Percival 2019).

To exploit cosmic voids as cosmological probes, their statistical properties have to be modelled reliably (Nadathur & Hotchkiss 2015a,b; Pollina et al. 2016). In this work we focus on void abundances. The same excursion-set approach used for the mass function of dark matter (DM) haloes can be used also to model the size function of cosmic voids (Sheth & van de Weygaert 2004). However, this model cannot accurately reproduce the number function of voids identified in cosmological simulations. Therefore, many studies have been conducted to better understand the evolution of voids over cosmic time and their statistics (Jennings, Li & Hu 2013; Pisani et al. 2015; Achitouv, Neyrinck & Paranjape 2015; Pycke & Russell 2016; Wojtak, Powell & Abel 2016). Moreover, the distribution of luminous tracers, such as e.g. galaxies and galaxy clusters, that are used to identify the voids, is biased with respect to the distribution of the underlying DM. It has been shown that the tracer bias plays a crucial role in determining the void profiles and size distributions. Having a reliable model to account for the effect of the tracer bias is thus mandatory to extract robust cosmological constraints from void statistics (Pollina et al. 2017, 2018; Nadathur & Percival 2019).

Recently, Ronconi et al. (2019) tested the void size function model developed by Sheth & van de Weygaert (2004), as revisited by Jennings et al. (2013), on a series of unbiased simulated tracer catalogues, and extended the model to the case of voids identified in the distribution of DM haloes. In this work, we further validate the model on a larger set of catalogues with different mass and redshift selections. Moreover, we provide a new parametrization of the void size function model as a function of the large-scale effective linear bias of the tracers. This represents a crucial ingredient to extract cosmological constraints from the statistical distribution of voids detected from real galaxy or cluster catalogues, when no direct information on the DM field is available. Finally, we investigate the cosmological constraints that can be inferred from the void size function at different redshifts.

Our work is organized as follows. In Section 2, we outline the methods employed for the identification of voids and the procedure of data reduction. In Section 3.1, we present the theoretical definition of cosmic voids and some of the existing models developed for the void size function, then we describe the method adopted to rescale the abundances of voids identified in the mass tracer distribution as a function of the tracer bias. In Section 4, we apply

the techniques previously described to simulated halo catalogues with different redshift and mass selections. We provide a relation between the effective linear bias of all the tracers and the one estimated inside voids, which is the one we use to rescale the void size function model. Then, we measure the void size function in all our halo catalogues, and compare it to the new theoretical model, exploiting the void abundances to test the possibility of deriving constraints on the main cosmological parameters. Finally, in Section 5, we summarize our results and discuss future developments of this work.

2 VOID CATALOGUES

In this section, we present the set of Λ cold dark matter (Λ CDM) N -body simulations used in our work, and the methods applied to build and clean the catalogues of cosmic voids. With the cleaning algorithm, we aim at aligning the objects included in the void catalogue with the adopted definition of cosmic void, which is fundamental to derive measured void size functions in agreement with theory predictions.

2.1 Simulated halo catalogues

In this work, we make use of simulated halo catalogues extracted from a set of high resolution N -body simulations (Baldi 2012) of the standard Λ CDM cosmology, performed with the C-GADGET module (Baldi et al. 2010). We adopt a model consistent with WMAP7 constraints (Komatsu et al. 2011), with $\sigma_8 = 0.809$, $h_0 = 0.703$, $\Omega_\lambda = 0.7289$, $\Omega_M = 0.2711$, $\Omega_b = 0.0451$, and a power spectrum with an initial scalar amplitude of $\mathcal{A}_s = 2.194 \times 10^{-9}$ and a primordial spectral index of $n_s = 0.96$. The simulations followed the dynamical evolution of 2×1024^3 particles: half of them are DM particles, while the other half is composed by non-collisional gas particles. Specifically, the catalogue covers a volume of $(1 \text{ Gpc } h^{-1})^3$, with a mass resolution of $\sim 6 \times 10^{10} M_\odot h^{-1}$ for the DM particles. To test the procedure described in Section 3.2, we built a set of DM halo catalogues with a friends-of-friends (FoF) algorithm,¹ applying five different mass selection cuts: $\{2 \times 10^{12}, 2.5 \times 10^{12}, 5 \times 10^{12}, 7.5 \times 10^{12}, 10^{13} M_\odot h^{-1}\}$, at three different redshifts $\{z = 0, 0.55, 1\}$. These mass cuts are applied to the FoF mass in order to inspect a sufficiently wide range of values for tracers' bias.² The redshifts are chosen instead to span a significant fraction of cosmic time over which FoF haloes with masses greater than $10^{12} M_\odot h^{-1}$ are resolved. This range allows to test our methodology on haloes corresponding to common density peaks (low redshifts and low masses) and on newly forming haloes corresponding to rare density peaks. The results obtained for the halo catalogues with $M_{\min} = 2.5$ and $7.5 \times 10^{12} M_\odot h^{-1}$ are consistent with the ones of the other catalogues, and do not add any relevant information to the overall outcome of the paper. Thus, we will not show them in the figures, with the only exception of Fig. 4.

¹The algorithm makes use of a linking length $\ell = 0.2 \cdot \bar{d}$, where \bar{d} is the mean interparticle separation, gathering the CDM particles as primary tracers of the local mass density, and then attaching baryonic particles to the FoF group of their nearest neighbour.

²Other methods to measure halo masses were applicable in this case, e.g. using spherical overdensity masses. Anyway, the mass-cut criterion, as well as the redshift selection, are not relevant in our work and do not influence the outcomes of the manuscript.

2.2 Building and cleaning the void catalogues

Many different void finders have been developed over the last decades due to the non-general concordance in the definition of voids (see e.g. Colberg et al. 2008; Micheletti et al. 2014; Elyiv et al. 2015). In this paper, we make use of the Void Identification and Examination toolkit (VIDE, Sutter et al. 2015) to construct our void catalogues. VIDE belongs to the class of algorithms based on geometrical criteria. It implements an enhanced version of the ZOnes Bordering On Voidness (ZOBOV) algorithm (Neyrinck 2008). ZOBOV is a popular publicly available code that finds density depressions in a 3D set of points, without any free parameter or assumption about the void shape. The algorithm is based on a procedure called *Voronoi tessellation*, which associates to each tracer a cell of volume that is closer to it than to any other tracer. Then the local density minima are found, and the *watershed technique* is performed. Specifically, the shallower zones are merged together starting from the minima, forming a hierarchical tree of voids and subvoids. The process of rising the density threshold goes on until a deeper zone is encountered. The effective radius of voids is defined as the radius of a sphere containing the same volume as the watershed region, and the void centre is defined as the volume-weighted barycentre, \bar{X} , of the N Voronoi cells that define the void,

$$\bar{X} = \frac{\sum_{i=1}^N \bar{x}_i V_i}{\sum_{i=1}^N V_i}, \quad (1)$$

where \bar{x}_i are the coordinates of the i th tracer of that void, and V_i the volume of its associated Voronoi cell. Therefore, the void centre does not necessarily coincide with the position of a tracer.

Once a candidate void catalogue is built, we apply the pipeline introduced in Ronconi & Marulli (2017), which has been recently implemented in the COSMOBOLOGNALIB.³ The procedure standardizes the outcome of void finders so as to make them directly comparable to model predictions. The cleaning algorithm is totally independent of the void finder employed since it makes use of the positions of void centres only. The goal is to take a candidate list of void centres and produce a catalogue of non-overlapping spherical underdensities, ‘void’. Our cleaning algorithm can be divided in three main steps:

(i) The underdense regions that do not satisfy the following criteria are rejected from the catalogue: (i) the effective radii have to be greater than a given scale, R_{\min} , which is chosen to remove objects that are under a certain resolution threshold; (ii) the central density has to be lower than $(1 + \delta_v^{\text{NL}})\bar{\rho}$, where δ_v^{NL} is a given non-linear underdensity threshold (see Section 3.2), and $\bar{\rho}$ is the mean density of the tracers. In this way, we are rejecting voids that are not relevant for our analysis, that is those regions that cannot be defined as cosmic voids according to our definition.

(ii) The effective void radii are rescaled: the algorithm reconstructs the density profile of each void and the value of the radius is increased until the sphere reaches a specific density contrast threshold, δ_v^{NL} . This value is not universal, any other threshold

Table 1. Void counts in five logarithmic bins of void effective radii, \bar{R}_{eff} , in the range [18–60] $\text{Mpc } h^{-1}$, for DM halo catalogues with different mass and redshift selections, after the cleaning procedure has been applied.

$M_{\min} (M_{\odot} h^{-1})$	N_{tot}	$\bar{R}_{\text{eff}} (\text{Mpc } h^{-1})$				
		20.5	26.0	33.1	42.1	53.6
		$N(\bar{R}_{\text{eff}})$				
$z = 0.00$						
2×10^{12}	1063	719	288	53	3	0
5×10^{12}	1007	544	333	115	15	0
10^{13}	803	291	309	160	39	4
$z = 0.55$						
2×10^{12}	1053	690	301	56	6	0
5×10^{12}	943	444	356	120	22	1
10^{13}	693	196	256	176	49	7
$z = 1.00$						
2×10^{12}	1090	698	314	72	6	0
5×10^{12}	850	370	301	146	33	0
10^{13}	557	140	170	156	77	14

sufficiently high to enclose enough tracers and sufficiently low to identify the voids would be valid (this will be intensively discussed in Section 3.2).

(iii) When two voids do overlap (thus when the distance between void centres is less than the sum of their radii), the one with the higher central density is rejected, avoiding double countings. This choice favours the selection of larger, most underdense voids.

The effect of the cleaning procedure is to reshape the selected voids as spherical non-overlapping regions, centred in density depths of the tracer density field, embedding a fixed density contrast (see Section 3.1). As a consequence, the void number counts result lower with respect to the original output of VIDE. Moreover, as it can be seen from Table 1, (i) the total number of void counts tends to decrease for tracer catalogues with higher mass selections due to the lowering of the resolution, and (ii) the void radii are shifted towards higher values because of the consequent reduction of the mean mass density.

Our choice of modelling the underdensity regions as spheres is aimed at comparing void statistics directly to theoretical models. We do not need to reconstruct accurately the real shape of individual voids. Although real voids are not spherical objects, the mean void ellipticity is small in standard cosmological frameworks (Verza et al. 2019). We can thus reasonably assume that the voids’ geometry is spherical on average (Lavaux & Wandelt 2012).

3 THE VOID SIZE FUNCTION MODEL

In this section, we first present the theoretical model for the void size function that we use in our work. Secondly, we describe the method we adopt to rescale the model to make it directly comparable with the abundance of voids identified in the distribution of biased tracers. In the end, we focus on the measure of the tracer bias in overdensity and underdensity regions, since the value of the latter is required for the re-parametrization of the theoretical model.

³The COSMOBOLOGNALIB (Marulli, Veropalumbo & Moresco 2016) is a large set of *free software* C++ /PYTHON libraries that provide an efficient numerical environment for cosmological investigations of the large-scale structure of the Universe. Thanks to the large amount of classes and functions recently implemented, these libraries offer the necessary tools to analyse cosmic void catalogues and perform all the statistical analyses of this work. The libraries are freely available at the following GitHub repository: <https://github.com/federicomarulli/CosmoBolognaLib>.

