



**HAL**  
open science

## Detecting cosmic voids via maps of geometric-optics parameters

Marius Peper, Boudewijn F Roukema, Krzysztof Bolejko

► **To cite this version:**

Marius Peper, Boudewijn F Roukema, Krzysztof Bolejko. Detecting cosmic voids via maps of geometric-optics parameters. *Monthly Notices of the Royal Astronomical Society*, 2023, 525 (1), pp.91-106. 10.1093/mnras/stad2246 . hal-04076329

**HAL Id: hal-04076329**

**<https://hal.science/hal-04076329v1>**

Submitted on 23 Apr 2024

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

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

# Detecting cosmic voids via maps of geometric-optics parameters

Marius Peper <sup>1</sup>★, Boudewijn F. Roukema <sup>1,2</sup> and Krzysztof Bolejko<sup>3</sup>

<sup>1</sup>*Institute of Astronomy, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Grudziadzka 5, PL-87-100 Toruń, Poland*

<sup>2</sup>*Univ Lyon, Ens de Lyon, Univ Lyon1, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69007 Lyon, France*

<sup>3</sup>*School of Natural Sciences, College of Sciences and Engineering, University of Tasmania, Private Bag 37 Hobart TAS 7001, Australia*

Accepted 2023 July 19. Received 2023 July 1; in original form 2023 April 2

## ABSTRACT

Curved space–time geometric-optics maps derived from a deep photometric survey should contain information about the 3D matter distribution and thus about cosmic voids in the survey, despite projection effects. We explore to what degree sky-plane geometric-optics maps can reveal the presence of intrinsic 3D voids. We carry out a cosmological  $N$ -body simulation and place it further than a gigaparsec from the observer, at redshift 0.5. We infer 3D void structures using the watershed algorithm. Independently, we calculate a surface overdensity map and maps of weak gravitational lensing and geometric-optics scalars. We propose and implement a heuristic algorithm for detecting (projected) radial void profiles from these maps. We find in our simulation that given the sky-plane centres of the 3D watershed-detected voids, there is significant evidence of finding corresponding void centres in the surface overdensity  $\Sigma$ , the averaged weak-lensing tangential shear  $\bar{\gamma}_\perp$ , the Sachs expansion  $\theta$ , and the Sachs shear modulus  $|\sigma|$ . Recovering the centres of the 3D voids from the sky-plane information alone is significant given the Sachs expansion  $\theta$ , or the Sachs shear  $|\sigma|$ , mildly significant given the weak-lensing shear  $\bar{\gamma}_\perp$ , and not significant for the surface overdensity  $\Sigma$ . Void radii are uncorrelated between 3D and 2D voids; our algorithm is not designed to distinguish voids that are nearly concentric in projection. This investigation shows preliminary evidence encouraging observational studies of gravitational lensing through individual voids, either blind or with spectroscopic/photometric redshifts. The former case – blind searches – should generate falsifiable predictions of intrinsic 3D void centres.

**Key words:** gravitational lensing: weak – methods: numerical – dark matter.

## 1 INTRODUCTION

Spectroscopic redshifts to determine the (comoving) 3D structures of cosmic voids require much more telescope resources than photometric surveys alone. While the simplest interpretation of a (single-filter) photometric survey is that it shows only the projected galaxy positions and shapes on the sky, the fact that the Universe is inhomogeneous implies that the photometric map contains information on the 3D distribution of inhomogeneities, with effects that are generically referred to as gravitational lensing. While arguments exist in classical mechanics for expecting light rays to be deflected as if they were massive particles (e.g. Schneider, Ehlers & Falco 1992, Historical Remarks), general relativity better justifies the response of light to gravity. Relativistic gravitational lensing theory was developed many decades ago in its generic form of geometric optics (Sachs 1961; Sasaki 1993). Some striking examples of gravitational lensing include detection of the twin quasar QSO 0957+561 A/B in 1979 (Walsh, Carswell & Weymann 1979; Young et al. 1980; Gorenstein et al. 1984, 1988) and giant luminous arcs (Paczynski 1987) and an Einstein cross (Adam et al. 1989) in the 1980s. For an in-depth review, see Bartelmann & Schneider (2001). While these are strong lensing effects, in this paper we focus on weak lensing, which primarily causes changes in the shape and flux of background

galaxies. Here, we argue that geometric-optics parameters that are derived from a deep photometric extragalactic map should contain information that can be used to detect some of the physical, 3D cosmic voids in the map, despite the fact that the voids are projected on the sky plane together with foreground and background voids. By carrying out an  $N$ -body simulation and analysing it, using an a priori reproducible software method, we aim to explore to what degree maps of photometrically derived geometric-optics parameters can reveal intrinsic 3D voids. We consider both the conventional approximation of weak-lensing parameters and Sachs optical scalars derived directly from the evolving gravitational potential.

