Euclid: Cosmological forecasts from the void size function


(Affiliations can be found after the references)

Received 23 May 2022 / Accepted 10 August 2022

ABSTRACT

The Euclid mission – with its spectroscopic galaxy survey covering a sky area over 15 000 deg² in the redshift range 0.9 < z < 1.8 – will provide a sample of tens of thousands of cosmic voids. This paper thoroughly explores for the first time the constraining power of the void size function on the properties of dark energy (DE) from a survey mock catalogue, the official Euclid Flagship simulation. We identified voids in the Flagship volume-light cone, which closely matches the features of the upcoming Euclid spectroscopic data set. We modelled the void size function considering a state-of-the art methodology: we relied on the volume-conserving (Vdn) model, a modification of the popular Sheth & van de Weygaert model extended by means of a linear function of the large-scale galaxy bias. We found an excellent agreement between model predictions and measured mock void number counts. We computed updated forecasts for the Euclid mission on DE from the void size function and provided reliable void number estimates to serve as a basis for further forecasts of cosmological applications using voids. We analysed two different cosmological models for DE: the first described by a constant DE equation of state parameter, w, and the second by a dynamic equation of state with coefficients w₀ and wₑ. We forecast 1σ errors on w lower than 10% and we estimated an expected figure of merit (FoM) for the dynamical DE scenario FoM_{w_{0}=17} when considering only the neutrino mass as additional free parameter of the model. The analysis is based on conservative assumptions to ensure full robustness, and is a pathfinder for future enhancements of the technique. Our results showcase the impressive constraining power of the void size function from the Euclid spectroscopic sample, both as a stand-alone probe, and to be combined with other Euclid cosmological probes.

Key words. dark energy – cosmology: theory – galaxies: statistics – catalogs – surveys – methods: data analysis

1. Introduction

Cosmic voids are vast under-dense regions filling most of the volume of the present-day Universe. With sizes up to hundreds of megaparsecs (Gregory & Thompson 1978; Tikhonov & Karachentsev 2006; Thompson & Gregory 2011; Szapudi et al. 2015) they are the largest observable structures in the cosmic web (Zeldovich et al. 1982; Bond et al. 1996) – the pattern arising in the galaxy distribution. Voids constitute a unique cosmological probe: their interiors, spanning a large range of scales and featuring low matter density, make them particularly suited to study dark energy (DE) and modified gravity (Lee & Park 2009; Biswas et al. 2010; Li & Efstathiou 2012; Clar DE 2013; Spolyar et al. 2013; Cai et al. 2015; Pisani et al. 2015a; Pollina et al. 2016; Zivick et al. 2016; Achitouv et al. 2016; Falck et al. 2018; Sahlén et al. 2018; Palles et al. 2019; Perico et al. 2019; Verza et al. 2019; Contarini et al. 2021), as well as massive neutrinos (Massara et al. 2015; Barber{	extgreek{e}j} & Dalal 2016; Kreisch et al. 2019, 2022; Sahlén 2019; Schuster et al. 2019), primordial non-Gaussianity (Chan et al. 2019), and physics beyond the standard model (Peebles 2001; Reed et al. 2015; Yang et al. 2015; Baldi & Villaescusa-Navarro 2018). Cosmic voids are becoming an effective and competitive new probe of cosmology thanks to the advent of current and upcoming sky surveys such as 6dF
Galaxy Survey (6dFGS, Jones et al. 2009), VIMOS Public Extragalactic Redshift Survey (VIPERS, Guzzo et al. 2014), SDSS-III’s Baryon Oscillation Spectroscopic Survey (BOSS, Alam et al. 2017) and Extended Baryon Oscillation Spectroscopic Survey (eBOSS, Dawson et al. 2016) from the Sloan Digital Sky Survey (SDSS, Blanton et al. 2017), Dark Energy Survey (DES, Dark Energy Survey Collaboration 2016), Dark Energy Spectroscopic Instrument (DESI, DESI Collaboration 2016), Prime Focus Spectrograph (PFS, Tamura et al. 2016), the Roman Space Telescope (Spergel et al. 2015), Spectro-Photometer for the History of the Universe and Ices Explorer (SPHEREx, Doré et al. 2018), and Large Synoptic Survey Telescope (LSST, Ivezić et al. 2019). Studying voids requires redshift surveys of very large volume, deep enough in the red band to measure a huge number of redshifts also for low-mass galaxies, and to map in detail significant contiguous fractions of the observable Universe. The Euclid survey, expected to sample the sky over 15,000 deg², will provide a unique opportunity to capitalise on cosmic voids, to leverage on measurements of the galaxy distribution at large scales and to improve our knowledge on cosmology and fundamental physics. Voids hold the keys to shed light on some of today’s open problems in cosmology (Pisani et al. 2019, and references therein).


In this paper we consider the void size function, which describes the number density of voids as a function of their size. Over the last two decades, studies of the hierarchical evolution of the void population in the excursion-set framework have allowed the construction of a theoretical void size function model built from first principles, the so-called Sheth & van de Weygaert model (Sheth & van de Weygaert 2004), later extended by Jennings et al. (2013). The void size function and its link to voids detected in galaxy surveys have been explored in depth with cosmological simulations (Furlanetto & Piran 2006; Platen et al. 2007; Paranjape et al. 2012; Jennings et al. 2013; Pisani et al. 2015a; Ronconi et al. 2019; Contarini et al. 2019; Verza et al. 2019; Contarini et al. 2021) and recently this statistic has proved to be a promising tool to constrain cosmology (Pisani et al. 2015a; Sahlén 2019; Contarini et al. 2019; Verza et al. 2019; Kreisch et al. 2022). The void size function has already been measured in surveys (see e.g. Nadathur 2016; Mao et al. 2017b; Aubert et al. 2022; Hamaus et al. 2020), and used for extreme-value statistics cosmology constraints (Sahlén et al. 2016), but the void size function as a stand-alone probe has not yet been used to derive cosmological constraints.

In this work we focus on the power of the void size function from the Euclid survey to constrain cosmological parameters. This study relies on the largest Euclid-like light-cone, the Flagship simulation (Potter et al. 2017). The paper belongs to a series of companion papers investigating the scientific return that can be expected from voids observed by the Euclid mission. It aims at measuring and theoretically modelling the void size function from the Flagship simulation, providing a state-of-the-art forecast for void numbers to be expected from the Euclid survey. Our model allows us to estimate the constraining power of the void size function on the DE equation of state while also varying the total matter density of the Universe and the total mass of neutrinos. This analysis is focused on voids found in the spectroscopic galaxy distribution, for which the identification of voids is particularly accurate and reliable. We note that we leave for future work the measurement of the void size function in the photometric galaxy distribution from Euclid, for which the data treatment greatly differs from the spectroscopic one (see e.g. Pollina et al. 2019; Bonici et al. 2022).

The paper is organised as follows: in Sect. 2 we introduce the Flagship simulation, and describe the void finder and the cleaning algorithm used to obtain the void catalogue; in Sect. 3 we present the theoretical model of the void size function (Sect. 3.1), describe how to self-consistently align the measured void catalogue with the theoretical description (Sect. 3.2), discuss the Bayesian statistical analysis used to perform the cosmological forecasts (Sect. 3.3), and finally introduce the cosmological models considered in this work (Sect. 3.4). In Sect. 4 we fit the theoretical model to the measured void size function in the Flagship simulation (Sect. 4.1) to obtain constraints on the DE equation of state and the remaining considered cosmological parameters, for different adopted approaches (Sect. 4.2); we conclude giving a discussion and a summary of our results in Sect. 5.

2. Galaxy and void catalogues

We now introduce the main tools for our work: the simulation and the void catalogues. This section also includes a brief description of the void finder and of the void catalogue preparation.

2.1. Flagship simulation

In this work we employed the Euclid Flagship mock galaxy catalogue1 (Castander et al., in prep.). This catalogue was created running a simulation of two trillion dark matter particles in a periodic box of \( L = 3780 \, h^{-1} \) Mpc per side (Potter et al. 2017), with a flat \( \Lambda \)-cold dark matter (LCDM) cosmology characterised by the parameters \( \Omega_m = 0.319, \Omega_b = 0.049, \Omega_{\Lambda} = 0.681, \sigma_8 = 0.83, n_s = 0.96 \) and \( h = 0.67 \), as obtained by Planck in 2018 (Planck Collaboration VI 2020). The simulation box was converted into a light-cone and the dark matter haloes have been identified using the Rockstar halo finder (Behroozi et al. 2013). These haloes were populated with central and satellite galaxies.

---

1 Version 1.8.4.
using a halo occupation distribution (HOD) method, to reproduce all the observables relevant for Euclid’s main cosmological probes. Specifically, the HOD algorithm was calibrated exploiting several local observational constraints, using for instance the local luminosity function for the faintest galaxies (Blanton et al. 2003, 2005) and the galaxy clustering as a function of luminosity and colour (Zehavi et al. 2011). This galaxy sample is composed of more than two billion objects and presents a cut at magnitude $H < 26$ or on the $\text{H}\alpha$ flux $f_{\text{H}\alpha} > 2 \times 10^{-16} \text{ ergs s}^{-1} \text{ cm}^{-2}$, which mimics the observation range expected for Euclid. To match the completeness and the spectroscopic performance expected for the Euclid survey, we uniformly downsampled the galaxy catalogue to consider only 60% of the galaxies originally included in it. Furthermore we assumed our galaxy sample to have a purity of 100% and associated a Gaussian error of $\sigma_z = 0.001$ to the redshift of each galaxy (Euclid Collaboration 2020). The full catalogue spans a large redshift range, up to $z = 2.3$, and covers one octant of the sky (close to 5157 deg$^2$).