3.1 The size function model of voids detected in the DM distribution

Contrary to what happens in the case of overdensities, voids typically do not invert their expansion during their growth, so they cannot collapse and virialize like DM haloes. Instead, they expand at a *super-Hubble* rate, which is inversely proportional to the density enclosed in their boundary. Considering an initial negative top-hat perturbation, and modelling it as a set of concentric shells, the inner shells will expand faster than the outer ones. This implies that the shells near the centre of the underdensity will eventually reach the more external ones. This event is called *shell-crossing*. When this occurs, we can consider that a void is formed. After the shell-crossing, the void recovers the overall expansion rate, growing with the Hubble flow. This phenomenon is completely described by the spherical non-linear evolution model of an isolated spherically symmetric density perturbation. It can be demonstrated that in linear theory this event takes place at a fixed value of the density contrast, $\delta_v^{\text{NL}} \approx -2.71$, for an Einstein–de Sitter universe (EdS). Therefore, we can define voids as underdense, spherical, non-overlapping regions, which have gone through shell crossing.

The void size function, that is the comoving number density of cosmic voids as a function of their effective radii, has been modelled for the first time by Sheth & van de Weygaert (2004, hereafter the SvdW model), with the same excursion-set approach used to model the mass function of DM haloes (Press & Schechter 1974; Bond et al. 1991). The void size function in linear theory can be written as follows:

$$\left. \frac{dn}{d \ln r} \right|_{\text{lin}} = \frac{f_{\text{in}\sigma}(\sigma) d \ln \sigma^{-1}}{V(r) d \ln r}, \quad (2)$$

where $f_{\text{in}\sigma}$ is the fraction of the Universe occupied by cosmic voids, as predicted by the excursion-set theory:

$$f_{\text{in}\sigma} = 2 \sum_{j=1}^{\infty} j \pi x^2 \sin(j \pi \mathcal{D}) \exp \left[-\frac{(j \pi x)^2}{2} \right], \quad (3)$$

where

$$x \equiv \frac{\mathcal{D}}{|\delta_v^L|} \sigma, \quad (4)$$

and

$$\mathcal{D} \equiv \frac{|\delta_v^L|}{\delta_c^L + |\delta_v^L|}. \quad (5)$$

In the previous equations, σ is the square root of the mass variance, while δ_v^L and δ_c^L represent the shell-crossing threshold and the critical value for the collapse of an overdense shell in an EdS universe, respectively.⁴ The latter is expected to vary within $1.06 \leq \delta_c^L \leq 1.686$, since both the turnaround and the collapse density contrast value can be considered acceptable assumptions.

In order to derive the void size function in the non-linear regime, SvdW assumed that the total number of voids has to be conserved in the transition from linearity to non-linearity. This condition leads

to a correction in the void radius by a factor $C \propto (1 + \delta_v^{\text{NL}})^{-1/3}$:

$$\left. \frac{dn}{d \ln r} \right|_{\text{SvdW}} = \left. \frac{dn}{d \ln(Cr)} \right|_{\text{lin}}. \quad (6)$$

However, according to equation (6), the fraction of the volume occupied by voids can be larger than the total volume of the Universe. To address this issue, Jennings et al. (2013) proposed a *volume conserving* model (hereafter the Vdn model), in which the total volume occupied by cosmic voids is conserved in the transition to the non-linear regime. In particular, the Vdn model can be obtained as follows:

$$\left. \frac{dn}{d \ln r} \right|_{\text{Vdn}} = \left. \frac{dn}{d \ln r} \right|_{\text{lin}} \frac{V(r_L) d \ln r_L}{V(r) d \ln r}, \quad (7)$$

where the subscript L indicates a value derived in linear theory. Ronconi et al. (2019) showed that the Vdn model can predict accurately the measured void size function of unbiased tracers, provided that the void catalogue is appropriately cleaned from spurious voids and the void radii are rescaled to a fixed density threshold (see Section 2.2).

3.2 The size function model of voids detected in biased tracer distribution

The goal of the cleaning procedure is to make the measured void size function directly comparable to the Vdn model. The method is based on the requirement that the spherically averaged density contrast embedded inside the void effective radius has to coincide with the value defined by the theory. As shown in Ronconi et al. (2019), dealing with unbiased tracers, this value can correspond to that of the shell crossing in the non-linear regime, $\delta_{\text{sc}}^{\text{NL}} = -0.795$. With this prescription, the measured size function of voids identified in the DM field is consistent with the one predicted by the Vdn model.

It is important to notice that the choice to rescale the void radii to the specific density contrast characteristic of the shell-crossing is not universal. In fact, in order to have correspondence with the theory, it is only required to rescale the radii at a chosen density threshold with the cleaning algorithm and use the same density contrast (converted to linear theory) also in the theoretical void size function. This overdensity threshold $\delta_{\text{sc}}^{\text{NL}}$, which identifies the time at which cosmic voids form, has to be rescaled at redshift $z > 0$, using the growth factor:

$$\delta_{\text{sc}}^L(z) = \delta_{\text{sc}}^L(0) \frac{D(z)}{D(0)}, \quad (8)$$

where $\delta_{\text{sc}}^L(0) = -2.71$ is the shell-crossing density contrast in linear theory at $z = 0$, and $D(z)$ is the growth factor. Therefore, voids formed at lower density contrast values in the past. After shell crossing, the void radii continue to grow with cosmic time, and the enclosed volume becomes increasingly underdense. Choosing the threshold $\delta_v^{\text{NL}} = \delta_{\text{sc}}^{\text{NL}}(z)$, we are rescaling voids to the radii embedding the density contrast typical of the phenomenon of the shell crossing at that epoch. In other words, we are rescaling void radii to the size they had when they formed. This method cleans the catalogue of newly forming voids that have yet to meet the underdensity criterion. In fact, the ones that cannot be rescaled to this particular density contrast (because they are not enough underdense at that epoch) are rejected. However, every negative density threshold $-1 < \delta_v^{\text{NL}} < 0$ is allowed in principle, provided that the same value is used in the theoretical size function.

⁴In this paper, we indicate with the superscripts L and NL the density contrasts derived in linear and non-linear regimes, respectively. In absence of any superscript, we take for granted the reference to the non-linear counterpart. Moreover, with the subscript v we refer to the values measured inside voids, both for DM and biased mass tracers. We will use the subscript tr to indicate generically any type of mass tracers, and the subscript *halo* to indicate specifically the DM haloes.

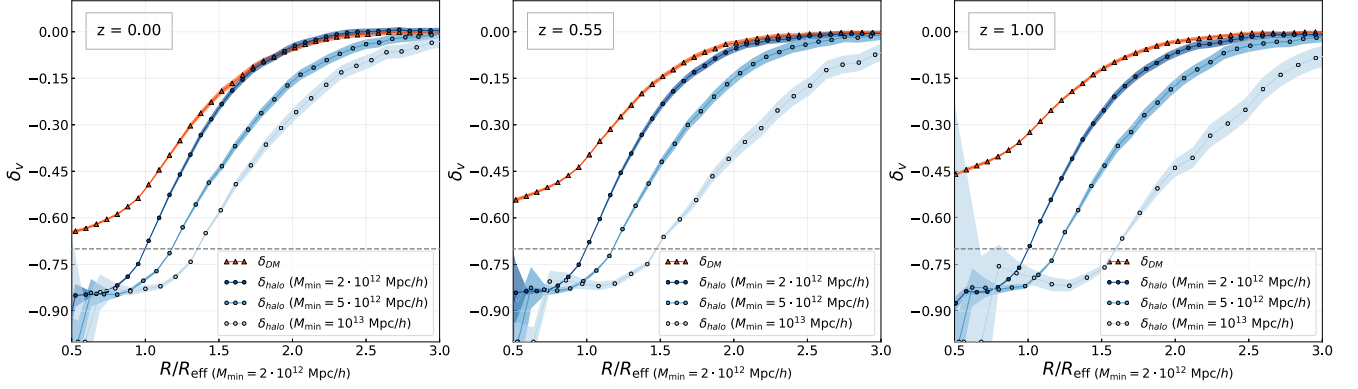


Figure 1. Spherically averaged density profiles measured from the centres of voids identified in the tracer distribution at redshifts $z = 0$ (left), $z = 0.55$ (centre), $z = 1$ (right). The red lines represent the median of the profiles computed in the DM particle distribution, while the blue ones indicate the profiles in the DM halo catalogues with different mass cuts. The horizontal dashed line indicates the value of the density contrast threshold ($\delta_{v, \text{tr}}^{\text{NL}} = -0.7$) selected in the cleaning procedure. All the profile radii are rescaled to the mean effective radius of the catalogue with $M_{\text{min}} = 2 \times 10^{12} M_{\odot} h^{-1}$, in order to show the effect of the rescaling procedure of the cleaning algorithm. The shaded areas represent 2σ confidence regions, that is two times the standard deviation of the distribution of the mean values.

Dealing with mass tracers, the effect of the tracer bias has to be taken into account to extract accurate cosmological constraints from the void number counts (see e.g. Pollina et al. 2017). Let us assume that the voids identified in the DM and the mass tracer density fields have the same radii when the phenomenon of the shell crossing occurs. This implies that voids found in the biased tracer distribution have a lower embedded density contrast with respect to the ones traced by DM particles. This is illustrated in Fig. 1, that shows the spherically averaged void density profiles⁵ as traced by either DM or DM haloes with different biases, at three different redshifts (see also Ronconi et al. 2019). With this assumption, and given that the DM density field within voids is linearly related to the density field traced by biased tracers (Pollina et al. 2018), the threshold at which the void radii have to be rescaled corresponds to:

$$\delta_{v, \text{tr}}^{\text{NL}} = b \delta_{v, \text{DM}}^{\text{NL}}, \text{ with } \delta_{v, \text{DM}}^{\text{NL}} = -0.795. \quad (9)$$

It is evident that for $b > 1$, the density contrast can reach values so low that the phenomenon of the shell crossing might not even happen, since the lowest minimum is $\delta_{v, \text{tr}}^{\text{NL}} = -1$ (corresponding to the state without any tracer). Therefore, it is not possible to perform this technique to rescale the void radii in the case of biased tracers. For this reason, we use a different density contrast in the rescaling procedure, fixing the threshold in the tracer distribution, instead of in the DM one. In particular, we set a threshold equal to -0.7 for all the considered halo catalogues. Thus, we rescale all the voids found in the tracer catalogues to an effective radius such that the spherically averaged density contrast they contain is $\delta_{v, \text{tr}}^{\text{NL}} = -0.7$. This choice is aimed at having not too small void radii (the higher is the threshold, the higher is the radius), in order to enclose a sufficient number of tracers. In fact, the resolution of the tracer catalogue does not allow us to identify voids with radii smaller than 2–3 times the mean inter-particle separation. At the same time, we require that the chosen threshold is not too high,

⁵All the void profiles with effective radii, R_{eff} , larger than two times the mean interparticle separation are stacked in these plots. As a result of the cleaning procedure, that rescales every void radii at the same level of density contrast, the profiles do not show a clear dependence on the void effective radius and they are therefore averaged together.

since the selected regions have to be enough underdense to be still classified as voids. Moreover, we want to keep the threshold low also to prevent the possible overlap between adjacent voids, that would make the cleaning algorithm to discard the smaller, thus the one with higher central density, to prevent double counts (see Section 2.2).