The discovery of cosmic voids goes well back over 40 yr (Gregory & Thompson 1978; Jöeveer, Einasto & Tago 1978), from galaxy surveys that indicated that large regions appear to be devoid from galaxies, with galaxies being located primarily in structures that are usually described as walls, filaments, and clusters. Modern observations indicate that the comoving volume of our Universe is dominated by cosmic voids. For example, measurements based on the Sloan Digital Sky Survey (SDSS) indicate that a total fraction of roughly 60 per cent of the volume consists of cosmic voids (Pan et al. 2012). Cosmic voids have recently gained in interest, as they provide different characteristics for testing cosmological models to those provided by overdense structures. Li, Zhao & Koyama (2012) use simulations to study voids in modified gravity models, namely a fifth force that would affect the size and the density of voids. Dark energy is suspected to influence the shape of a void (Bos et al.

\* E-mail: [marius.peper@astro.uni.torun.pl](mailto:marius.peper@astro.uni.torun.pl)

2012). Studying the shape, density profile, size, and abundance of cosmic voids should yield crucial information about our Universe (e.g. Peebles 2001; Nadathur & Hotchkiss 2014; Pisani et al. 2015).

Multiple strategies for detecting voids in either observations or simulations exist. Since voids consist of the absence of luminous matter rather than its presence, this is a challenging task. Several different methods are commonly used for tracing these underdense regions of the cosmic web. Early routines used the assumption of spherical structures (e.g. Kauffmann & Fairall 1991) to detect voids. This was justified from the theoretical description of a void evolving out of a tophat-filtered density fluctuation; ellipsoidal initial density profiles were found to generally evolve to become more spherical (Gunn & Gott 1972; Lilje & Lahav 1991; Sheth & van de Weygaert 2004), and near-spherical void profiles were shown to be stable in the absence of shell crossings (Sussman & Bolejko 2012). The watershed void finders (e.g. Neyrinck 2008; Aragón-Calvo et al. 2010) make no assumptions on the shape of a void and is close to being parameter-free. This has become a de facto standard for determining the shapes of cosmic voids. Watershed mechanisms detect local minima in the density distribution of the cosmic web and identify underdense structures by searching for successively higher density contours, effectively finding the overdense edges of voids. The properties of the resulting voids depend on the spatial number density of the tracer particles used to represent the matter distribution. Nadathur & Hotchkiss (2015b) show that randomly subsampling the density of dark matter particles will tend to bias the void statistics, and suggest the use of halo occupation distribution models instead. To apply a watershed void finder to observational galaxy data, spectroscopic redshifts are needed.

Sánchez et al. (2017) introduce a method of detecting voids from a multifilter photometric survey by analysing redshift slices whose thickness is based on the photometric redshift uncertainties. The first detector variable that we analyse here results in a roughly comparable method. We detect structures in the surface overdensity, which, under the approximation of a constant mass-to-luminosity ratio, can be inferred from the observed photometric survey.

However, we are primarily interested in other sky-plane variables that can yield information on the large-scale structure of the Universe: the gravitational lensing signal. Due to the nature of dark matter having very weak interactions apart from gravity, we cannot measure the dark matter distribution of the Universe directly from electromagnetic surveys. Photometric galaxy surveys are generally thought to provide a fair proxy for the real projected matter distribution, but with many caveats. In contrast, the gravitational interaction of the dark matter distribution with photons can be measured via lensing effects. A light bundle that transverses a cosmological structure will experience shear and expansion and in particular, cosmic voids should leave an imprint on the shear and expansion. This should make it possible to reconstruct the underdensity field of voids based on the lensing signal, without no dependence on assumptions about baryon cooling or star formation. Large-scale maps of the gravitational lensing signal could thus, in principle, be used to detect cosmic voids. This would provide a method independent of using the projected spatial distribution of galaxies, since the lensing signal depends on the full underlying mass density, no matter whether it is luminous or not.