The Euclid satellite will observe 15,000 deg$^2$ of the sky with patches that extend up to 6000 deg$^2$. The total area covered by the satellite will be significantly larger than the available Flagship area. By rescaling it is possible to compute the full predicting power from Euclid. The larger Euclid survey coverage will allow us to increase statistics, reducing the size of the error bar in particular for the high radius end of the void size function, and to better account for super-sample covariance. On the other hand, the Euclid survey is expected to have a less regular pattern than the Flagship box, which might impact the void statistics. Conversely to galaxies, voids are strongly sensitive to survey area specifics because of their extended nature: while contiguous regions are a great advantage for void search, as they provide larger voids, void statistics can be reduced in the case of patchy survey coverage, because voids touching survey edges must be excluded from the analysis. While the interplay between these different effects may have a role in final constraints, we do not expect this role to significantly impact the precision of constraints resulting from Euclid.

We focused our analysis on the expected sub-sample corresponding to spectroscopic data, selecting galaxies from redshift 0.9 to 1.8. We obtained a resulting mock catalogue composed of about $6.5 \times 10^9$ galaxies, having the spatial distribution of a shell of sphere octant.

### 2.2. Void finding and catalogue preparation

We identified cosmic voids in the Flagship light-cone with the public Void IDentification and Examination toolkit\(^2\) (VIDE, Sutter et al. 2015), a parameter-free watershed void finding algorithm based on the code ZOnes Bordering On Voidness (ZOBOV, Neyrinck 2008). VIDE provides a robust density field estimation via the Voronoi tessellation of tracer positions, which allows us to identify local minima and their surrounding density depressions in the tracer density field. With the watershed algorithm (Platen et al. 2007), VIDE constructs the void catalogue and provides void properties, such as the void barycentre, the effective radius, the ellipticity, etc. VIDE can be launched on any catalogue of tracers, both on simulation boxes with periodic boundary conditions and on galaxies from real surveys. It is also capable to handle a survey selection function and a mask. These features make VIDE a very flexible tool to study voids in data and simulations. VIDE has been extensively used for cosmological applications relying on voids in the past decade (see e.g. Sutter et al. 2012, 2014a; Leclercq et al. 2015; Hamaus et al. 2016, 2017, 2020; Pollina et al. 2019; Fang et al. 2019; Hawken et al. 2020).

We built void catalogues using VIDE from the galaxy sample both with real and redshift-space coordinates given by true and observed redshifts, and note that the redshift-space catalogue is identical to the one used in our companion paper, Hamaus et al. (2022). In the true redshift catalogue, the galaxy redshift corresponds to the cosmological one only, in the observed redshift catalogue it corresponds to the cosmological plus Doppler shift due to peculiar velocity.

Despite VIDE being a parameter-free algorithm, the theoretical model of the void size function requires voids with the same level of embedded underdensity, so we further processed the void catalogue. We applied therefore to both the obtained void catalogues a cleaning algorithm\(^3\) (Ronconi & Marulli 2017). The goal of this procedure is to conform observed voids with their theoretical counterpart. The main steps of this cleaning pipeline are: (i) the rejection of spurious voids, i.e. with central density too high or with radius below the spatial resolution of the tracer catalogue, (ii) the rescaling of voids to a specific radius $R_{\text{eff}}$ to match a specific spherical density contrast within the sphere, $\delta_{\text{eff}}$, in the tracer distribution, (iii) the removal of overlapped voids, i.e. voids whose distance between centres is smaller than the sum of their radii. We underline that, during the rescaling procedure, any negative value of density contrast $\delta_{\text{NL}}$ can in principle be chosen to resize underdensities, as long as the theoretical model is consistently calculated using the same threshold (see Sect. 3.1). When dealing with observed voids, the threshold can be fixed to a suitable value chosen based on survey features. We considered the following reasoning to select this value: on the one hand the more negative the threshold, the more the identified underdensities are free of contamination by Poisson noise (see also Neyrinck 2008; Cousinou et al. 2019, for a discussion on spurious voids and possible treatments) and the stronger the impact of the cosmology on the void size function; on the other hand, an excessively negative threshold entails both a low statistic and a higher uncertainty in the rescaled void radius, caused by the sparsity of galaxies tracing such extreme underdense regions. In particular, for this analysis we followed the choice of Contarini et al. (2019, 2021), selecting a threshold $\delta_{\text{NL}} = -0.7$, which ensures a good compromise on the aforementioned effects. We verified the robustness of our method by also performing the entire analysis using $\delta_{\text{NL}} = -0.6$, finding consistent results.

VIDE takes into account the presence of a survey mask, and prevents voids from including volumes outside the survey extent. We applied the mask following the simulated $\sim 5000$ deg$^2$ octant. While the actual Euclid data will be more complex (due to e.g. more elaborate survey mask and survey-related systematic effects), this methodology at least partially accounts for mask effects in our pipeline, preparing the analysis of future Euclid data. Aiming at a very conservative void selection at the edges of the survey’s footprint, we applied an additional cut to ensure the mask is not affecting the cleaning procedure: we removed all

---

\(^2\) https://bitbucket.org/cosmicvoids/vide_public

\(^3\) This algorithm is an improved version of the code developed by Ronconi & Marulli (2017) and is inserted in the free software C++/Python libraries CosmoBolognaLib V5.5 (Marulli et al. 2016), available at https://gitlab.com/federicomarulli/CosmoBolognaLib. In this version of the code, the cleaning procedure can be applied to catalogues with comoving coordinates and the void rescaling is performed by taking into account the variation of the tracer density with redshift.
voids whose centre is closer than $30 \, h^{-1} \, \text{Mpc}$ to the edge and corrected the model accordingly for the selected volume. We then pruned voids at low and high redshifts to further avoid selection effects given by redshift boundaries of the light-cone, and we divided the sample in six redshift bins. This number is found as the optimal compromise between maximising the number of redshift shells and keeping void numbers in bins high enough to avoid falling in the shot-noise dominated regime. In order to have shells with roughly the same number of cleaned voids identified by the VIDE void finder with $R > f_{\text{cut}}(z)$ MGS and of voids obtained after the cleaning procedure with $R_{\text{cf}} > f_{\text{cut}}(z)$ MGS, respectively. In the last row we show the total volume of all redshift shells, the mean MGS and $f_{\text{cut}}(z)$ values and the total void counts corresponding to the entire range of redshifts. A table with equi-spaced redshift bins is provided in Appendix A to serve as a reference for future forecast analyses needing void numbers.

<table>
<thead>
<tr>
<th>$z$ range</th>
<th>Shell volume [($h^{-1}$ Gpc)$^3$]</th>
<th>MGS [($h^{-1}$ Mpc)]</th>
<th>$f_{\text{cut}}(z)$</th>
<th>All voids</th>
<th>Voids after cleaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950–1.035</td>
<td>1.157</td>
<td>10.28</td>
<td>2.30</td>
<td>4989</td>
<td>343</td>
</tr>
<tr>
<td>1.035–1.126</td>
<td>1.329</td>
<td>11.02</td>
<td>2.24</td>
<td>4935</td>
<td>343</td>
</tr>
<tr>
<td>1.126–1.208</td>
<td>1.269</td>
<td>11.74</td>
<td>2.18</td>
<td>4232</td>
<td>342</td>
</tr>
<tr>
<td>1.208–1.318</td>
<td>1.796</td>
<td>12.63</td>
<td>2.12</td>
<td>5302</td>
<td>341</td>
</tr>
<tr>
<td>1.318–1.455</td>
<td>2.363</td>
<td>13.51</td>
<td>2.06</td>
<td>5935</td>
<td>342</td>
</tr>
<tr>
<td>1.455–1.700</td>
<td>4.490</td>
<td>14.45</td>
<td>2.00</td>
<td>8435</td>
<td>343</td>
</tr>
<tr>
<td>0.950–1.700</td>
<td>12.40</td>
<td>13.69</td>
<td>2.15</td>
<td>33828</td>
<td>2054</td>
</tr>
</tbody>
</table>

Notes. The first column represents the minimum and the maximum redshift values for each bin, while the second and the third columns provide the volume in units of ($h^{-1}$ Gpc)$^3$ corresponding to each shell of the sky octant, and the mean separation between galaxies (MGS), respectively. The fourth column reports the factor, $f_{\text{cut}}(z)$, used to select voids unaffected by the incompleteness of counts. The last two columns show the number counts of voids identified by the VIDE void finder with $R > f_{\text{cut}}(z)$ MGS and of voids obtained after the cleaning procedure with $R_{\text{cf}} > f_{\text{cut}}(z)$ MGS, respectively. In the last row we show the total volume of all redshift shells, the mean MGS and $f_{\text{cut}}(z)$ values and the total void counts corresponding to the entire range of redshifts. A table with equi-spaced redshift bins is provided in Appendix A to serve as a reference for future forecast analyses needing void numbers.

3. Theory and methods

In this section we introduce the theoretical background necessary for this work. We first discuss the model of the void size function, then we present the prescriptions applied to extend this model to voids identified in the distribution of biased tracers. We describe the Bayesian statistical analysis used to provide forecasts on the DE equation of state and on the sum of neutrino masses. Finally, we present the cosmological scenarios considered in our analysis.