To model the theoretical size function of voids identified in the mass tracer field, we follow the prescription described in Ronconi et al. (2019). This is based on the reasonable assumption that voids identified in the DM and the tracer field are equal in number, and that their centre positions are approximately the same.⁶ Since the Vdn model can predict the number of voids with a certain radius, the simplest procedure to apply is to rescale the theoretical size function dividing the chosen threshold by the bias value:

$$\delta_{v, \text{DM}}^{\text{NL}} = \frac{\delta_{v, \text{tr}}^{\text{NL}}}{b}, \text{ with } \delta_{v, \text{tr}}^{\text{NL}} = -0.7. \quad (10)$$

We convert $\delta_{v, \text{DM}}^{\text{NL}}$ to its linear counterpart, with the fitting formula provided by Bernardeau (1994):

$$\delta_{v, \text{DM}}^L = C [1 - (1 + \delta_{v, \text{DM}}^{\text{NL}})^{-1/C}], \quad (11)$$

with $C = 1.594$. This equation is exact for models with null cosmological constant Λ , and is a good fit for any values of Λ , especially for the underdense regions. The density contrast given by equation (11), $\delta_{v, \text{DM}}^L$, has to be used in equations (4) and (5). This is basically equivalent to expand the radii of voids identified in the DM field (embedding the same density contrast -0.7), in order to match the same radius of the ones identified in the tracer field. In this way, we are able to recover the theoretical size function taking into account the effect of the bias, that shifts the size function to higher void radii.⁷

⁶We tested this hypothesis using a catalogue of voids identified in the DM density field and cleaning it using the corresponding distribution of DM haloes as tracer. The results obtained are in agreement with the ones found with the voids identified in the biased tracer distribution. Therefore this assumption can be considered statistically valid, even if the correspondence between void centres in different mass density fields is not always exact.

⁷A new algorithm to rescale the void size function model as a function of the tracer bias has been implemented in the COSMOBLOGICALIB. The code

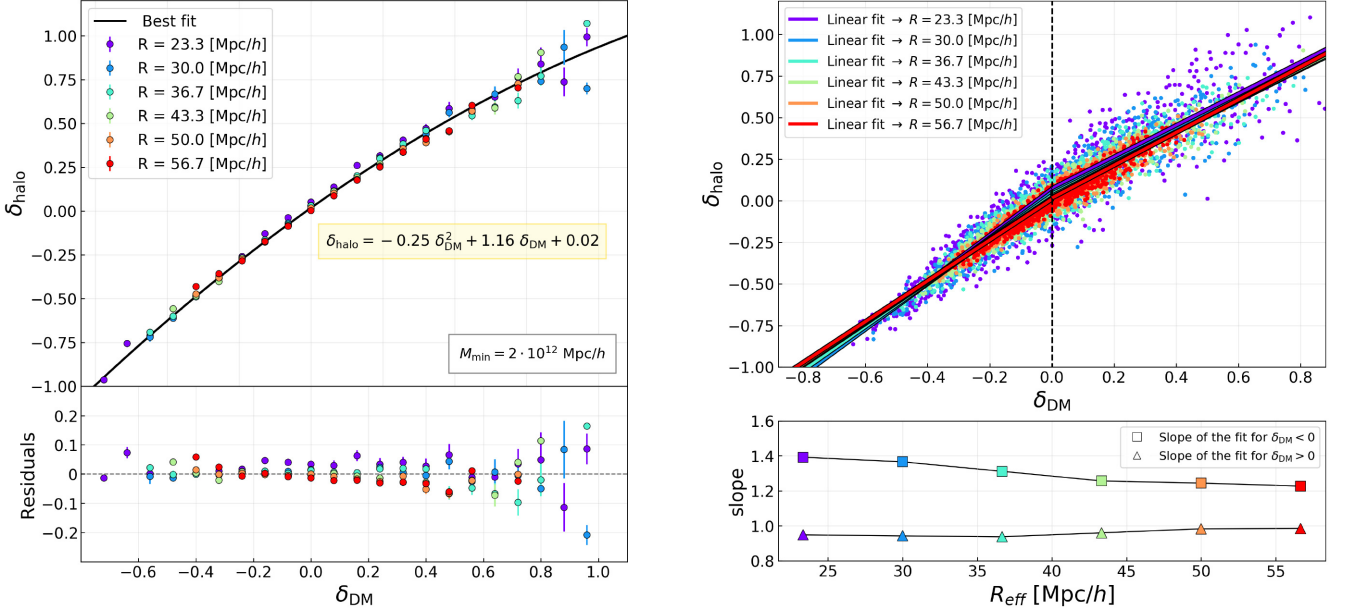


Figure 2. Relation between the density contrast computed in the DM distribution (δ_{DM}) and in the tracer distribution (δ_{halo}). The data points are computed as the spherically averaged density contrast for 1000 random positions in the halo catalogue with $M_{\text{min}} = 2 \times 10^{12} M_{\odot} h^{-1}$ at $z = 0$, averaging in different radius bins. The different colours refer to the different radius sizes of the spheres used to compute the density contrast. Left: in the upper sub-panel, the data are computed as the median of the values of δ_{halo} in different bins of δ_{DM} , with error bars computed as the ratio between the standard deviation and the square root of the number counts in each bin. The points are fitted with a second-order polynomial, whose equation is reported in the yellow insert and represents the non-linear bias function. In the lower sub-panel are reported the residuals from the quadratic fit. Right: in the upper sub-panel, the points with $\delta_{\text{DM}} > 0$ and < 0 of each radius bin are fitted separately with a linear relation. In the lower sub-panel is shown the variation of the slope of each fit as a function of the radius of the sphere used to compute the averaged density contrast.

3.3 The bias of tracers in overdensity and underdensity regions

The bias of cosmic tracers is a non-linear stochastic function described by the conditional probability of tracer density contrast, δ_{tr} , given the mass density contrast δ_{DM} (see e.g. Dekel & Lahav 1999; Di Porto et al. 2016, and references therein). This is shown in Fig. 2, where the density contrast of a halo catalogue analysed in this work ($M_{\text{min}} = 2 \times 10^{12} M_{\odot} h^{-1}$ at $z = 0$) is plotted against the corresponding DM density contrast, smoothing the density field at 1000 random positions with top-hat spherical filters with different radii. As shown in the left-hand panel of Fig. 2, the data are well fitted by a second-order polynomial. However, a linear model is accurate enough to describe separately the points in the overdensity and the underdensity regions. Indeed, fitting all the points with a second-order polynomial the reduced chi square is $\bar{\chi}^2 = 1.977$, while fitting $\delta_{\text{DM}} > 0$ and < 0 separately with a linear relation we obtain $\bar{\chi}^2 = 1.758$ and 2.780 , respectively. The slope of the former, $b[\delta_{\text{DM}} > 0]$, represents the linear bias that can be approximately inferred e.g. from the tracer large-scale two-point correlation function (2PCF). The slope of the latter, $b[\delta_{\text{DM}} < 0]$, represents the bias of the tracers inside cosmic voids, which is the

requires in input the values of the radii at which the model is computed, the redshift of the sample, the size function model to use (e.g. SvdW and Vdn) and the effective bias of the catalogue, b_{eff} . The latter can be automatically converted to b_{rel} using the relation calibrated in this work (see Section 4.1). Moreover, a new notebook is provided to explain, step by step, how to clean a void catalogue, and how to measure and model the void size function, according to the method described in this paper.

value we actually need in order to properly rescale the void size function, as we will explain in the next section. As shown in the right-hand panel of Fig. 2, $b[\delta_{\text{DM}} < 0] > b[\delta_{\text{DM}} > 0]$.

Since $b[\delta_{\text{DM}} < 0]$ is generally not directly measurable, we shall calibrate a relation between $b[\delta_{\text{DM}} > 0]$ and $b[\delta_{\text{DM}} < 0]$ to be able to model the size function of voids detected from real tracer catalogues. Specifically, we search for a relation between the effective linear bias of the tracers used to detect cosmic voids, $b_{\text{eff}} \sim b[\delta_{\text{DM}} > 0]$ (that we measure from the tracer 2PCF at large scales, as described in Appendix A), and the linear bias of tracers *inside* the detected voids. A convenient estimate of the latter can be assessed through the ratio between $\delta_{\text{v,halo}}^{\text{NL}}$ and $\delta_{\text{v,DM}}^{\text{NL}}$ at a distance of R_{eff} from void centres (Ronconi et al. 2019):

$$b_{\text{punct}} \equiv \frac{\delta_{\text{v,tr}}^{\text{NL}}(R = R_{\text{eff}})}{\delta_{\text{v,DM}}^{\text{NL}}(R = R_{\text{eff}})} \quad (12)$$

The *punctual bias* given by equation (12) characterizes the relation between the density contrast measured in the tracer field and in the DM field *punctually*, that is at $R = R_{\text{eff}}$. Since in our analysis the value of $\delta_{\text{halo}}(R_{\text{eff}})$ is fixed at -0.7 , then $\delta_{\text{DM}}(R_{\text{eff}})$ is exactly the value we need to rescale the void size function model (see Section 3.2).

An alternative method to estimate $b[\delta_{\text{DM}} < 0]$ is the one employed by Pollina et al. (2018). They found a linear relation between the density profiles of tracers and DM inside voids. The slope of this relation, b_{slope} , provides an estimate of the tracer bias in underdensity regions. We discuss about this method in Appendix B, but choose not to use it in our analysis as it is more prone to uncertainties.

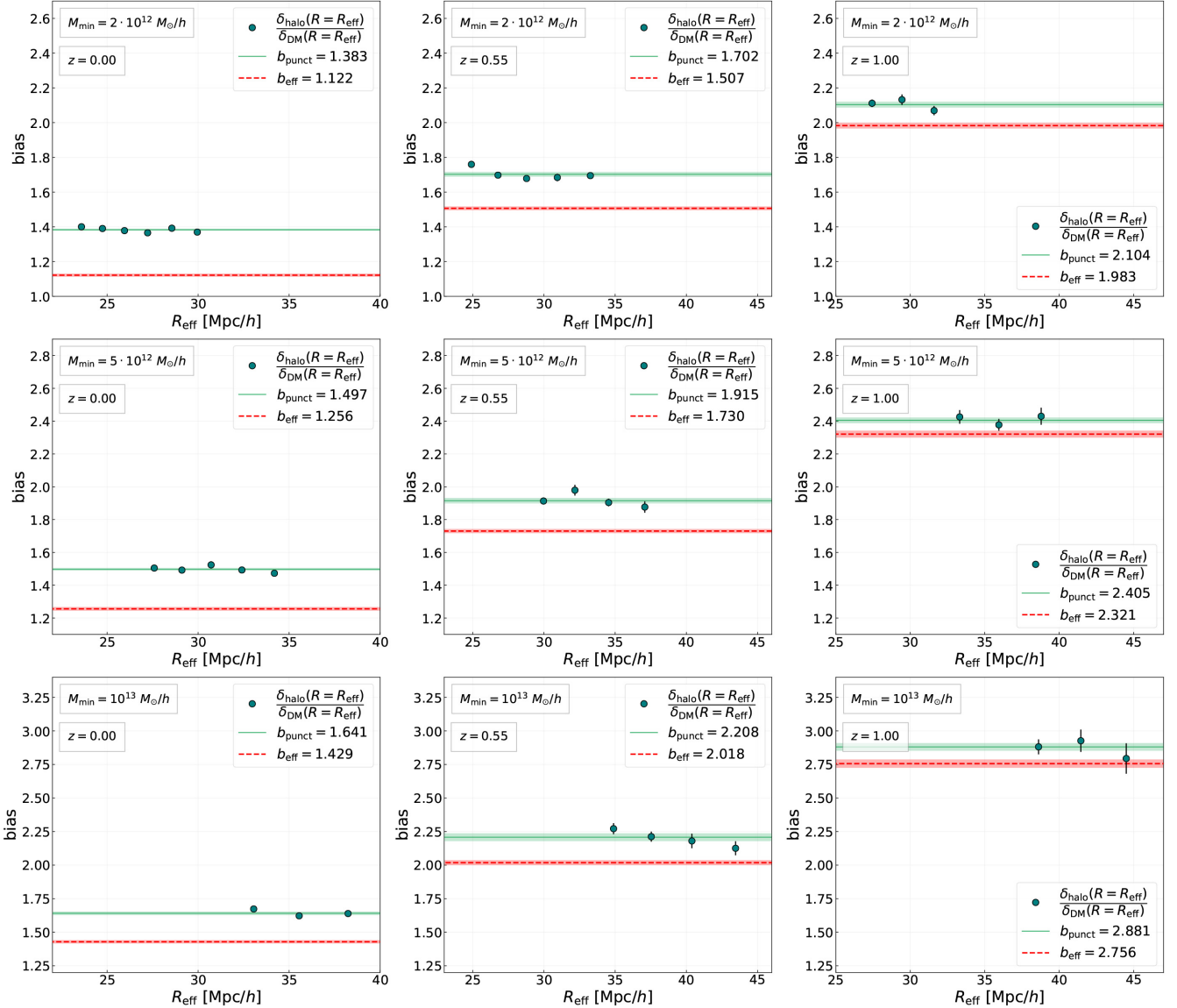


Figure 3. Measure of the tracer bias estimated as the ratio between the density contrast computed in the halo (δ_{halo}) and in the DM (δ_{DM}) density fields, at a distance of $1R_{\text{eff}}$ from the void centres. The different panels show the results obtained from the halo catalogues with $M_{\min} = 2 \times 10^{12}$, 5×10^{12} , and $10^{13} M_{\odot} h^{-1}$ (rows from top to bottom), at redshifts $z = 0, 0.55$, and 1 (columns from left to right). The dark green points represent the median of the ratio for different radius bins, with error bars representing the 1σ uncertainty. The green lines are the weighted fit of the data, b_{punct} , while the red dashed lines show the effective bias, b_{eff} . The shaded regions show the 1σ errors on the bias values.