While in this work, we consider lensing parameters that are measurable from source galaxies far beyond the voids that we aim to detect, other observational methods of constructing the lensing signal have been proposed. Lewis & Challinor (2006) argue that since the cosmic microwave background (CMB) is lensed, the lensing potential can be reconstructed based on the observed CMB power

spectrum. Another method was proposed by Croft, Romeo & Metcalf (2018) to use Lyman  $\alpha$  forest observations to obtain lensing signals in the foreground of redshift slices of the forest. While our interest is in detecting lensing from individual voids, the lensing signal from underdensities in deep projections has been detected in Dark Energy Survey (DES) Science Verification data (Gruen et al. 2016), which motivates the search for lensing from individual voids.

In this work, we consider an idealized scenario in which we calculate detector variables without taking any detector errors into account, to see if it is possible to detect individual voids by their lensing signals. In reality, the derivation of the surface overdensity  $\Sigma$ , the averaged weak-lensing tangential shear  $\bar{\gamma}_T$ , the Sachs expansion  $\theta$ , and the Sachs shear modulus  $|\sigma|$  from observations involves many sources of random and systematic error, which will need to be taken into account in observational analyses. If our initial approach is successful, then the next step will be to make tests taking realistic observational noise and analysis into account.

We present a software pipeline (which aims to be fully reproducible on any unix-like operating system with sufficient RAM and disc space; Akhlaghi et al. 2021) to generate a cosmological  $N$ -body simulation, to detect galaxies and voids in it, and to ray-trace geometrics-optics parameters. The source package is provided as a frozen record at zenodo.8103985<sup>1</sup> and in live<sup>2</sup> and archived<sup>3</sup> GIT repositories. In Section 2.1 we briefly describe this overall pipeline, which extends that used in Peper & Roukema (2021).

We describe our simulation geometry in Section 2.2. We first detect intrinsic 3D voids using the watershed algorithm (Section 2.3.1). We independently try to detect voids in projection, ‘photometrically’ (in the absence of spectroscopic and photometric redshift information), from either the surface overdensity, conventional weak-lensing, or other geometric-optics signals. Our generic void profile search algorithm is defined in Section 2.3.2. We compare the sky-plane positions and radii of the photometric voids to those of the intrinsic 3D voids, using Monte Carlo simulations to check if this association is better than random (Section 2.4). In Section 2.5 we present our four detector variables, including a modification of our default void profile search algorithm specific to the weak-lensing shear, in Section 2.5.2). We present our results in Section 3, discuss these in Section 4, and conclude in Section 5. This version of the paper was produced with git commit d8952e7 of the source, after downloading, configuring, compiling, and running on a computer with a Little Endian x86\_64 architecture.

## 2 METHOD

### 2.1 Software pipeline

We use a highly reproducible software pipeline, following the Maneage template for reproducibility (Akhlaghi et al. 2021), that generates a realistic distribution of galaxies using a succession of several different cosmological tools. The software pipeline extends the galaxy formation pipeline presented in Peper & Roukema (2021). The pipeline includes a full simulation chain, starting with the generation of initial conditions with MPGRAFIC and running an  $N$ -body simulation with RAMSES (Teyssier 2002; Prunet et al. 2008). This simulation is processed as in Roukema, Quinn & Peterson (1993), Roukema et al. (1997), but using more recent software packages: dark

<sup>1</sup><https://zenodo.org/record/8103985>

<sup>2</sup><https://codeberg.org/mpeper/lensing>

<sup>3</sup>[swh:1:rev:169cb31e75ba6bcb69dd31741cf53f4c6795f91](https://doi.org/10.5281/zenodo.8103985)

matter haloes are detected and their merger-history tree is built with ROCKSTAR and CONSISTENT-TREES (Behroozi, Wechsler & Wu 2013a; Behroozi et al. 2013b) and semi-analytical galaxy formation recipes are applied with SAGE (Croton et al. 2016). We detect voids traced by the resulting spatial distribution of galaxies using a watershed void finder with REVOLVER (Nadathur et al. 2019). The coding of our sky-plane void-profile search algorithm (Sections 2.3.2 and Section 2.5.2) is original to this paper. The full details of the analysis are provided in a live git repository<sup>2</sup> and a frozen Zenodo record.<sup>1</sup> We used fixed pseudo-random seeds in most steps in the pipeline, as indicated in the configuration files. A detailed discussion of the pipeline can be found in Peper & Roukema (2021); here we primarily focus on new steps that are added to the analysis.