3.1. Theoretical void size function

To estimate the constraining power of the void size function, i.e. the distribution function of void radii, we first need a theoretical model. The void size function model most widely used in the literature relies on the excursion-set formalism, developed within the framework of the halo mass function (Peacock & Heavens 1990; Cole 1991; Bond et al. 1991; Mo & White 1996). This model was first proposed by Sheth & van de Weygaert (2004) and extended by Jennings et al. (2013). The distribution of fluctuations that become voids, i.e. the multiplicity function, is obtained as the conditional first crossing distribution
of the matter density contrast filtered at decreasing Lagrangian radius in a double barrier problem: a fluctuation becomes a void at a radius $R_L$ if the density contrast first crosses the void formation threshold $\delta_c^L$ at $R_L$, without having crossed the threshold for collapse $\delta_c^L$ at any larger scale\(^4\) (Sheth & van de Weygaert 2004). The multiplicity function of Sheth & van de Weygaert (2004) is derived for spherical fluctuations in Lagrangian space, i.e. the initial density field linearly evolved to the epoch of interest, while the observed voids live in the fully nonlinear evolved density field in comoving coordinates, i.e. the Eulerian space. The spherical approximation allows us to easily go back and forth from Lagrangian to Eulerian space in all the computations.

The void size function probes the inner region of cosmic voids and in contrast to the collapsing case, i.e. halo formation (Monaco 1995; Sheth & Tormen 2002), the spherical approximation is accurate enough for this purpose, at least for voids of scales detectable by Euclid (Icke 1984; Verza et al. 2019). The void case is different: if the initial underdensity identifying a void is deep enough, its evolution is not marked by any specific event, and it continues its outward-directed expansion forever. It is common to consider the shell-crossing\(^5\) condition as the event that identifies the void formation (Blumenthal et al. 1992; Sheth & van de Weygaert 2004; Jennings et al. 2013), but this condition strictly depends on the initial density profile of the underdensity. For an initial density profile of the underdensity \(J\) (Jennings et al. 2013), but this condition strictly depends on the initial density profile of the underdensity.

The linear threshold for collapse is fixed at $\delta_c^L = 1.686$, according to the collapse of a spherical fluctuation. This value corresponds in an Einstein–de Sitter model to the full collapse in linear theory, when the halo virialises. The void case is different: if the initial underdensity identifying a void is deep enough, its evolution is not marked by any specific event, and it continues its outward-directed expansion forever. It is common to consider the shell-crossing\(^5\) condition as the event that identifies the void formation (Blumenthal et al. 1992; Sheth & van de Weygaert 2004; Jennings et al. 2013), but this condition strictly depends on the initial density profile of the underdensity.

The linear threshold for collapse is fixed at $\delta_c^L = 1.686$, according to the collapse of a spherical fluctuation. This value corresponds in an Einstein–de Sitter model to the full collapse in linear theory, when the halo virialises. The void case is different: if the initial underdensity identifying a void is deep enough, its evolution is not marked by any specific event, and it continues its outward-directed expansion forever. It is common to consider the shell-crossing\(^5\) condition as the event that identifies the void formation (Blumenthal et al. 1992; Sheth & van de Weygaert 2004; Jennings et al. 2013), but this condition strictly depends on the initial density profile of the underdensity.

The linear threshold for collapse is fixed at $\delta_c^L = 1.686$, according to the collapse of a spherical fluctuation. This value corresponds in an Einstein–de Sitter model to the full collapse in linear theory, when the halo virialises. The void case is different: if the initial underdensity identifying a void is deep enough, its evolution is not marked by any specific event, and it continues its outward-directed expansion forever. It is common to consider the shell-crossing\(^5\) condition as the event that identifies the void formation (Blumenthal et al. 1992; Sheth & van de Weygaert 2004; Jennings et al. 2013), but this condition strictly depends on the initial density profile of the underdensity.

The multiplicity function, as given by Sheth & van de Weygaert (2004), is:

$$f_{\text{L},\sigma}(\sigma) = 2 \sum_{j=1}^{N_s} \exp \left( -\frac{(\pi x)^2}{2} \right) \pi x^2 \sin \left( \pi D \right),$$

with

$$D = \frac{\delta_c^L}{\delta_c^L + \delta_v^L}, \quad \chi = \frac{D}{\delta_c^L} \sigma,$$

where $\sigma$ is the square root of the variance of linear matter perturbations on the Lagrangian scale $R_L$. All these quantities are computed in the linear regime, on which the excursion-set formalism relies. The void size function in Lagrangian space is then readily derived as (Sheth & van de Weygaert 2004; Jennings et al. 2013):

$$\frac{dn_L}{d\ln r_L} = f_{\text{L},\sigma}(\sigma) \frac{\ln \sigma^{-1}}{V(r_L)} d \ln r_L,$$

where $V(r_L) = 4\pi r_L^3/3$ is the volume of the spherical fluctuation of radius $r_L$. Conversely to the case of the halo mass function, the void size function in Eulerian space is different from the one in Lagrangian space. Firstly, the expansion of voids from linear to nonlinear theory has to be taken into account. The evolution of perturbations in the nonlinear regime provides the conversion from the linear to the nonlinear shell radius:

$$\frac{r}{r_L} = \left( \frac{\rho}{\rho_c} \right)^{1/3},$$

where $\rho$ is the mean density of the Universe and $\rho_c$ is the average density within the void. Second, to prevent the fraction of the volume occupied by voids from exceeding unity in the transition from linearity to nonlinearity, we fix the void volume fraction of the Universe to be equal in the linear and in the nonlinear regimes (Jennings et al. 2013):

$$V(r) \, dn = V(r_L) \, dn_{L,|r_L=1}(r).$$

With this requirement the model ensures void volume conservation — hereafter Vdn model, following Jennings et al. (2013) — and from Eq. (3) we can derive the final definition of the theoretical void size function adopted in this paper:

$$\frac{dn}{d\ln r} = f_{\text{L},\sigma}(\sigma) \frac{\ln \sigma^{-1}}{V(r)} d \ln \eta_L |_{r_L=1}(r).$$

3.2. Methodology

To compare the measured and the theoretical void size functions, we need to link objects found by the void finder in the tracer distribution with the ideal spherical and isolated voids described by the void size function theoretical model (Jennings et al. 2013). Any watershed void finder defines a region spanning from its density minimum to its overdense ridge (Roerdink & Meijster 2000; Neyrinck 2008; Platen et al. 2007; Sutter et al. 2015). On the contrary, the theoretical voids are matter density fluctuations for which the mean density contrast in a sphere reaches a specific threshold value at a radius $R_{\text{thr}}$. Previous papers attempted to mitigate this difference by modifying the threshold of the model (Pisani et al. 2015a; Sahlén et al. 2016; Sahlén 2019), in particular considering marginalisation over the threshold, for cosmological uses of the model.

It is useful to recall that the Vdn model describes voids evolving in the total matter density field, but that in our case (and when dealing with data) we can only identify voids in the galaxy density field. Therefore, to align these objects to those modelled by the theory, we need to relate the characteristic density threshold used in the theoretical model, $\delta_c^L$, to the corresponding one in the galaxy density field. To accomplish this purpose, we relied on the following two steps for data preparation: first, we measured the mean density profiles of cosmic voids to find the radius of the sphere at which the mean density contrast reaches the desired value $\delta_{c,\text{thr}}$ in the galaxy distribution.
(Jennings et al. 2013; Ronconi & Marulli 2017; Ronconi et al. 2019; Contarini et al. 2019; Verza et al. 2019), i.e. the resized radius, $R_{\text{eff}}$. Second, it is necessary to find the corresponding density contrast in the underlying matter density distribution, within the resized radius. Recently the properties of voids in the galaxy distribution, as well as of galaxies and tracer bias within cosmic voids, have been explored extensively (Furlanetto & Piran 2006; Sutter et al. 2014b; Neyrinck et al. 2014; Pollina et al. 2017, 2019; Contarini et al. 2019, 2021). To recover the matter density contrast corresponding to the threshold value in the galaxy density field, we need to model the galaxy distribution inside cosmic voids taking into account tracer bias. To describe tracer bias, i.e. the bias of the cosmological objects chosen to trace voids, various possibilities have been considered, including a full theoretical discussion (see discussion in Verza et al. 2019 and Desjacques et al. 2018 for an extensive review), or a robust modelling of bias inside voids based on simulations (Pollina et al. 2017, 2019; Contarini et al. 2019, 2021). In this analysis we chose to rely on the latter, following Contarini et al. (2019, 2021). These works showed that it is possible to extend the Vdn model by considering a linear relationship between tracer and matter density contrast in cosmic voids, $\delta^{\text{NL,DM}}_{v,\text{DM}}$ and $\delta^{\text{NL,DM}}_{v,\text{tr}}$, with a dependence only on the large-scale effective bias $b_{\text{eff}}$:

$$\delta^{\text{NL,DM}}_{v,\text{DM}} = \frac{\delta^{\text{NL,DM}}_{v,\text{tr}}}{\mathcal{F}(b_{\text{eff}})} \rightleftharpoons \delta^{\text{NL,DM}}_{v,\text{DM}^*},$$