4 RESULTS

In this section, we first estimate the linear bias of DM haloes inside voids, b_{punct} . We then model the relation between this bias and the effective linear bias of all the tracers used to detect the voids, b_{eff} . The latter is estimated from the tracer 2PCF at large scales, as explained in Appendix A. We want a size function model that can be promptly compared with real data measures. To this end, it is crucial to obtain a relation between b_{eff} and b_{punct} that can be applied, independently of the tracer used to sample the underlying DM density field. Afterwards, we measure the void size function in all our simulated catalogues, and compare the measurements with the theoretical model. Finally, we perform a Bayesian Markov Chain Monte Carlo (MCMC) statistical analysis, exploiting the calibrated bias scaling relation to construct the likelihood function. With this

approach, we investigate the constraining power of the method by assessing the posterior probability of two cosmological parameters, namely Ω_{M} and σ_8 , at varying redshift.

4.1 The bias of DM haloes inside voids

Fig. 3 shows the ratio between the density contrast of haloes and DM, $\delta_{\text{halo}}/\delta_{\text{DM}}$, measured at $R = R_{\text{eff}}$ and averaged over voids of similar effective radii, together with their weighted average values (equation 12), b_{punct} , for all the considered simulated catalogues. Therefore, in this figure, the points are obtained by computing the ratio $\delta_{\text{v, tr}}^{\text{NL}}(R = R_{\text{eff}})/\delta_{\text{v, DM}}^{\text{NL}}(R = R_{\text{eff}})$ for each void of the catalogues (with R_{eff} being the effective radius of that specific void) and binning the result as a function of R_{eff} . Then, to compute the value

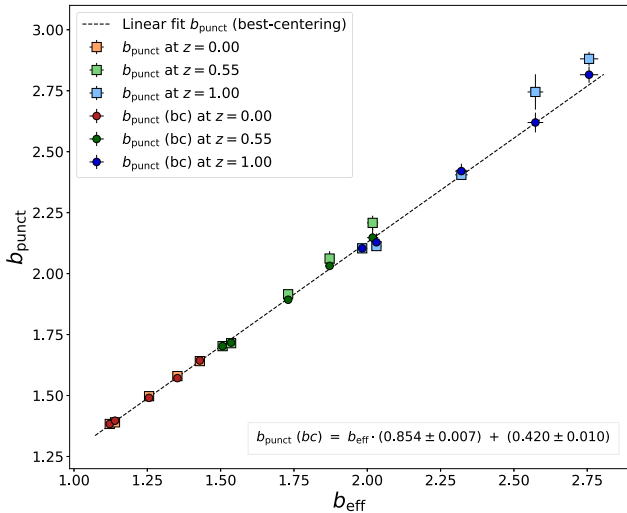


Figure 4. Relation between the effective bias (b_{eff}) and the bias measured inside voids (b_{punct}), at different redshifts. The points correspond to the data reported in Table 2, with 1σ errors. The squares are the values of b_{punct} , obtained with the method presented in Fig. 3, while the circles are estimated with the best-centering technique (see Section 4.1) and correspond to $b_{\text{punct}}(\text{bc})$. The black line is the linear fit of the $b_{\text{punct}}(\text{bc})$ values. The best-fitting parameters are reported in the label in the lower right corner.

of b_{punct} , we perform a weighted fit of these data with a constant. For comparison, we show also the effective tracer bias, b_{eff} , estimated from the 2PCF at large scales, as explained in Appendix A. As shown in Fig. 3, the $\delta_{\text{halo}}/\delta_{\text{DM}}$ ratio decreases as a function of R_{eff} , especially at high redshifts. In particular, it tends to b_{eff} at large radii, in agreement with the results obtained by Pollina et al. (2017, 2018). Nevertheless, we find that an average constant value of b_{punct} is sufficient to properly rescale the void size function, as we will show in Section 4.2.

Since in most cases it is not possible to infer the underlying DM distribution inside voids, it is worth to search for a relation between b_{punct} and b_{eff} , which can be accurately estimated e.g. from clustering measurements. This relation is displayed in Fig. 4. As it can be seen, the data can be well fitted by a simple linear model.

However, the b_{punct} values estimated in the higher bias halo catalogues tend to systematically depart from the fit, at all redshifts. The reason of this slight deviation is related to the method used to find the void centres. In fact, if the detected voids are traced by too few tracers, the VIDE method might not be sufficiently accurate to localize their centres. Computing the spherically averaged density contrast starting from a point that is not a local minimum of the density field causes systematic errors in the bias measurements. This is a natural consequence of the cleaning procedure: when rescaling the void radii, the selected threshold might be reached at smaller radii if overdense regions are included in the measurement, due to a bad centring. This is an issue especially for catalogues with a high-mass selection.

As a possible strategy to alleviate the problem, we repeat our bias measurements using in all cases, the centre positions of the voids detected in the catalogues with the lowest mass cut. We will refer to this method as our best-centering technique, and we will call $b_{\text{punct}}(\text{bc})$ the corresponding bias. As shown in Fig. 4, these bias values (shown as coloured circles) are in better agreement with

Table 2. The values of the bias with 1σ uncertainties measured in the overdensity (b_{eff}) and in the underdensity (b_{punct} and $b_{\text{punct}}(\text{bc})$) regions, for all the halo catalogues with different mass selections and redshifts.

	b_{eff}	b_{punct}	$b_{\text{punct}}(\text{bc})$
$M_{\text{min}} (M_{\odot} h^{-1})$ $z = 0.00$			
2×10^{12}	1.122 ± 0.006	1.383 ± 0.006	1.383 ± 0.006
2.5×10^{12}	1.140 ± 0.009	1.390 ± 0.005	1.397 ± 0.004
5×10^{12}	1.256 ± 0.011	1.497 ± 0.008	1.491 ± 0.007
7.5×10^{12}	1.353 ± 0.011	1.580 ± 0.014	1.571 ± 0.009
10^{13}	1.429 ± 0.012	1.641 ± 0.013	1.644 ± 0.012
$M_{\text{min}} (M_{\odot} h^{-1})$ $z = 0.55$			
2×10^{12}	1.507 ± 0.011	1.702 ± 0.014	1.702 ± 0.014
2.5×10^{12}	1.536 ± 0.011	1.715 ± 0.018	1.717 ± 0.013
5×10^{12}	1.730 ± 0.013	1.915 ± 0.017	1.893 ± 0.012
7.5×10^{12}	1.872 ± 0.015	2.062 ± 0.030	2.032 ± 0.017
10^{13}	2.018 ± 0.019	2.208 ± 0.029	2.148 ± 0.037
$M_{\text{min}} (M_{\odot} h^{-1})$ $z = 1.00$			
2×10^{12}	1.983 ± 0.017	2.104 ± 0.017	2.104 ± 0.017
2.5×10^{12}	2.301 ± 0.017	2.113 ± 0.017	2.128 ± 0.036
5×10^{12}	2.321 ± 0.021	2.405 ± 0.018	2.420 ± 0.031
7.5×10^{12}	2.573 ± 0.027	2.745 ± 0.072	2.620 ± 0.041
10^{13}	2.756 ± 0.031	2.881 ± 0.028	2.816 ± 0.033

a linear model. Therefore, we use them to calibrate the relation between the bias measured on large scales and the one computed inside cosmic voids, obtaining the following equation:

$$b_{\text{punct}}(\text{bc}) = b_{\text{eff}} \cdot (0.854 \pm 0.007) + (0.420 \pm 0.010). \quad (13)$$

This relation can be used to estimate the bias of the tracers inside voids from the effective bias of the whole tracer population. Hereafter, the bias obtained using equation (13) will be called $f(b_{\text{eff}}) \equiv b_{\text{rel}}$. All the different bias values are reported in Table 2.

It is important to notice that the best-centering technique is not employable with real mocks, since in that case it is not possible to use more numerous tracers to improve the centre of a void. Nevertheless, in our work we choose to rely on this technique to obtain a better calibration of the relation between b_{punct} and b_{eff} . Indeed, it is convenient to calibrate the latter with $b_{\text{punct}}(\text{bc})$ to minimize the deviation of the data associated to the catalogues with higher mass selections from the linear fit. Using the best-centering technique to alleviate the problem of the sparsity of the tracers, we are able to extend our pipeline also to catalogues with lower spatial resolution.

4.2 The void size function

Here we measure the void size function of our cleaned catalogues and compare it with the theoretical predictions given by the re-parametrized Vdn model. We reject the voids that are too close to the boundaries of the simulation box, as their radii cannot be accurately rescaled by our cleaning algorithm, and we correct consequently the effective volume of the box. The theoretical size function is modelled taking into account the effect of the bias of DM haloes inside voids, as described in Section 3.2.

Fig. 5 displays our results. The new re-parametrized void size function model accurately describes all our measurements, in the full range of redshift and mass (thus bias) selections. This represents the main outcome of our analysis. We show both the size function models obtained by rescaling with b_{punct} and b_{rel} , that appear

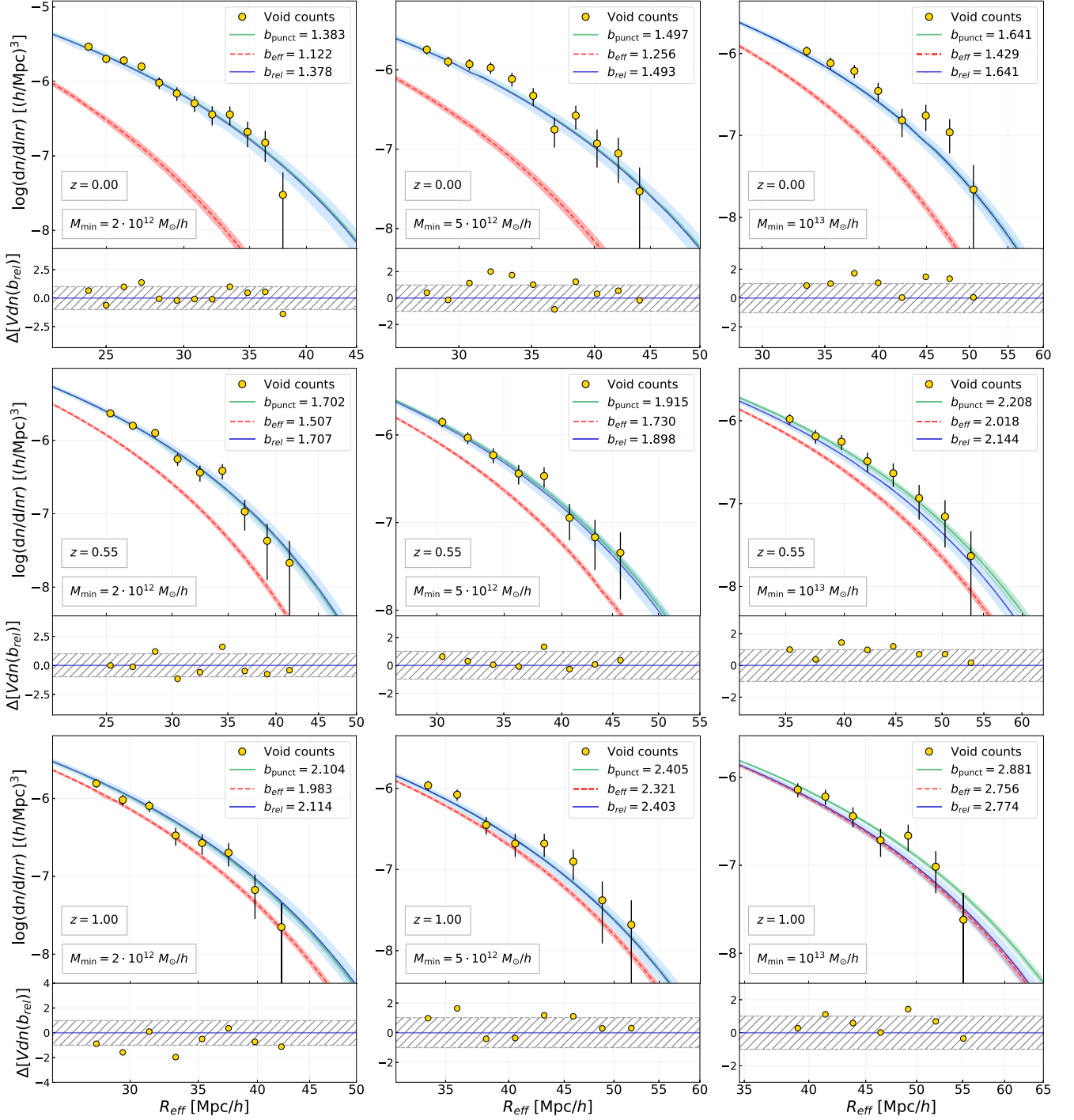


Figure 5. The measured size function of the voids (yellow dots) identified in the DM halo catalogues with $M_{\min} = 2 \times 10^{12}$, 5×10^{12} , and $10^{13} M_{\odot} h^{-1}$ (rows from top to bottom), at redshifts $z = 0, 0.55$, and 1 (columns from left to right). Voids with $R_{\text{eff}} < 2.5$ times the mean interparticle separation are rejected from the analysis. Upper sub-panels: the blue dashed lines represent the void size function obtained by rescaling the Vdn model with b_{rel} , that is the value of the bias computed from the relation shown in Fig. 4. The green solid lines show the model rescaled with the value of b_{punct} . The red dashed lines represent the model rescaled with the effective bias, b_{eff} . In all cases, the shaded areas indicate the variation of the model obtained applying 1σ errors on the value of the tracer bias. Lower sub-panels: the residuals of the void counts, computed as the ratio between the difference $\text{data} - \text{model}$ and the data errors , where the data are the measured void size function and the model is given by the re-parametrization of the Vdn model with b_{rel} . The grey dashed areas indicate the regions in which the discrepancy between the data and the model is within the data errors.

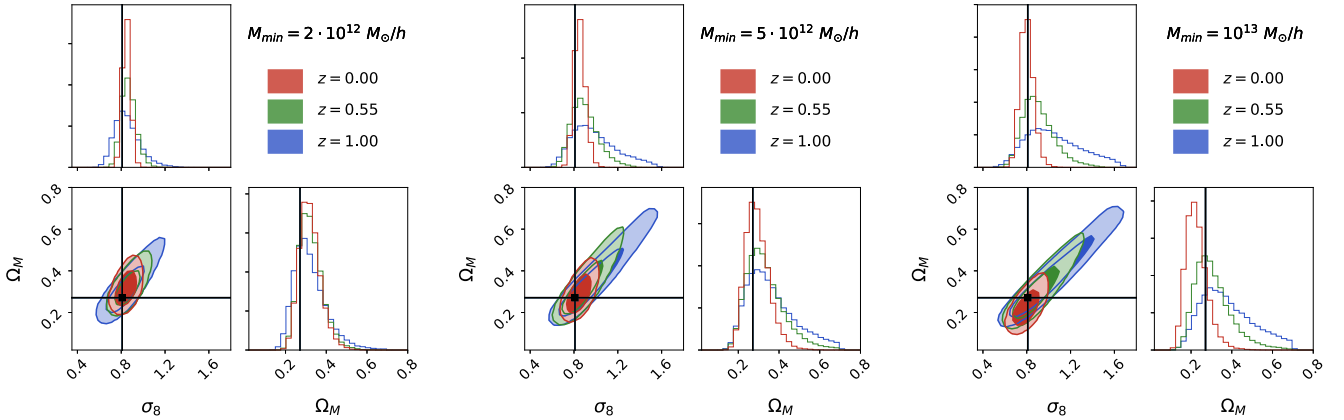


Figure 6. 68 per cent and 95 per cent contours in the σ_8 – Ω_M plane, for the halo catalogues with $M_{\min} = 2 \times 10^{12} M_{\odot} h^{-1}$ (left), $5 \times 10^{12} M_{\odot} h^{-1}$ (centre), and $10^{13} M_{\odot} h^{-1}$ (right). The colour of ellipses corresponds to different redshifts: red for $z = 0$, green for $z = 0.55$, and blue for $z = 1$. The prior distributions are uniform for σ_8 and Ω_M , and Gaussian for b_{eff} , A_{rel} , and B_{rel} . The histograms (top and bottom right panels) show the posterior distributions of σ_8 and Ω_M , respectively. The black lines represent the true *WMAP7* values ($\sigma_8 = 0.809$ and $\Omega_M = 0.2711$).

fully consistent, especially at low-redshift and bias values. The uncertainty in the identification of void centres in low-density tracer catalogues causes the slight discrepancies that can be seen at high redshifts and biases, which in any case appear not statistically significant. For comparison, we also show the model obtained by rescaling the Vdn model with the effective bias of the full DM halo population, b_{eff} . As it is clearly evident in the figure, this case underpredicts systematically the measured size function at all redshifts and biases.

The final goal of this paper is to investigate the cosmological constraints that can be derived from the void size function at different redshifts. To mimic real data analyses, we suppose to have access only to the tracer density field. With no information about the underlying total matter distribution, we have to rely on the relation found in Section 4.1. We first estimate the effective bias of the sample, b_{eff} , and we consider the coefficients shown in equation (13), A_{rel} and B_{rel} , that are the offset and the slope of the calibrated relation, respectively. These coefficients are necessary to convert b_{eff} into b_{rel} , which in turn is required to re-parametrize the Vdn model, as shown in Fig 5. Then, we perform a Bayesian statistical MCMC analysis of the measured void size function by sampling the posterior distribution of the parameters σ_8 and Ω_M . We assume uniform prior distributions for σ_8 and Ω_M , and we leave as free parameters also b_{eff} , A_{rel} , and B_{rel} , assuming in this case Gaussian prior distributions centred at the estimated values of these parameters, with standard deviations equal to their relative 1σ uncertainties. The results for our simulated catalogues with three different mass cuts and redshifts are reported in Fig. 6. The true values of the cosmological parameters are within the 68 per cent levels in all cases. In Appendix C, we investigate the systematics in the cosmological constraints possibly caused by the uncertainties in the estimation of the tracer bias, while in Appendix D, we show the outcome of combining the posterior distributions at different redshifts to achieve tighter constraints on the cosmological parameters.

5 CONCLUSIONS

In this work, we have extended the prescriptions developed by Ronconi et al. (2019) to model cosmic void abundances. In partic-

ular, we have focused on the relation between the size-abundance distribution of voids and the bias of the tracers used to detect them. We have also investigated the cosmological constraints that can be derived from void number counts at different redshifts.

The main steps and results of our work can be summarized as follows.

- (i) We have run the finding and cleaning algorithms on simulated DM halo catalogues, selected with different mass cuts to probe different tracer biases and redshifts.
- (ii) We have compared the tracer bias measured on large scales to the one measured inside cosmic voids, finding a tight relation (see equation 13) between the two.
- (iii) Then, we have provided a new parametrization of the Vdn model as a function of the large-scale tracer bias. We have calibrated the model on our simulated catalogues at different redshifts and biases.
- (iv) Finally, we have explored the constraining power of the void size function. Specifically, we have performed a Bayesian statistical inference analysis, fitting the measured size function with the new calibrated model, obtaining constraints on σ_8 and Ω_M .

In this paper, we have investigated one of the possible cosmological applications of cosmic void statistics, that is void abundances. The bias relation calibrated in Section 4.1 allows to construct the likelihood for the statistical inference analysis as a function of the large-scale effective bias of the sample. This work lays the foundations for the cosmological exploitation of the void size function, when the voids are identified in the distribution of biased tracers, such as in real data catalogues.

We note that another interesting application of our method would be to combine results using tracers with significantly different bias, e.g. combining optical surveys to HI surveys, where the bias can be negative (Pénin, Umeh & Santos 2018; Maartens, Clarkson & Chen 2018). This test is beyond the scope of the paper and we leave it as a future development of our work.