(7)

where $\delta^{\text{NL,DM}}_{v,\text{DM}}$ is the value of the threshold in the dark matter field to be used in the Vdn model, after its conversion in linear theory (Jennings et al. 2013). For ACDM and the DE equations of state considered in this work, the conversion from nonlinear to linear density contrast in the matter field is cosmology and redshift independent with very high accuracy (Jennings et al. 2013; Pace et al. 2017), allowing us to exploit the fast and precise Beanardeau (1994) fitting formula:

$$\delta^{*}_v = C \left[ 1 - \left( 1 + \frac{\delta^{\text{NL}}_v}{C} \right)^{-1} \right], \quad \text{with C = 1.594}. \rightleftharpoons \delta^{\text{NL}}_v = \frac{\delta^{*}_v C}{1 - \left( 1 + \frac{\delta^{\text{NL}}_v}{C} \right)^{-1}},$$

(8)

Comparing the void density profiles computed both in the Friend-of-Friends (FoF) halos and dark matter particle field, Contarini et al. (2019) found that the function $\mathcal{F}$, reported in Eq. (7), is well modelled as a linear relation of the large-scale effective bias $b_{\text{eff}}$:

$$\mathcal{F}(b_{\text{eff}}) = B_{\text{slope}} b_{\text{eff}} + B_{\text{offset}}.$$

(9)

where $B_{\text{slope}}$ and $B_{\text{offset}}$ are the values of the first and second coefficients of the linear function, respectively. This relation will be calibrated in this work using the samples of galaxies and voids extracted from the Flagship light-cone and the resulting values of $B_{\text{slope}}$ and $B_{\text{offset}}$, together with the associated uncertainties, will be presented in Sect. 4.1. The parametrisation introduced in equations Eq. (7), (9) was tested also in Contarini et al. (2021), using different selection criteria for the halo identification and verifying its negligible dependence on the cosmological model. In particular, this relation is tested varying the neutrino mass and the parameters of the $f(R)$ class of modified gravity models, in the form introduced by Hu & Sawicki (2007). The quantity represented by the function $\mathcal{F}(b_{\text{eff}})$ parametrises the value of the tracer effective bias measured inside cosmic voids and it has been denoted $h_{\text{punc}}$\footnote{We note that $h_{\text{punc}}$ and $\mathcal{F}(b_{\text{eff}})$ refer to the same quantity, i.e. the value of the tracer bias computed inside cosmic voids. This quantity represents the relation between the void density profiles computed using biased (e.g. galaxies) and unbiased (e.g. dark matter particles) mass tracers (see Pollina et al. 2017, 2019; Contarini et al. 2019, 2021). Nevertheless, we keep two different notations since their computation is different: $h_{\text{punc}}$ is measured for each bin of redshift, while $\mathcal{F}(b_{\text{eff}})$ is given as a function of $b_{\text{eff}}$ and varies linearly with it by construction (see Sect. 4.1).} in Contarini et al. (2019, 2021). To facilitate the reader’s comprehension of the adopted methodology, we summarise in Fig. 1 all the negative density contrasts mentioned in this paper and their relative relations.

To convert the underdensity threshold of the Vdn model according to the function $\mathcal{F}(b_{\text{eff}})$, we first need to compute accurately the large-scale effective linear bias of our galaxy sample. For this estimate we followed the same prescriptions described in Marulli et al. (2013, 2018). In particular, we exploited the galaxy two-point correlation function (2PCF), performing a Bayesian statistical analysis to infer the effective bias, $b_{\text{eff}}$. We computed the angle-averaged galaxy 2PCF $\xi(r)$ in real space creating a random catalogue 10 times larger than the original one and using the Landy & Szalay estimator (Landy & Szalay 1993).

We then estimated the covariance matrix, which measures the variance and correlation between the different bins of the 2PCF. For this purpose we applied the Bootstrap method, dividing the original catalogues in 125 sub-catalogues and constructing 100 realisations by resampling from the sub-catalogues, A162, page 6 of 20
with replacement. In the end we performed a full Markov chain Monte Carlo (MCMC) analysis of the 2PCF, using a Gaussian likelihood function. The 2PCF model, $\xi_{\text{mod}}(r)$, is computed as follows:

$$
\xi_{\text{mod}}(r) = b_{\text{eff}}^2 \xi_{m}(r),
$$

where $\xi_{m}(r)$ is the matter 2PCF, which is estimated by Fourier transforming the matter power spectrum, $P_m(k)$, computed with the Code for Anisotropies in the Microwave Background (CAMB\(^7\), Lewis et al. 2000). Then we accurately evaluated the effective bias parameter $b_{\text{eff}}$ by sampling its posterior distribution with the MCMC modelling in the range of scales of $[20\,\text{–}\,40]\,\hMpc$. We underline that the relative error associated to $b_{\text{eff}}$ is expected to be relatively small because of the strategy used to compute this quantity relying on the galaxy catalogue in real space and assuming the true cosmological parameters of the simulation. A more complete and realistic treatment will be performed in the future, including in the analysis the modelling of the multipoles of the 2PCF, which will allow us to take into account the effects of redshift-space and geometrical distortions (see e.g. Scoccimarro 2004; Taruya et al. 2010; Beutler et al. 2017). Additionally, an alternative methodology to extract Flagship galaxy bias is to follow e.g. Tutusaus et al. (2020), who parametrised the Flagship galaxy bias as a function of $z$, albeit for the photometric redshift selection.

3.3. Bayesian statistical analysis

In this work we used a reliable forecast method for the sensitivity of the void size function in the Euclid survey to constrain the cosmological model, based on a parameter extraction from Bayesian analysis with MCMC (Perotto et al. 2006; Wang et al. 2009; Lahav et al. 2010; Martinelli et al. 2011; de Bernardis et al. 2011; Wolz et al. 2012; Hamann et al. 2012; Khedekar & Majumdar 2013; Audren et al. 2013).

In order to forecast the sensitivity of void counts with an MCMC analysis in Euclid, we have to consider that the Flagship simulation covers about one third of the Euclid survey. We obtained the Euclid predicted void number counts relying on the theoretical void size function model validated on the Flagship simulation (see Sect. 4.1), that is assuming a fiducial ΛCDM cosmology with the cosmological parameters of the Flagship and the calibration in redshift space of the VdM model described in Sect. 4.1. We assumed the same binning of void radii employed in our Flagship analysis but consider a survey area matching the one expected for Euclid (roughly three times the Flagship area), rescaling the Poissonian errors of the void number counts consistently by a factor $\sqrt{3}$.

This allows us to use MCMC analysis to explore the likelihood distribution in the parameter space without any assumption on the Gaussianity of parameters and local approximations around the fiducial value, as in Fisher forecasts. Moreover, according to the Cramér–Rao inequality, the Fisher matrix gives a lower bound on the error on a parameter (Kendall et al. 1987), while the MCMC is proven to be more realistic, in particular in the presence of degeneracies (Perotto et al. 2006; Wolz et al. 2012; Audren et al. 2013; Sellentin et al. 2014). Finally, this kind of approach allows us to compute unbiased constraints, with confidence contours centred on the Flagship simulation cosmological parameters and on the calibrated nuisance parameters $B_{\text{shape}}$ and $B_{\text{offset}}$.

According to Bayes’s theorem, given a set of data $D$, the distribution of a set of parameters $\Theta$ in the cosmological model considered is given by the posterior probability:

$$
P(\Theta|D) \propto \mathcal{L}(D|\Theta) \cdot p(\Theta),
$$

where $\mathcal{L}(D|\Theta)$ is the likelihood and $p(\Theta)$ the prior distribution. Since in this work we consider the number counts of cosmic voids, the likelihood can be assumed to follow Poisson statistics (Sahlén et al. 2016):

$$
\mathcal{L}(D|\Theta) = \prod_{i,j} N(r_i, z_j|\Theta)^{N(r_i, z_j|\Theta)} \exp \left[-N(r_i, z_j|\Theta)\right],
$$

where the product is over the radius and redshift bins, labelled as $i$ and $j$ respectively. The $N(r_i, z_j|\Theta)$ quantity corresponds to the number of voids in the $i$th radius bin and $j$th redshift bin, while $N(r_i, z_j|\Theta)$ corresponds to the expected value in the cosmological model considered, given a set of parameters $\Theta$. In our work, the former is obtained from the Flagship analysis (with the void size function model validated on the Flagship simulation, but considering that the Euclid area will be three times larger), while the latter is given by the predictions of the void size function model varying the considered cosmological parameters $\Theta$.