Finally it is important to notice that the relation between b_{punct} and b_{eff} provided in equation (13) is valid in the Λ CDM framework only. Extending our method to other cosmological scenarios (e.g. for

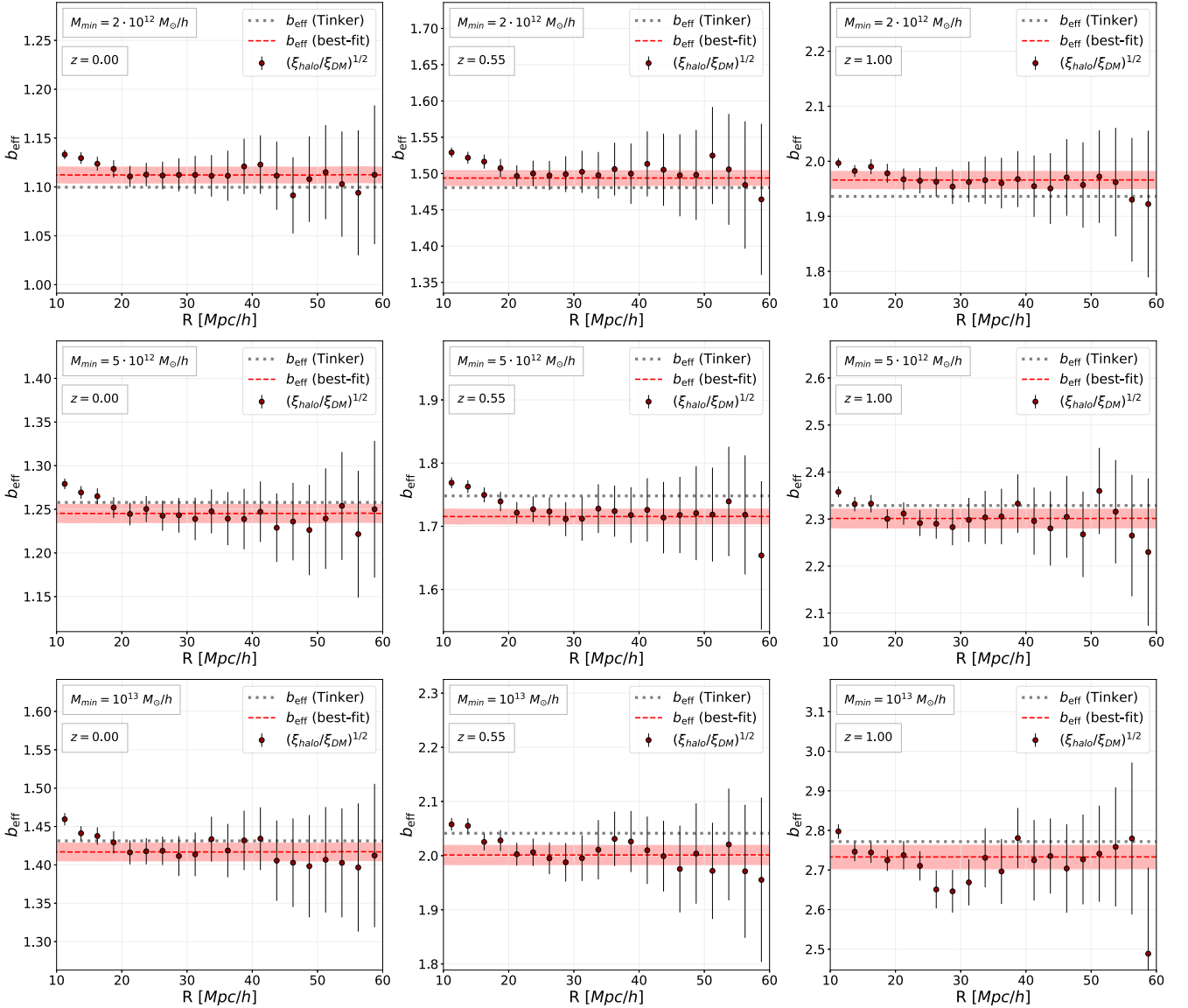


Figure 7. The halo bias for the catalogues with $M_{\min} = 2 \times 10^{12}$, 5×10^{12} , and $10^{13} M_{\odot} h^{-1}$ (rows from top to bottom), at redshifts $z = 0, 0.55$, and 1 (columns from left to right). The black points represent the square root of the ratio between the autocorrelation function of the haloes and the DM particles (equation A5). The error bars are the diagonal elements of the covariance matrix estimated with Bootstrap. The red shaded areas show the 1σ uncertainties on the best-fitting bias values estimated with the MCMC modelling, fitting in the range of radii of $[20\text{--}40] \text{ Mpc } h^{-1}$. The dashed grey lines show the theoretical predictions given by the Tinker et al. (2010) model.

constraining modified DE models) requires to calibrate the relation using appropriate N -body simulations.

ACKNOWLEDGEMENTS

We acknowledge the anonymous referee for the useful comments, which significantly contributed to the improvement of the quality of the manuscript. TR is grateful for the support of his supervisors, Andrea Lapi and Matteo Viel. FM and LM acknowledge the grants Agenzia Spaziale Italiana (ASI) number I/023/12/0, ASI-Istituto Nazionale di Astrofisica (INAF) number 2018-23-HH.0 and Progetti di Ricerca di Interesse Nazionale del Ministero dell’Istruzione dell’Università e della Ricerca (PRIN MIUR) 2015 ‘Cosmology and Fundamental Physics: illuminating the Dark Universe with Euclid’.

REFERENCES

- Achitouv I., 2017, *Phys. Rev. D*, 96, 083506
 Achitouv I., Neyrinck M., Paranjape A., 2015, *MNRAS*, 451, 3964
 Adermann E., Elahi P. J., Lewis G. F., Power C., 2017, *MNRAS*, 468, 3381
 Adermann E., Elahi P. J., Lewis G. F., Power C., 2018, *MNRAS*, 479, 4861
 Alcock C., Paczynski B., 1979, *Nature*, 281, 358
 Baldi M., 2012, *MNRAS*, 422, 1028
 Baldi M., Pettorino V., Robbers G., Springel V., 2010, *MNRAS*, 403, 1684
 Barreira A., Cautun M., Li B., Baugh C. M., Pascoli S., 2015, *J. Cosmol. Astropart. Phys.*, 8, 028
 Bernardeau F., 1994, *ApJ*, 427, 51
 Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, *ApJ*, 379, 440
 Bos E. G. P., van de Weygaert R., Dolag K., Pettorino V., 2012, *MNRAS*, 426, 440
 Cai Y.-C., Padilla N., Li B., 2015, *MNRAS*, 451, 1036

Cai Y.-C., Taylor A., Peacock J. A., Padilla N., 2016, *MNRAS*, 462, 2465
 Clampitt J., Jain B., 2015, *MNRAS*, 454, 3357
 Clampitt J., Cai Y.-C., Li B., 2013, *MNRAS*, 431, 749
 Colberg J. M. et al., 2008, *MNRAS*, 387, 933
 Davies C. T., Cautun M., Li B., 2018, *MNRAS*, 480, L101
 Dekel A., Lahav O., 1999, *ApJ*, 520, 24
 Di Porto C. et al., 2016, *A&A*, 594, A62
 Elyiv A., Marulli F., Pollina G., Baldi M., Branchini E., Cimatti A., Moscardini L., 2015, *MNRAS*, 448, 642
 Falck B., Koyama K., Zhao G.-B., Cautun M., 2018, *MNRAS*, 475, 3262
 Hamaus N., Sutter P. M., Lavaux G., Wandelt B. D., 2014, *J. Cosmol. Astropart. Phys.*, 12, 013
 Hamaus N., Sutter P. M., Lavaux G., Wandelt B. D., 2015, *J. Cosmol. Astropart. Phys.*, 11, 036
 Hamaus N., Pisani A., Sutter P. M., Lavaux G., Escoffier S., Wandelt B. D., Weller J., 2016, *Phys. Rev. Lett.*, 117, 091302
 Hamaus N., Cousinou M.-C., Pisani A., Aubert M., Escoffier S., Weller J., 2017, *J. Cosmol. Astropart. Phys.*, 7, 014
 Hawken A. J. et al., 2017, *A&A*, 607, A54
 Icke V., 1984, *MNRAS*, 206, IP
 Jennings E., Li Y., Hu W., 2013, *MNRAS*, 434, 2167
 Komatsu E. et al., 2011, *ApJS*, 192, 18
 Kreisch C. D., Pisani A., Carbone C., Liu J., Hawken A. J., Massara E., Spergel D. N., Wandelt B. D., 2019, *AAS*, 51, 203.06
 Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64
 Lavaux G., Wandelt B. D., 2012, *ApJ*, 754, 109
 Lee J., Park D., 2009, *ApJ*, 696, L10
 Maartens R., Clarkson C., Chen S., 2018, *J. Cosmol. Astropart. Phys.*, 2018, 013
 Marulli F. et al., 2013, *A&A*, 557, A17
 Marulli F., Veropalumbo A., Moresco M., 2016, *Astron. Comput.*, 14, 35
 Marulli F. et al., 2018, *A&A*, 620, A1
 Massara E., Villaescusa-Navarro F., Viel M., Sutter P. M., 2015, *J. Cosmol. Astropart. Phys.*, 11, 018
 Melchior P., Sutter P. M., Sheldon E. S., Krause E., Wandelt B. D., 2014, *MNRAS*, 440, 2922
 Micheletti D. et al., 2014, *A&A*, 570, A106
 Nadathur S., Hotchkiss S., 2015a, *MNRAS*, 454, 2228
 Nadathur S., Hotchkiss S., 2015b, *MNRAS*, 454, 889
 Nadathur S., Percival W. J., 2019, *MNRAS*, 483, 3472
 Nadathur S., Carter P., Percival W. J., 2019, *MNRAS*, 482, 2459
 Neyrinck M. C., 2008, *MNRAS*, 386, 2101
 Norberg P., Baugh C. M., Gaztañaga E., Croton D. J., 2009, *MNRAS*, 396, 19
 Pénin A., Umeh O., Santos M. G., 2018, *MNRAS*, 473, 4297
 Pisani A., Sutter P. M., Hamaus N., Alizadeh E., Biswas R., Wandelt B. D., Hirata C. M., 2015, *Phys. Rev. D*, 92, 083531
 Platen E., van de Weygaert R., Jones B. J. T., 2007, *MNRAS*, 380, 551
 Pollina G., Baldi M., Marulli F., Moscardini L., 2016, *MNRAS*, 455, 3075
 Pollina G., Hamaus N., Dolag K., Weller J., Baldi M., Moscardini L., 2017, *MNRAS*, 469, 787
 Pollina G. et al., 2018, *MNRAS*, 487, 2836
 Press W. H., Schechter P., 1974, *ApJ*, 193, 437
 Pycke J.-R., Russell E., 2016, *ApJ*, 821, 110
 Ronconi T., Marulli F., 2017, *A&A*, 607, A24
 Ronconi T., Contarini S., Marulli F., Baldi M., Moscardini L., 2019, preprint (arXiv:1902.04585)
 Sahlén M., 2019, *Phys. Rev. D*, 99, 063525
 Sahlén M., Silk J., 2018, *Phys. Rev. D*, 97, 103504
 Sánchez C. et al., 2017, *MNRAS*, 465, 746
 Sheth R. K., van de Weygaert R., 2004, *MNRAS*, 350, 517
 Sutter P. M., Lavaux G., Wandelt B. D., Weinberg D. H., 2012, *ApJ*, 761, 187
 Sutter P. M., Lavaux G., Hamaus N., Wandelt B. D., Weinberg D. H., Warren M. S., 2014, *MNRAS*, 442, 462
 Sutter P. M. et al., 2015, *Astron. Comput.*, 9, 1
 Szapudi I. et al., 2015, *MNRAS*, 450, 288
 Tikhonov A. V., Karachentsev I. D., 2006, *ApJ*, 653, 969

Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, *ApJ*, 724, 878
 van de Weygaert R., van Kampen E., 1993, *MNRAS*, 263, 481
 Verza G., Pisani A., Carbone C., Hamaus N., Guzzo L., 2019, preprint (arXiv:1906.00409)
 Wojtak R., Powell D., Abel T., 2016, *MNRAS*, 458, 4431
 Zivick P., Sutter P. M., Wandelt B. D., Li B., Lam T. Y., 2015, *MNRAS*, 451, 4215

APPENDIX A: MEASURING THE LINEAR BIAS

In this appendix, we describe the methods employed in this work to estimate the large-scale effective linear bias of the tracers used to identify the voids. We followed the same prescriptions as in Marulli et al. (2013, 2018), exploiting the 2PCF of the DM haloes of our simulated catalogues, and performing a Bayesian statistical analysis to infer the effective bias, b_{eff} .

The angle-averaged 2PCF $\hat{\xi}(r)$ is computed using the Landy & Szalay (1993) estimator:

$$\hat{\xi}(r) = \frac{N_{\text{RR}} \text{OO}(r)}{N_{\text{OO}} \text{RR}(r)} - 2 \frac{N_{\text{RR}} \text{OR}(r)}{N_{\text{RR}} \text{RR}(r)} + 1, \quad (\text{A1})$$

where $\text{OO}(r)$, $\text{RR}(r)$, and $\text{OR}(r)$ are the binned numbers of object–object, random–random, and object–random pairs with distance $r \pm \Delta r$, while $N_{\text{OO}} = N_{\text{O}}(N_{\text{O}} - 1)/2$, $N_{\text{RR}} = N_{\text{R}}(N_{\text{R}} - 1)/2$, and $N_{\text{OR}} = N_{\text{O}}N_{\text{R}}$ are the total numbers of object–object, random–random, and object–random pairs in the sample, respectively, and N_{O} and N_{R} are the total number of objects and random objects, respectively. The Landy & Szalay (1993) estimator provides an unbiased estimate of the 2PCF in the limit $N_{\text{R}} \rightarrow \infty$, with minimum variance.

We computed the covariance matrix $C_{i,j}$, which measures the variance and correlation between the different bins of the 2PCF, with the Bootstrap method, dividing the original catalogues in 125 sub-catalogues, and constructing 100 realizations by resampling

Table A1. Table of the values b_{slope} and c_{off} obtained by the linear fitting of δ_{DM} and δ_{halo} , as shown in Fig. A1. Here are presented the results for all the halo catalogues (with five mass cuts) at redshifts $z = 0, 0.55$, and 1. We report also the values of \tilde{b}_{punct} computed with equation (B2).

	b_{slope}	c_{off}	\tilde{b}_{punct}
$M_{\text{min}} (M_{\odot} h^{-1})$ $z = 0.00$			
2×10^{12}	1.405 ± 0.009	0.017 ± 0.004	1.373 ± 0.011
2.5×10^{12}	1.415 ± 0.008	0.015 ± 0.003	1.386 ± 0.010
5×10^{12}	1.521 ± 0.009	0.016 ± 0.003	1.486 ± 0.011
7.5×10^{12}	1.562 ± 0.007	0.002 ± 0.002	1.561 ± 0.009
10^{13}	1.661 ± 0.006	0.005 ± 0.002	1.650 ± 0.008
$M_{\text{min}} (M_{\odot} h^{-1})$ $z = 0.55$			
2×10^{12}	1.685 ± 0.009	-0.001 ± 0.003	1.687 ± 0.011
2.5×10^{12}	1.789 ± 0.010	0.004 ± 0.003	1.719 ± 0.012
5×10^{12}	1.901 ± 0.012	-0.001 ± 0.003	1.903 ± 0.014
7.5×10^{12}	2.032 ± 0.019	-0.002 ± 0.004	2.037 ± 0.023
10^{13}	2.163 ± 0.013	-0.001 ± 0.003	2.167 ± 0.016
$M_{\text{min}} (M_{\odot} h^{-1})$ $z = 1.00$			
2×10^{12}	2.086 ± 0.010	-0.003 ± 0.002	2.095 ± 0.011
2.5×10^{12}	2.129 ± 0.011	-0.005 ± 0.002	2.145 ± 0.012
5×10^{12}	2.433 ± 0.019	-0.005 ± 0.003	2.452 ± 0.021
7.5×10^{12}	2.690 ± 0.016	-0.005 ± 0.002	2.711 ± 0.017
10^{13}	2.806 ± 0.025	-0.006 ± 0.003	2.830 ± 0.028

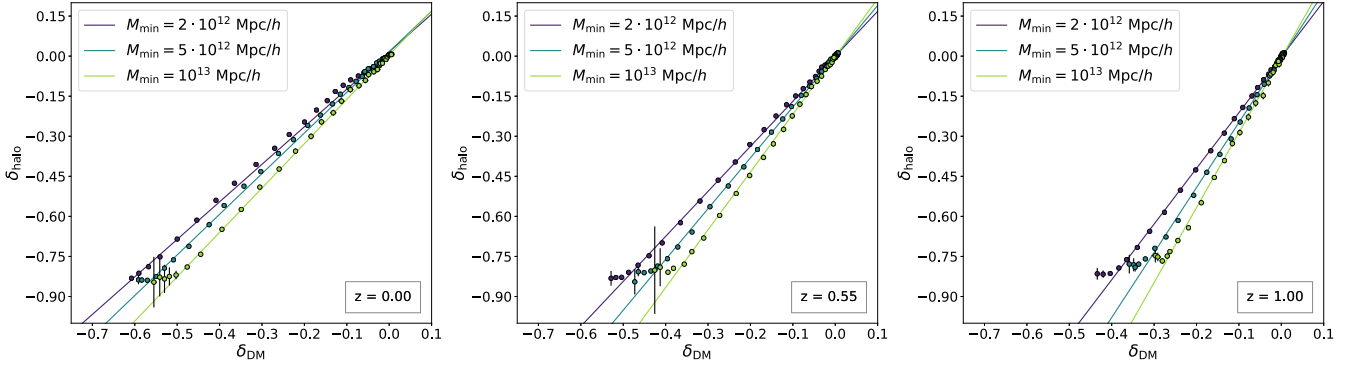


Figure A1. The ratio of the stacked density profiles shown in Fig. 1, that is $\delta_{\text{halo}}^{\text{NL}}$ as a function $\delta_{\text{v,DM}}^{\text{NL}}$, at redshifts $z = 0$ (left), $z = 0.55$ (centre), and $z = 1$ (right). Different colours correspond to the halo catalogues with $M_{\text{min}} = 2 \times 10^{12} M_{\odot} h^{-1}$ (in violet), $M_{\text{min}} = 5 \times 10^{12} M_{\odot} h^{-1}$ (in blue), and $M_{\text{min}} = 10^{13} M_{\odot} h^{-1}$ (in green). The black error bars represent 1σ uncertainties. As expected, the slope of the fit becomes steeper with higher mass cuts (thus, the value of bias inside voids).

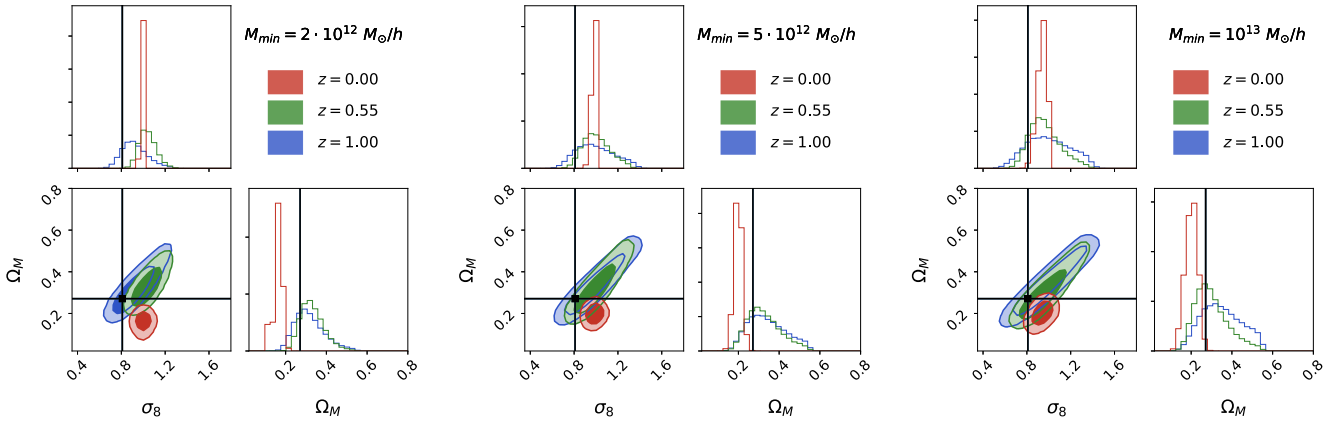


Figure A2. 68 per cent and 95 per cent contours in the σ_8 – Ω_M plane, for the halo catalogues with $M_{\text{min}} = 2 \times 10^{12} M_{\odot} h^{-1}$ (left), $5 \times 10^{12} M_{\odot} h^{-1}$ (centre), and $10^{13} M_{\odot} h^{-1}$ (right), obtained by re-parametrizing the Vdn model directly with b_{eff} , thus without converting this value by means of equation (13). The colour of ellipses corresponds to different redshifts: red for $z = 0$, green for $z = 0.55$, and blue for $z = 1$. The prior distributions are uniform for σ_8 and Ω_M , and Gaussian for b_{eff} . The histograms (top and bottom right panels) show the marginalized posterior distributions of σ_8 and Ω_M , respectively. The black lines represent the true *WMAP7* values ($\sigma_8 = 0.809$ and $\Omega_M = 0.2711$).

from the sub-catalogues, with replacement. We constructed the random catalogue by extracting the object coordinates randomly, preserving the same 3D coverage and the same geometry of the initial catalogue. In particular, we build the random catalogue to be four times larger than the DM halo sample, since this proportion allows to have sufficiently small Poissonian errors in the DR counts, compared to the errors in DD. We also performed tests with different sizes of the random catalogue, finding consistent results.

The covariance matrix is defined as follows:

$$C_{i,j} = \mathcal{F} \sum_{k=1}^{N_R} (\xi_i^k - \bar{\xi}_i)(\xi_j^k - \bar{\xi}_j), \quad (\text{A2})$$

where the subscripts i and j run over the 2PCF bins, while k refers to the 2PCF of the k th of N_R catalogue realizations, and $\bar{\xi}$ is the mean 2PCF of the N_R samples. \mathcal{F} is the normalization factor, which takes into account the fact that the N_R realizations might not be independent (Norberg et al. 2009), and is $\mathcal{F} = 1/(N_R - 1)$ in the case of the Bootstrap method.

Finally, we performed a full MCMC analysis of the 2PCF, using a Gaussian likelihood function \mathcal{L} , defined as:

$$-2 \ln \mathcal{L} = \sum_{i=1}^N \sum_{j=1}^N (\xi_i^d - \xi_i^m) C_{i,j}^{-1} (\xi_j^d - \xi_j^m), \quad (\text{A3})$$

where $C_{i,j}^{-1}$ is the inverse of the covariance matrix, N is the number of comoving separation bins at which the 2PCF is estimated, and the superscripts d and m stand for data and model, respectively. The 2PCF model, $\xi^m(r)$, is computed as follows:

$$\xi_m(r) = b_{\text{eff}}^2 \xi_{\text{DM}}(r), \quad (\text{A4})$$

where $\xi_{\text{DM}}(r)$ is the DM 2PCF, which is estimated by Fourier transforming the power spectrum, $P_{\text{DM}}(k)$, computed with the CODE FOR ANISOTROPIES IN THE MICROWAVE BACKGROUND (CAMB, <http://camb.info>). An accurate estimate of the effective bias parameter, b_{eff} , and its uncertainty are assessed by sampling its posterior distribution.

Fig. 7 shows the results of this analysis. The data points are the square root of ratio between the tracer and matter 2PCFs:

$$b = \sqrt{\xi_{\text{tr}}/\xi_{\text{m}}}, \quad (\text{A5})$$

while the dashed red lines show the best-fitting values and uncertainties of b_{eff} , estimated from the median and quartiles of the posterior distribution sampled with the MCMC analysis.

We compared these values to the theoretical effective bias of DM haloes, computed as follows:

$$b_{\text{eff}}(z) = \frac{\int_{M_{\min}}^{M_{\max}} dM b(M, z) \Phi(M, z)}{\int_{M_{\min}}^{M_{\max}} dM \Phi(M, z)}, \quad (\text{A6})$$

where $\Phi(M, z)$ is the halo mass function of the catalogue, and M_{\min} and M_{\max} are the lowest and largest masses in the sample, respectively. To compute the linear bias $b(M, z)$, we relied on the theoretical model developed by Tinker et al. (2010).