In performing the MCMC analysis, the mapping between redshift and comoving distance changes with the cosmological parameters assumed at each step of the chain. This introduces geometrical distortions for all the considered sets of cosmological parameters (different from the true one). We used a fiducial cosmology to build up the void catalogue, and, in computing the likelihood, we theoretically accounted for the distortion effects on the quantities we measured. In particular, geometrical distortions can be modelled with two effects: they vary the inferred survey comoving volume and introduce the Alcock–Paczyński (Alcock & Paczynski 1979) distortion. The effect on the survey volume impacts the number of voids expected in the survey. The theoretical void size function model predicts the number density of voids in each radius and redshift bin. Therefore, to obtain the total number of voids, the number density has to be multiplied by the volume, which is impacted by the cosmology. On the other hand, the Alcock–Paczyński (Alcock & Paczynski 1979) distortion affects the size of voids and introduces an anisotropy between the orthogonal and the parallel direction with respect to the line-of-sight. These quantities change according to (Sánchez et al. 2017a):

$$
\begin{align*}
\hat{r}_q &= \frac{H(z)}{H^\prime(z)} r_q = q_\parallel^{-1} r_\parallel, \\
\hat{r}_\perp &= \frac{D_A(z)}{D_L(z)} r_\perp = q_\perp^{-1} r_\perp;
\end{align*}
$$

where $r_\parallel$ and $r_\perp$ are the comoving distances between two objects at redshift $z$ projected along the parallel and perpendicular direction with respect to the line-of-sight, $H(z)$ is the Hubble parameter and $D_A(z)$ the comoving angular-diameter distance. The primed quantities refer to the calculation at the fiducial cosmology, the non-primed to the true cosmology, assumed in a MCMC step. It follows that the volume of a sphere with radius $R$ appears modified according to $V = q_\parallel^{3/2} q_\perp^{-3} R^3$ (Ballinger et al. 1996; Eisenstein et al. 2005; Xu et al. 2013; Sánchez et al. 2017a; Hamaus et al. 2020; Correa et al. 2021), so the void size function expected in the survey is shifted accordingly. We checked

\(^7\) http://camb.info
the validity of this relationship varying the cosmological model considered two cosmological models, extending the standard $\Lambda$CDM with a DE equation of state, focusing in particular the DE equation-of-state parameters. We assumed the void centres to remain at the same locations at different cosmologies. While void shapes can suffer from symmetric geometrical distortions, this marginally affects the identification of void centres, and the effect is even smaller since the void size function is an averaged quantity. Furthermore, the variation caused by the change of the cosmological parameters on void radii is taken into account by the modelling of the Alcock–Paczynski effect, therefore the cleaning procedure (see Sect. 2.2) is applied only once to the void sample, considering a fiducial $\Lambda$CDM cosmology. We note that the combination of the two effects – volume effect acting on the expected number density, and the Alcock–Paczynski effect acting on the void sizes – enhances the constraining power of the void size function.

### 3.4. Cosmological models

The aim of this work is to investigate the constraining power of the void number count statistic on cosmological parameters, focusing in particular the DE equation-of-state parameters. We considered two cosmological models, extending the standard $\Lambda$CDM with different DE equation of states. The first model, $w_0\Lambda$CDM, implements a constant DE equation of state $w$; the second one, $w_0w_a\Lambda$CDM, parametrises dynamical DE models with the popular Chevallier–Polarski–Linder (CPL) equation of state (Chevallier & Polarski 2001; Linder 2003):

$$w_{\text{CPL}}(z) = w_0 + w_a \frac{z}{z + 1}. \quad (14)$$

Both cosmological models consider a flat universe and do not include spatial fluctuations of the DE, which are negligible given the scales considered in this work (see e.g. Khoraminezhad et al. 2020). We performed the MCMC analysis of each cosmological model focusing on different sets of free cosmological parameters: together with the DE equation of state parameters (i.e. $w$ or $w_0$ and $w_a$, depending on the cosmological model) the density parameter $\Omega_m$ or the sum of neutrino masses $M_\nu$ are allowed to vary. Moreover, we analysed both the cases with two different approaches: firstly, fixing the parameters of the extended Vdn model, $B_{\text{slope}}$ and $B_{\text{offset}}$, to the median values obtained from the calibration performed with Flagship data (label: fixed calibration); secondly, allowing $B_{\text{slope}}$ and $B_{\text{offset}}$ to vary in the parameter space described by a 2D Gaussian distribution centred on their median values and given by the calibration with the Flagship simulation (label: relaxed calibration).

The two adopted approaches are meant to demonstrate the impact of the calibration that will be performed in Sect. 4.1 on the cosmological forecast. In this work the constraints on the parameters $B_{\text{slope}}$ and $B_{\text{offset}}$ are indeed limited to the statistical relevance of the number counts of voids identified by means of the Flagship galaxies. The case in which the cosmological forecasts are computed fixing $B_{\text{slope}}$ and $B_{\text{offset}}$ to their exact calibrated values represents therefore an optimistic evaluation of the results that we may obtain in the future thanks to the usage of larger mock catalogues, or by means of a fully theoretical modelling of the tracer bias inside cosmic voids (see Sect. 3.2).

The cosmological model considered for the analysis is characterised by a primordial comoving curvature power spectrum amplitude fixed to the Flagship simulation value, $A_s = 2.11 \times 10^{-9}$. We followed the strategy to fix this parameter in order to mimic the future application to real data, which will be supported by the impressive constraints obtained from the study of CMB anisotropies by Planck Collaboration VI (2020). Thanks to this approach, for each MCMC step we derived $\sigma_{B_{\text{slope}}}$ and $\sigma_{B_{\text{offset}}}$ fixed in measurements to the corresponding one in the matter distribution, as described in Sect. 3.2.

First of all, we verified the calibration of the relation $\mathcal{F}(b_{uv})$ reported in Eq. (9) using the Flagship simulation. To this end,
we extracted the value of $B_{\text{slope}}$ and $B_{\text{offset}}$ by leaving them as free parameters with uniform priors of the extended Vdn model and fitting the measured void number counts in the selected redshift bins, considering also a Gaussian prior for $b_{\text{eff}}$ at each redshift. We notice that, since the error on the effective bias only corresponds to a few percent of its value, the variation allowed for this parameter during the fit is small. All the remaining cosmological parameters were kept fixed to the Flagship simulation values during this calibration.

With this prescription we obtained the confidence levels reported on the left panel of Fig. 2, for the void size function measured in both real and redshift space in light blue and orange, respectively. The resulting coefficients for the calibrated relations are:

\begin{align}
\mathcal{F}(b_{\text{eff}}) &= (0.96 \pm 0.04) b_{\text{eff}} + (0.44 \pm 0.07) \quad \text{(16)} \\
\mathcal{F}(b_{\text{eff}}) &= (0.96 \pm 0.03) b_{\text{eff}} + (0.26 \pm 0.06), \quad \text{(17)}
\end{align}

for the redshift-space and the real-space void abundance, respectively.

We show on the right panel of Fig. 2 the corresponding linear relations obtained with these calibrations, with a shaded area representing an uncertainty of 2$\sigma$. As a comparison, we present in the same plot the values computed for $b_{\text{punct}}$, leaving it as the only free parameter of the model and fitting separately the measurements at different redshifts. This analysis is aimed at testing the precision of the calibrated relations for each redshift: in the right plot of Fig. 2 the markers with best match to the linear relations correspond in Fig. 3 to the redshift bins for which the calibrated coefficients $B_{\text{slope}}$ and $B_{\text{offset}}$, for real (blue) and redshift space (orange), while the shaded regions indicate an uncertainty of 2$\sigma$ on the relationships. The markers represent the calibration obtained for each bin of redshift, leaving $b_{\text{punct}}$ as the only free parameter of the void size function model when fitting the measured void number counts. This alternative calibration provides a value of $b_{\text{punct}}$ for each redshift of the sample and is associated with the value of the effective bias $b_{\text{eff}}$ of the Flagship galaxies at that specific redshift. As a comparison we also show the linear function calibrated using FoF dark matter haloes in real space by Contarini et al. (2019), displayed with a dashed grey line.

Finally, we report also the calibration performed in Contarini et al. (2019) using the dark matter haloes of the COUpled Dark Energy Cosmological Simulations (CoDECs, Baldi 2012), and represented in grey in the right panel of Fig. 2. At lower redshifts the calibration we measure in this paper is in good agreement with the calibration from the CoDECs simulation, characterised by a cosmology consistent with the results of the seven-year Wilkinson Microwave Anisotropy Probe (WMAP7, Komatsu et al. 2011), but it slightly deviates from the latter at higher redshift values. The reason for this minor deviation is twofold. Firstly it is linked to the kind of cosmic tracers (i.e. dark matter haloes or galaxies) and the selection criteria (i.e. minimum mass or magnitude) used to identify voids (Contarini et al. 2021). Secondly it is related to the fact that in Contarini et al. (2019) the calibration was performed for redshift from 0 to 1, while here we are testing this relationship beyond this range. The physics underlying the function $\mathcal{F}(b_{\text{eff}})$ and its relation with the mass tracers used to identify voids will be investigated in future papers.

More importantly, since the void size function will be measured on real data from the Euclid survey, we have to deal with voids detected in redshift space. The overall effect of RSD on voids, relevant for the void size function, is an apparent enlargement of the voids’ volume, due to the elongation along the line of sight. This is reflected in a mean shift of the measured void size function towards greater radii. Even if this effect can in principle be theoretically modelled (Pisani et al. 2015b; Correa et al. 2021), we decide to parametrise it empirically as described below. Indeed, the theoretical approach requires knowledge of the void matter density profile for the entire void population, which has to be characterised in simulations and may introduce some model dependencies. We found that the parametrisation of $\mathcal{F}(b_{\text{eff}})$ can be exploited to encapsulate also the modifications on the void sizes caused by the enlargement of cosmic voids in redshift space. This approach has the advantage of being both simple to model and robust, allowing us to take into account,
Fig. 3. Comparison between the measured void number counts as a function of $R_{\text{eff}}$ (the void radii rescaled by the cleaning algorithm), and the theoretical predictions given by the extended Vdn model, in six different redshift bins. The dark green circles and the dark red diamonds represent the measured void size functions in real and redshift space, respectively, while the corresponding model predictions are depicted in light blue and orange. The shaded regions indicate the uncertainty of 2σ assigned to the model through the calibration of the extended Vdn parameters. Bottom panels: report the residuals computed as the difference of data points from the relative theoretical model, divided by the Poissonian error associated with each data point. The hatched regions represent a band with amplitude 2 useful to check if the data points, considered with a 2σ error, are compatible with the main theoretical curve.

with the same parameter, both the impact of tracer bias in voids and of the RSD. Moreover, this approach is fully agnostic and does not require any assumption about the void density profile, nor any other modelling, making it particularly suited to survey analyses.