APPENDIX B: TESTING DIFFERENT BIAS ESTIMATES INSIDE VOIDS

Pollina et al. (2018) found a linear relation between the spherically averaged density profiles of biased tracers inside voids, $\delta_{v, \text{tr}}^{\text{NL}}$, and the underlying DM, $\delta_{v, \text{DM}}^{\text{NL}}$. The slope of this relation provides an alternative estimate of the bias of void tracers. We repeated the analysis of Pollina et al. (2018) finding consistent results, as shown in Fig. A1, that reports this relation measured in three DM halo catalogues at three different redshifts.

Even if the offset of the linear fit is small, it can be taken into account to recover a better estimate of the bias inside voids. In particular, to obtain a value comparable with b_{punct} (see equation 12), we can recover the bias at a distance of one effective radius, R_{eff} , from void centres, that is where $\delta_{v, \text{tr}}^{\text{NL}} = -0.7$ (thus the requested threshold we chose for the cleaning procedure). Specifically, from

the fitting of the data we obtained:

$$\delta_{v, \text{tr}}^{\text{NL}} = b_{\text{slope}} \delta_{v, \text{DM}}^{\text{NL}} + c_{\text{off}}, \quad (\text{B1})$$

and we computed equation (B1) at $R = R_{\text{eff}}$, imposing $\delta_{v, \text{tr}}^{\text{NL}}(R_{\text{eff}}) = -0.7$. Then, dividing both sides by $\delta_{v, \text{tr}}^{\text{NL}}(R_{\text{eff}})$, we can derive a bias value equivalent to b_{punct} as follows:

$$\tilde{b}_{\text{punct}} \equiv \frac{b_{\text{slope}}}{1 - \frac{c_{\text{off}}}{\delta_{v, \text{tr}}^{\text{NL}}(R_{\text{eff}})}}. \quad (\text{B2})$$

In Table A1, we report the results obtained for all the analysed halo catalogues. As expected, we found similar values for \tilde{b}_{punct} and b_{punct} . Nevertheless, the method described here is not particularly accurate for the following reasons (see Fig. 2 as a reference):

- (i) First, using the stacked profiles of voids we cannot distinguish a possible variation of the bias as a function of void radii.
- (ii) Second, the slope of the fit depends on the radial extension of the profiles: the wider the regions are embedded, the more the bias computed will tend to the one of the overdensities. There is not a preferential value for the maximum radius of the profiles.
- (iii) Third, a linear fit to the relation between δ_{halo} and δ_{DM} is not accurate enough in the full range of δ_{DM} .

This method can be refined considering different bins of void radii and using different linear fits for each of them, to account for the variation of the bias as a function of R_{eff} .

APPENDIX C: TESTING SYSTEMATICS CAUSED BY UNCERTAINTIES ON THE CALIBRATED BIAS RELATION

Here, we first test the effect of using the value of b_{eff} instead of b_{rel} to recover the cosmological parameters. In particular, we repeat the MCMC analysis of the measured size functions performed in Section 4.2 employing a wrong theoretical model, that is a Vdn model re-scaled with the linear bias inferred from the tracer large-scale 2PCF. As demonstrated in Fig. 5, the model obtained using the tracers bias b_{eff} cannot fit properly the measured void abundances, unless this value is previously converted in b_{rel} by means of the relation in equation (13). As shown in Fig. A2, the contour levels achieved with the effective bias are on average smaller with respect to the ones presented in Fig. 6. In fact, the uncertainties associated to the theoretical model re-parametrized directly with b_{eff} are smaller, since the errors of A_{rel} and B_{rel} are not included in the model. As expected, the contour levels obtained using the wrong bias value tend to shift from the real values of σ_8 and Ω_M , especially for low redshifts and mass cuts. Indeed, in these cases the values of b_{eff} and b_{rel} are significantly different from each other, whereas at higher redshifts and mass cuts they tend to be more similar, as showed in Fig. 3.

We secondly assess the systematic errors on the cosmological constraints caused by uncertainties in the estimation of the coefficients of the conversion relation, calibrated in Section 4.1. This is particularly useful in the perspective of a future application on real surveys. To propagate a possible systematic error on the equation (13) to the final cosmological constraints, we repeated the MCMC analysis described in Section 4.2 assuming different values for the coefficient A_{rel} and B_{rel} . In particular, to test the cases with the major discrepancy from the calibrated relation, we increased or decreased both the parameters by 1σ , where 1σ is the uncertainty derived by the weighted fit of the data in Fig. 4. Specifically, we set $A'_{\text{rel}} = 0.420 + 0.010$ and $B'_{\text{rel}} = 0.854 + 0.007$ in the first

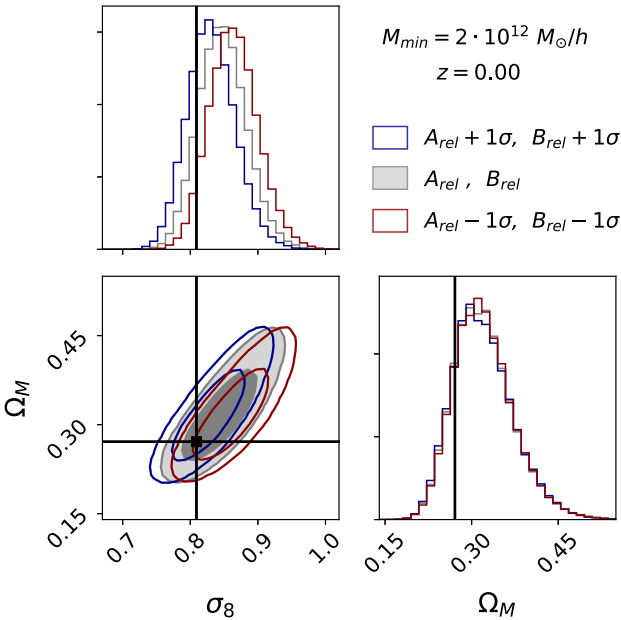


Figure B1. 68 per cent and 95 per cent contours in the σ_8 – Ω_M plane, for the halo catalogue with $M_{\min} = 2 \times 10^{12} M_{\odot} h^{-1}$ at $z = 0$. The histograms (top and bottom right panels) show the posterior distributions of σ_8 and Ω_M , respectively. The grey filled contours represent the confidence levels obtained using equation (13), while the blue and red contours indicate the results obtained by converting the value of b_{eff} shifting both the values of A_{rel} and B_{rel} by $+1\sigma$ and -1σ , respectively. The black lines represent the true WMAP7 values ($\sigma_8 = 0.809$ and $\Omega_M = 0.2711$).

case, whereas $A''_{\text{rel}} = 0.420 - 0.010$ and $B''_{\text{rel}} = 0.854 - 0.007$ in the second case. In Fig. B1, we report the results for the catalogue with $M_{\text{min}} = 2 \times 10^{12} M_{\odot} h^{-1}$ at $z = 0$. As shown in this figure, the real values of σ_8 and Ω_M are within the 68 per cent confidence levels obtained in both cases. Moreover, the posterior distribution of Ω_M is almost unchanged, while σ_8 results shifted towards greater values using a conversion relation with A''_{rel} and B''_{rel} and towards lower values for the case with A'_{rel} and B'_{rel} . We obtained the same results also for the catalogue with higher redshift and mass selections. The larger is the tracer bias, the larger is the discrepancy of the modified relation from the one calibrated in equation (13). Indeed, shifting both the values of A_{rel} and B_{rel} by $+1\sigma$ and -1σ , the resulting linear equations tend to move even further away from the calibrated relation with b_{eff} . This causes a systematic error that has more impact on the theoretical size functions associated to the catalogues with higher b_{eff} . Nevertheless, we verified that even in these cases the constraints are still consistent with the real values of σ_8 and Ω_M . We can conclude that, even with a systematic error of $\pm 1\sigma$ on the values of the coefficients in the calibrated relation, the void size function still provides reliable cosmological constraints.

APPENDIX D: COMBINING SAMPLES AT DIFFERENT REDSHIFTS

In this appendix, we test the constraining power given by the combination of the posterior probabilities obtained performing the Bayesian statistical analysis of the measured void size functions for a DM halo catalogue at different redshifts. Thanks to the redshift dependence of the degeneracy directions, it is possible to derive tighter constraints on the values achieved for σ_8 and Ω_M .

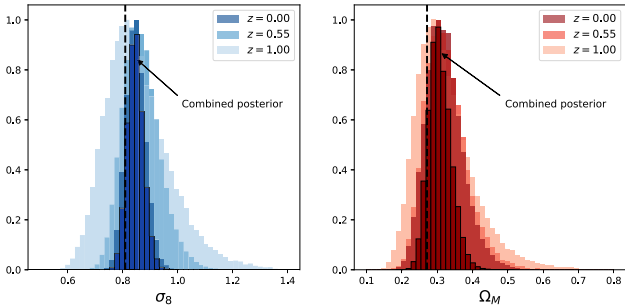


Figure D1. Normalized posterior probabilities of σ_8 (left) and Ω_M (right) computed for the halo catalogue with $M_{\text{min}} = 2 \times 10^{12} M_{\odot} h^{-1}$ at redshift $z = 0, 0.55$, and 1 . The histograms with black outlines represent the combined distributions achieved by multiplying all the posterior probabilities relative to different redshifts. The black dashed lines indicate the true *WMAP7* values ($\sigma_8 = 0.809$ and $\Omega_M = 0.2711$).

Table D1. Mean and standard deviation of the posterior distributions for the parameters σ_8 and Ω_M , computed from the Bayesian statistical analysis of the measured void size functions for the DM halo catalogues with $M_{\text{min}} = 2 \times 10^{12}, 5 \times 10^{12}$, and $10^{13} M_{\odot} h^{-1}$ at $z = 0, 0.55$, and 1 . The last line of each table reports the results obtained by combining the posterior distributions at the three different redshifts.

	σ_8		Ω_M	
	Mean	St. dev.	Mean	St. dev.
$M_{\text{min}} = 2 \times 10^{12} M_{\odot} h^{-1}$				
$z = 0.00$	0.848	0.036	0.321	0.052
$z = 0.55$	0.868	0.068	0.325	0.059
$z = 1.00$	0.856	0.118	0.322	0.081
Combined	0.846	0.030	0.308	0.032
$M_{\text{min}} = 5 \times 10^{12} M_{\odot} h^{-1}$				
$z = 0.00$	0.853	0.051	0.295	0.059
$z = 0.55$	0.901	0.124	0.333	0.093
$z = 1.00$	0.989	0.204	0.360	0.116
Combined	0.856	0.045	0.293	0.041
$M_{\text{min}} = 1 \times 10^{13} M_{\odot} h^{-1}$				
$z = 0.00$	0.802	0.062	0.228	0.054
$z = 0.55$	0.910	0.147	0.310	0.091
$z = 1.00$	1.047	0.247	0.378	0.125
Combined	0.822	0.053	0.256	0.040

Despite our samples cannot be considered completely independent, we multiply the posterior probabilities at different redshifts as if they were achieved from independent data, in order to reproduce the results that would be obtained from separate redshift shells in real surveys. We show in Fig. D1, the results for the halo catalogue with $M_{\text{min}} = 2 \times 10^{12} M_{\odot} h^{-1}$ obtained by multiplying the posterior distributions for the parameters at $z = 0, 0.55$, and 1 . Table D1 reports the mean values and the standard deviations of the posterior distributions of σ_8 and Ω_M at these redshifts also for the catalogues with $M_{\text{min}} = 5 \times 10^{12}$ and $10^{13} M_{\odot} h^{-1}$, together with the analogous quantities obtained for the combined posterior probability. As expected, by joining the information at different redshifts, we can achieve more precise constraints on the cosmological parameters, as shown by the decreasing of the width of the combined posterior distributions.

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