It is worth noting that the relation obtained for voids in redshift space shows a greater offset but almost the same slope with respect to its analog in real space. This difference reflects the increase of void sizes in redshift space. It also opens the way to test theoretical implementations in future work, indicating that a simple modelling of those effects should suffice to extract robust constraints.

Equipped with these calibrated relations, we now have all the elements necessary to compare the measured void size function with the theoretical predictions given by the extended Vdn model, in which the underdensity threshold is converted as
described in Sect. 3.2. Figure 3 provides the main results of our Flagship analysis. We show the comparison between the measured void number counts and the corresponding theoretical void size functions, both in real and redshift space, for the six equi-populated bins in redshift. The Poissonian errors related to the data are represented by the error bars, while the uncertainty related to the theoretical model is shown as a shaded region. The latter is computed associating an error to the theoretical model is shown as a shaded region. The data are represented by the error bars, while the uncertainty related to the theoretical model is shown as a shaded region. The latter is computed associating an error to the theoretical model. In Sect. 3.3, we report the constraints obtained by fixing the calibration parameters with blue contours marked by a solid line and the results obtained by relaxing the calibration constraints with light-blue contours marked by a dashed line (see Sect. 3.4).

To test the goodness of the fits shown in Fig. 3 we computed the reduced \( \chi^2 \) using the weighted sum of squared deviations of the two data sets from their corresponding models and dividing the results by the degrees of freedom of the two systems, \( \nu = N_{\text{bin}} - N_{\text{par}} \), where \( N_{\text{bin}} \) is the number of bins in radius used to compute the void counts and \( N_{\text{par}} \) is the number of free parameters of the model. In our case \( N_{\text{bin}} = 50 \) and \( N_{\text{par}} = 2 \), since we are fitting the void counts simultaneously for all the redshift shells and we are considering \( B_{\text{slope}} \) and \( B_{\text{offset}} \) only as free parameters of the model. The results are \( \chi^2 = 1.60 \) and \( \chi^2 = 1.02 \) for real and redshift space, respectively.

### 4.2. Cosmological forecasts

In this section we provide the cosmological forecasts obtained using the void size function in redshift space in the perspective of the \textit{Euclid} mission. We applied the statistical analysis described in Sect. 3.3 to derive constraints on the parameters of the two cosmological models analysed, labelled as \( w \text{CDM} \) and \( \text{wCDM} \), following the two approaches described in Sect. 3.4. For the model \( w \text{CDM} \) we assumed a flat prior for all the remaining free cosmological parameters of the model, and for the model \( \text{wCDM} \) we assumed a Gaussian prior distribution with standard deviation \( \sigma = 5 \) for \( w_{\text{T}} \) and \( \sigma = 15 \) for \( v_r \), both centred on the true values of these parameters, given by the Flagship simulation cosmology \((w_{\text{T}} = -1, v_r = 0)\). We preferred to use very wide Gaussian priors instead of uniform ones to improve the numerical stability of the whole pipeline, but we tested that uniform priors yield consistent results. The remaining cosmological parameters analysed in this work \((\Omega_m, \sigma_8)\) were included in the void size function modelling with uniform prior distributions.

In Fig. 4 we present the 68% and 95% confidence levels of the constraints on the model \( w \text{CDM} \). In the left plot we show the \textit{Euclid} forecasts from a void size function model characterised by \( w \) and \( \Omega_m \) as free cosmological parameters. We represent with different colours and borders the results obtained with the two approaches described in Sect. 3.4: in blue with solid contours the forecasts obtained by fixing the extended Vdn parameters \( B_{\text{slope}} \) and \( B_{\text{offset}} \), in light-blue with dashed contours those obtained by relaxing the calibration constraints by means of a 2D Gaussian prior on \( B_{\text{slope}} \) and \( B_{\text{offset}} \), which distribution is represented in the left panel of Fig. 2. In the right plot we represent the same forecasts but considering a void size function model with neutrino total mass \( M_\nu \) as free parameter instead of the matter density \( \Omega_m \). In this case we show the fixed and the relaxed calibration approach results in red and orange, respectively. In both the presented cases \( \sigma_8 \) is computed as derived parameter. As expected, the effect of relaxing the calibration constraints is to broaden the confidence contours.
\textit{w}_0\textit{w}_a\textit{CDM} model - fixed calib.,  
\textit{w}_0\textit{w}_a\textit{CDM} model - relaxed calib.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Same as Fig. 4 but for the cosmological model labelled as \textit{w}_0\textit{w}_a\textit{CDM}, having a dynamical DE component described by the CPL parametrization (see Sect. 3.4).}
\end{figure}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|c|}
\hline
\textbf{Model} & \textbf{w} & \textbf{\textsigma{w}} & \textbf{\Omega_m} & \textbf{\textmu{}\text{eV}} & \textbf{B_{slope}} & \textbf{B_{offset}} & \\
\hline
\textbf{Fixed calib.} & -1.01^{+0.09}_{-0.11} & 0.83 \pm 0.03 & 0.319^{+0.005}_{-0.004} & 0 & 0.96 & 0.44 & \\
& -0.99^{+0.06}_{-0.04} & 0.83^{+0.1}_{-0.2} & 0.319 & <0.03 & 0.96 & 0.44 & \\
\textbf{Relaxed calib.} & -1.0 \pm 0.1 & 0.84 \pm 0.04 & 0.318^{+0.008}_{-0.005} & 0 & 0.96 \pm 0.02 & 0.44 \pm 0.04 & \\
& -0.98^{+0.10}_{-0.07} & 0.83^{+0.02}_{-0.03} & 0.319 & <0.06 & 0.95 \pm 0.02 & 0.46 \pm 0.04 & \\
\hline
\end{tabular}
\caption{Cosmological forecasts computed for the \textit{Euclid} mission from the void size function for the cosmological model \textit{w}CDM.}
\end{table}

\textbf{Notes.} In this table we report the results of the two analysis strategies adopted in this work: considering the parameters \textit{B}_{slope} and \textit{B}_{offset} fixed to the respective median calibrated values (label: fixed calib.) or with a multivariate Gaussian with the same median value but a constraining power given by the calibration procedure with Flagship (label: relaxed calib.). For each of the two cases we present, in the upper and lower line, the forecasts obtained fixing \textit{M}_e or \textit{\Omega}_m to the Flagship simulation true values, respectively. All the constraints are reported with errors with a 1σ confidence level.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|}
\hline
\textbf{Model} & \textbf{w} & \textbf{\textsigma{w}} & \textbf{\Omega_m} & \textbf{\textmu{}\text{eV}} & \textbf{B_{slope}} & \textbf{B_{offset}} & \textbf{FoM}_{\textit{w}_0\textit{w}_a\textit{CDM}} & \\
\hline
\textbf{Fixed calib.} & -1.0 \pm 0.2 & -0.1^{+0.1}_{-0.1} & 0.84^{+0.04}_{-0.03} & 0.32 \pm 0.01 & 0 & 0.96 & 0.44 & 4.9 & \\
& -1.0^{+0.2}_{-0.6} & -0.1^{+0.3}_{-0.8} & 0.319 & <0.08 & 0.96 & 0.44 & 17 & \\
\textbf{Relaxed calib.} & -0.8^{+0.5}_{-0.6} & -0.9^{+0.6}_{-0.6} & 0.86 \pm 0.04 & 0.32 \pm 0.01 & 0 & 1.01^{+0.04}_{-0.04} & 0.55^{+0.08}_{-0.05} & 0.78 & \\
& -0.9^{+0.3}_{-0.2} & -0.5^{+0.9}_{-1.3} & 0.86^{+0.02}_{-0.03} & 0.319 & <0.08 & 0.99^{+0.01}_{-0.04} & 0.38^{+0.07}_{-0.01} & 2.3 & \\
\hline
\end{tabular}
\caption{Same as Table 2 but for the \textit{w}_0\textit{w}_a\textit{CDM} scenario.}
\end{table}

\textbf{Notes.} In this case we present in the last column also the values computed with Eq. (18) to estimate the FoM for the DE equation of state.

In Fig. 5 we show the same contours represented in Fig. 4 but considering the \textit{w}_0\textit{w}_a\textit{CDM} scenario. The free cosmological parameters of the void size function model are the coefficients of the DE equation of state, \textit{w}_0 and \textit{w}_a, together with \textit{\Omega}_m (left plot) or \textit{\textmu{}} (right plot). Also in this case the relaxation of the constraining condition of the calibration parameters causes an enlargement of the confidence contours. In this scenario however, the strongest impact of the calibration constraints is on the \textit{w}_1=\textit{w}_a parameter plane, in particular along the diagonal where these parameters become degenerate. The effect of the calibration constraints on \textit{\Omega}_m and \textit{\textmu{}} has a lower impact.

In Tables 2 and 3 we report the values, with relative 1σ errors, of the cosmological constraints derived for the \textit{w}CDM and \textit{w}_0\textit{w}_a\textit{CDM} scenario, respectively. The constraints on the sum of neutrino masses \textit{\textmu{}} are expressed as a 1σ upper limit. For each table we show the results for the two approaches followed in this paper: fixing and relaxing the calibration constraints on the void size function model. The calibration parameter are reported in the columns \textit{B}_{slope} and \textit{B}_{offset} for completeness. Notice that each quantity reported without any uncertainty is considered fixed in the specific scenario presented in that table row.
For the $w_0 w_a$CDM scenario, in order to evaluate the constraining power of the void size function on the DE equation of state, we derived the figure of merit (FoM) for the coefficients of the CPL parametrisation $w_0$ and $w_a$. We computed this value by following Wang (2008):

$$\text{FoM}_{w_0 w_a} = \frac{1}{\sqrt{\det \text{Cov}(w_0, w_a)}},$$

(18)

where $\text{Cov}(w_0, w_a)$ represents the covariance matrix of the DE equation of state parameters. We notice that this definition is in agreement with the one adopted in Euclid Collaboration (2020). We underline that, contrary to other constraint accuracy indicators (e.g. the relative or percentage errors), the higher the FoM value, the better the precision on the analysed cosmological parameters. We report this quantity in the last column of Table 3.

As a first exploration of the cosmic void statistics combined power, we now compare the forecasts from the void size function provided in this work with other Euclid forecasts. We present as a first comparison the results of the $\Omega_{de} - w$ confidence contour with the model-calibrated forecasts presented in Hamaus et al. (2022). The latter are computed by modelling the observable distortions of average shapes in redshift space via RSD and the Alcock–Paczyński effect, for voids to be measured in the Euclid spectroscopic galaxy distribution. Contrary to the model-independent case, in the presented approach the nuisance parameters of the model have been calibrated by means of Flagship data. In this comparison we consider the wCDM scenario with fixed neutrino mass and we focus on the $\Omega_{de} - w$ parameter space. Given the assumption of flat spatial geometry, to compute the corresponding $\Omega_{de}$ forecasts, we converted the $\Omega_{de}$ obtained in the MCMC analysis as $\Omega_{de} = 1 - \Omega_m$.

As a second comparison we take the results of Fisher analysis reported by the inter-science taskforce for forecasting (IST/F, Euclid Collaboration 2020) obtained in the optimistic setting for the single probes weak lensing and galaxy clustering. We considered in this case the $w_0 w_a$CDM scenario with fixed neutrino mass and we focus on the $\Omega_m - \sigma_8$ degeneracy. To compute the IST confidence contour we made use of the publicly available\(^8\) Fisher matrices and we marginalised on the parameters not reported in the plot with the code Cosmofish (Raveri et al. 2016). We recall that amplitude of density fluctuations at $z = 0$, $\sigma_8$, is computed as a derived parameter in our analysis and its variation is given by the modifications caused by the free cosmological parameters of the model to the total matter power spectrum. We also stress the fact that a larger set of cosmological parameters is used in IST forecasts. This includes in particular the baryon matter energy density, $\Omega_h$, the dimensionless Hubble parameter, $h$, and the spectral index of the primordial density power spectrum, $n_s$. The impact on forecasts when including these parameters in the model will be tested in future work.

We show the presented comparisons in Fig. 6, representing in blue the forecasts obtained in this work considering a void size function model with fixed calibrated parameters. In the left panel we compare our results with the $\Omega_{de} - w$ confidence contour computed with the model-calibrated forecasts presented Hamaus et al. (2022; in magenta). In the right panel we show instead the comparison of $\Omega_m - \sigma_8$ confidence contour provided by IST forecasts considering the optimistic setting for weak lensing (in orange) and galaxy clustering (in purple). See Appendix B for an analogous comparison considering the cosmological forecasts presented above but with less optimistic modelling approaches.

In both panels we can appreciate the comparable extension of the presented contours and in the latter we can notice in particular the strong complementarity of the void size function forecasts with those of the Euclid standard probes. While a more accurate analysis would require proper accounting of covariance between analysed cosmological constraints, Fig. 6 shows how the presented probes explore the parameter space differently and motivates investigation on combination to be performed in future works.

5. Conclusions and discussion

In this work we presented state-of-the-art forecasts for cosmological constraints from the void size function to be expected from the Euclid mission. We measured the void number counts from the Flagship mock galaxy spectroscopic catalogue in redshift bins and matched the measurements with the theoretical definition given by the Vdn model (Sheth & van de Weygaert 2004; Jennings et al. 2013). We employed an extension of the Vdn model that conservatively accounts for the effects of the galaxy large-scale bias, $b_g$, on the void effective radii. With this method, we parametrised the Vdn model’s characteristic threshold $\delta_c$ according to the prescriptions of Contarini et al. (2019), also verifying the calibration of the function $F(b_g)$. The parametrisation method further allowed us to account for the modifications on the void sizes caused by the volume change of cosmic voids in redshift space.

We showed that the extended Vdn calibrated on Flagship data is effective in predicting the measured void number counts both in real and redshift space. Indeed we obtained a remarkable agreement between the measured and predicted void size functions, for all the redshift bins and all the spatial scales considered in our analysis. We also performed a MCMC analysis, estimating the constraints from void number counts on two main cosmological models: assuming in one case a scenario characterised by a constant equation-of-state parameter ($w$CDM) and in the other case a scenario with a dynamical DE component described by the CPL parametrisation ($w_0 w_a$CDM). For each scenario we presented the Euclid cosmological forecasts considering both approaches: by fixing the extended Vdn model parameters and by relaxing their boundaries to those provided by the calibration with Flagship mock catalogues. The former represents the ideal situation in which the simulations used to calibrate the void size function model allow us to have no uncertainties nor systematic errors on the calibration parameters; or alternatively, the case in which the value of the tracer bias inside voids is fully determined thanks to theoretical modelling.

In the wCDM scenario we forecasted relative percentage errors on the constant DE component, $w$, below the 10% for every analysed case. In the $w_0 w_a$CDM scenario, with the optimistic approach of fixing the model calibration parameters, we computed a FoM of $\Omega_m = 4.9$ or 17, in the case of leaving $\Omega_m$ or $M_0$ respectively, as additional free cosmological parameters of the model. As a reference, the corresponding FoM values computed by the IST/F for spectroscopic galaxy clustering and weak lensing are 55 and 44, respectively (see Table 13 in Euclid Collaboration 2020, for flat-$w_0 w_a$CDM cosmology and optimistic scenario).

The marginalised constraints on the derived parameter $\sigma_8$ are lower than 5% in every analysed case, while the relative errors on $\Omega_m$ are of the order of 2% in the wCDM scenario and of 3% in the $w_0 w_a$CDM scenario. The $1\sigma$ upper limit on $M_0$ is instead of

---

\(^8\) See https://github.com/euclidist-forecasting/fisher_for_public.
0.03 eV in the most optimistic case of the wCDM scenario and of 0.08 eV in the w0wCDM scenario. We recall that, in the cosmological models with free neutrino mass, the total matter energy density was fixed to the Flagship simulation true value, therefore the degeneracy of \( \Omega_\gamma \) with \( \Omega_m \) is not considered in the results.

Our analysis showcases the impressive constraining power of the void size function from the Euroid survey, strongly complementing the Euroid primary probes. This complementarity will make the combination powerful in particular for weak lensing and galaxy clustering, additionally enhancing robustness to systematic effects in both cases.

In this work we considered extremely conservative assumptions when analysing the void sample. Such conservative assumptions dramatically reduce the statistical power of our void catalogues, to ensure strong reliability: in the future, modelling improvements will allow a more efficient void selection, critically enhancing results while maintaining full robustness.

Among the conservative choices in modelling the void size function and in building the likelihood we recall the treatment of both the threshold value and of the minimum void radius accepted for the analysis. In particular, we selected a low value of the underdensity threshold \( \delta_{\text{crit}}^{\text{NL}} = -0.7 \) to avoid the shallowest voids in the sample, characterised by higher Poissonian noise contamination. We also strictly restricted the range of considered radii to avoid modelling poorly sampled voids of the Flagship galaxy catalogue, in order to prevent the inclusion of spatial scales affected by a loss of void counts. This conservative approach allows us to have a sample of voids composed by a limited number of objects but characterised by high purity. Different techniques will be tested in the future to better model the scales affected by numerical incompleteness (see e.g. Cousinou et al. 2019) and include them in the analysis, safely obtaining access to much larger statistics. A better modelling of these effects will lead to further improvements in the constraining power of the void size function.

Further future prospects to expand this work include exploiting void number count forecasts to predict constraints from other void applications (such as the stacked void-galaxy cross-correlation function, see Hamaus et al. 2022, void lensing, see Bonici et al. 2022, the void-void correlation function, see Kreisch et al. 2022), and subsequently combine joint constraints from voids with other Euroid probes (primary and not, e.g. galaxy clustering, galaxy weak lensing, cluster counts and clustering, baryonic acoustic oscillations, supernova distance measurements, CMB cross-correlations, etc.).

Other areas to explore include considering other cosmological parameters for the likelihood modelling, a more realistic treatment of observational effects (a more complex survey mask, a more realistic \( \sigma_8 \) and further survey-related systematic effects). Moreover, to prepare the application to real-survey data, the large-scale effective bias can be recovered from redshift-space catalogues. To this purpose, a number of methods are available (see e.g. Chan et al. 2012; Sheth et al. 2013; Lazeyras et al. 2016; Sánchez et al. 2017a) and Desjacques et al. 2018 for an extensive review), which would allow us to reliably model linear bias as a function of redshift, \( b(z) \). These methods can be exploited in the context of the Euroid mission, deriving tight constraints on \( b(z) \), especially from the cross-correlation between galaxy clustering and weak lensing, capable of breaking the degeneracies with cosmological parameters (Tutusaus et al. 2020).

Finally, a purely theoretical treatment of the impact of RSD and of the Alcock–Paczynski effect on the void size function, based on work from companion papers (Hamaus et al. 2022), as well as a comprehensive analysis of the systematic effects linked to these corrections is relevant to consider in future works.

For the future, a possible approach to model the measured Euroid void number counts would be to apply directly the relation \( \mathcal{F}(b_{\text{est}}) \) calibrated on the redshift-space void sample of the Flagship mock (see Eq. (16)). A close match of the galaxy properties in the mock catalogue with galaxies to be observed by
Euclid would make this approach particularly effective. Future works will aim to explore these methodologies and their impact on void constraints, along with testing the range of applicability of the $\mathcal{F}(b_{\text{eff}})$ relation and its possible cosmological dependency (expected however to be mild from Contarini et al. 2021).

This paper – with a first analysis on a full mock, the Euclid Flagship simulation – shows the impressive constraining capability of void number counts to tackle the properties of DE and neutrinos, demonstrating for the first time the feasibility of the technique with an end-to-end data-like application, and setting the ground for a robust use of the void size function for cosmology with Euclid.

Acknowledgements. We acknowledge the grant ASI n.2018-33-HH.0 SC, FM and LM acknowledge the use of computational resources from the parallel computing cluster of the Open Physics Hub (https://site.unibo.it/openphysicshub/en) at the Physics and Astronomy Department in Bologna. GV is supported by Universtà degli Studi di Padova and in part by the project “Combining Cosmic Microwave Background and Large Scale Structure data: an Integrated Approach for Addressing Fundamental Questions in Cosmology”, funded by the MIUR Progetti di Rilevante Interesse Nazionale (PRIN) Bando 2017 – grant 2017YYJZAH. AP is supported by NASA ROSES grant 12-EAA2-12-0004, and NASA grant 15-WPRST15-0008 to the Nancy Grace Roman Space Telescope Science Investigation Team “Cosmology with the High Latitude Survey”. NH is supported by the Excellence Cluster ORIGINS, which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC-2094 – 390783311. MS acknowledges support by the P. E. Filén fellowship and a fellowship at the Swedish Consortium for Advanced Study (SCAS). LM acknowledges support from PRIN MIUR 2017 WSC32 “Zooming into dark matter and proto-galaxies with massive lensing clusters”. AR acknowledges funding from Italian Ministry of Education, University and Research (MIUR) through the “Dipartimenti di eccellenza” project Science of the Universe. He is supported in part by the project “Combining Cosmic Microwave Background and Large Scale Structure data: an Integrated Approach for Addressing Fundamental Questions in Cosmology”, funded by the MIUR Progetti di Rilevante Interesse Nazionale (PRIN) Bando 2017 – grant 2017YYJZAH We acknowledge use of the Python libraries Matplotlib (Harris et al. 2020), ChainConsumer (Hinton 2020), and Plotlib (Hunter 2007) and the Integrated Python Library for Cosmology (ChainConsumer 2016) workflows will aim to explore these methodologies and their impact on the cosmological parameters of DE and neutrinos, demonstrating for the first time the feasibility of the technique with an end-to-end data-like application, and setting the ground for a robust use of the void size function for cosmology with Euclid.
Appendix A: Void count measures

Table A.1 provides the number of voids identified in the redshift-space distribution of galaxies, in different redshift bins. We note that, conversely to Table 1, it relies on generic equi-spaced $\Delta z = 0.1$ bins, to facilitate the use for future forecasts of different void applications and combination with other probes. Here we report void number counts obtained both before and after the application of the cleaning procedure. Moreover, we show the number of cosmic voids considering both an optimistic and a pessimistic cut on smaller void radii, that is voids with radius over 1 time the MGS, and voids over 2 times the MGS, respectively. The lowering of void counts in the outermost bins is caused by survey mask effects at redshift boundaries of the simulated light-cone.

Table A.1. Void counts in nine equi-spaced bins in redshift, measured in the redshift-space mock galaxy catalogue, provided as a reference for future forecast analyses.

<table>
<thead>
<tr>
<th>$z$ range</th>
<th>shell volume ($h^{-1}$ Gpc$^3$)</th>
<th>MGS ($h^{-1}$ Mpc)</th>
<th>all voids</th>
<th>voids after cleaning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R &gt; $MGS</td>
<td>$R_{\text{eff}} &gt; $MGS</td>
</tr>
<tr>
<td>0.9 – 1.0</td>
<td>1.308</td>
<td>10.28</td>
<td>8928</td>
<td>4845</td>
</tr>
<tr>
<td>1.0 – 1.1</td>
<td>1.427</td>
<td>11.02</td>
<td>8987</td>
<td>5253</td>
</tr>
<tr>
<td>1.1 – 1.2</td>
<td>1.531</td>
<td>11.74</td>
<td>7735</td>
<td>4690</td>
</tr>
<tr>
<td>1.2 – 1.3</td>
<td>1.622</td>
<td>12.63</td>
<td>7167</td>
<td>4140</td>
</tr>
<tr>
<td>1.3 – 1.4</td>
<td>1.700</td>
<td>13.51</td>
<td>6575</td>
<td>3703</td>
</tr>
<tr>
<td>1.4 – 1.5</td>
<td>1.766</td>
<td>14.45</td>
<td>5636</td>
<td>3152</td>
</tr>
<tr>
<td>1.5 – 1.6</td>
<td>1.821</td>
<td>15.45</td>
<td>5132</td>
<td>2719</td>
</tr>
<tr>
<td>1.6 – 1.7</td>
<td>1.867</td>
<td>16.48</td>
<td>4389</td>
<td>2286</td>
</tr>
<tr>
<td>1.7 – 1.8</td>
<td>1.904</td>
<td>17.63</td>
<td>2248</td>
<td>851</td>
</tr>
<tr>
<td>0.9 – 1.8</td>
<td>14.95</td>
<td>13.69</td>
<td>56797</td>
<td>31639</td>
</tr>
</tbody>
</table>

Notes. The first column represents the minimum and the maximum redshift values for each bin, while the second and the third columns provide the volume in units of ($h^{-1}$ Gpc$^3$) corresponding to each shell of the sky octant, and the MGS, respectively. The next two columns show the number of voids identified by the VIDE void finder, selected with an effective radius greater than 1 and 2 times the MGS, respectively. In the last two columns we provide the void number counts obtained after the cleaning procedure, to be modelled using the void size function theory. The latter are reported with the same radius selections as described before. In the last row we show the total shells’ volume, the mean MGS and the total void counts corresponding to the entire range of redshifts.
Appendix B: Different setting forecasts

We show in Fig. B.1 the same forecast comparison presented in Fig. 6 but using less optimistic settings for the analyses. In particular, we report here the Euclid forecasts on a flat $wCDM$ and $w0w_aCDM$ cosmology with massless neutrinos, in the left and right panels respectively, indicating the void size function constraints as light-blue contours with dashed borders.

In the left panel of Fig. B.1 we compare our relaxed-calibration results (see Sect. 3.4 and Sect. 4.1), with the Euclid model-independent forecasts of Hamaus et al. (2022), represented in magenta with solid borders. The latter are computed by means of the void-galaxy cross-correlation function in redshift space, with a calibration-independent approach: the two nuisance parameters of the model are left free to vary and therefore are constrained by the data directly (instead of being fixed to the values calibrated with mock catalogues, i.e. the Flagship light-cone). The authors emphasize that the calibration-independent approach is to be preferred, as fixing the nuisance parameters to the mock values may introduce a prior dependence on the model parameters assumed in the mocks, possibly yielding biased cosmological constraints and underestimated relative uncertainties.

Finally, in the right panel of Fig. B.1 we show the comparison of our relaxed-calibration constraints with the Fisher forecasts computed with the pessimistic setting described in Euclid Collaboration (2020) for the probes weak lensing (WL, in orange) and spectroscopic galaxy clustering (GC$s$, in purple). According to the authors, the pessimistic configuration used for these constraints differs from the optimistic one mainly for a stronger cut of the maximum angular mode for the weak lensing angular power spectrum, $\ell_{\text{max}}(\text{WL}) = 1500$ (5000 in the optimistic setting), and of the power spectrum maximum scale, $k_{\text{max}}(\text{GC}s) = 0.25 \, h^{-1} \, \text{Mpc}^{-1}$ ($0.3 \, h^{-1} \, \text{Mpc}^{-1}$ in the optimistic setting). Analogously to Fig. 6, the presented confidence contours marginalised on the analysed parameter space are comparable and partially complementary.

Fig. B.1. Same as Fig. 6 but for different forecast settings. In this case the confidence contours obtained in this work from the void size function model (light-blue contours with dashed lines) are computed relaxing the constraints given by calibration parameters. The Euclid forecasts derived with void cross-correlation are computed with a model-independent approach, while IST forecasts are computed with the pessimistic setting described (Euclid Collaboration 2020).