## 2.2 Simulation geometry

We use a simulation for a standard  $\Lambda$ CDM model (cold dark matter cosmological model with a cosmological constant  $\Lambda$ ) with Hubble constant  $H_0 = 70.0$ , current dark energy parameter  $\Omega_{\Lambda 0} = 0.7$ , and current matter density  $\Omega_{m0} = 0.3$ ,  $N_{\text{part}} = 256^3$  particles, and a comoving box size  $L_{\text{box}} = 120 \text{ Mpc } h^{-1}$ , where  $h := H_0 / (100 \text{ km s}^{-1} \text{ Mpc}^{-1})$ . This yields a dark matter particle mass of  $m_{\text{DM}} = 0.86 \times 10^{10} M_{\odot} h^{-1}$ . We require at least 10 particles to detect a halo. Since our simulated volume is a standard 3-torus ( $T^3$ ) simulation, we detect dark matter haloes, generate a merger-history tree, and detect intrinsic voids by interpreting the simulation's spatial section as  $T^3$ .

For computational convenience, for the 'observational' steps in which we detect voids in a sky-plane map of a detector variable, we interpret the (projected or ray-traced) simulated volume as the fundamental domain of a 2-torus ( $T^2 := S^1 \times S^1 \times \mathbb{R}$ ), where the two multiply connected directions lie in the sky plane. We informally refer to the fundamental domain as the 'box'. The foreground and background of the box, at lower and higher redshifts, respectively, are implicitly assumed to be a homogeneous (structure-free), simply connected  $\Lambda$ CDM background, i.e. they are assumed to be transparent and flat, with no effect on gravitational lensing. This simplification helps focus on the primary questions of our analysis; future analyses should include the effects of the full past time cone. We assume that this box is observed at high redshift, with its centre at redshift  $z' = 0.5$ , corresponding to a comoving distance of  $\chi_{\text{OL}} := 1322.0 \text{ Mpc } h^{-1}$  from the observer, which is a large distance compared to the box size. We use both an observer-centred Euclidean comoving coordinate system, with the observer at  $(x, y, z) = (0, 0, 0)$  and the centre of the simulated volume on the  $y$  axis of the  $\Lambda$ CDM simply connected space at  $(0, \chi_{\text{OL}}, 0)$ , and a simulation-centred system shifted by  $\chi_{\text{OL}}$ . To model light rays detected by the observer we convert  $(x, y, z)$  to  $(x = \chi \sin \theta' \cos \varphi', y = \chi \sin \theta' \sin \varphi', z = \chi \cos \theta')$ , where  $\chi$ , the comoving radial distance, together with  $\theta' \in [0, \pi]$  and  $\varphi' \in [0, 2\pi]$  are spherical coordinates of the spatial part of a flat  $\Lambda$ CDM model, and the simulation's centre is at  $(\chi = \chi_{\text{OL}}, \theta' = \pi/2, \varphi' = \pi/2)$ . We compute each of our detector variables on a grid with  $N_{\text{grid}} = 120^2$  'pixels' that we place on the middle plane of the box ( $y = 1322.0 \text{ Mpc } h^{-1}$ ), at  $(x, z)$  positions in the grid. We model light rays emitted from an observer-centred spherical surface near the back face of the simulated box ( $T^2$  slice), through to a second observer-centred spherical surface close to the front face of the box. The light rays's spatial paths are assumed to be straight in the non-perturbed space, i.e. with constant  $\theta'$  and  $\varphi'$ . We avoid approximately 5 per cent of the front and back parts of the box to minimize edge effects. Projected variables for a given pixel are computed along the line of sight of a light ray passing from the back spherical surface, through the pixel,

to the front spherical surface. In other words, the grid approximately corresponds to what is often referred to as 'the sky plane' for small solid angles, although it is (in this construction) a genuinely flat plane in comoving space. For brevity, we will use the term 'sky plane' as equivalent to this grid.

For the optical scalar calculations (Section 2.5.3), we trace our light rays geometrically under the assumption of a flat  $\Lambda$ CDM model, but calculate the optical scalars with a longitudinal Newtonian gauge approximation of an inhomogeneous model, with the line element

$$ds^2 = a^2 [-(1 + 2\Phi)d\tau^2 + (1 - 2\Phi)(d\chi^2 + \chi^2 d\Omega^2)], \quad (1)$$

where  $\tau$  is conformal time,  $\Phi$  is a potential, and  $d\Omega$  is the solid spherical angle element  $(d\theta')^2 + (\cos \theta' d\phi')^2$ .

## 2.3 Void detection

### 2.3.1 Intrinsic 3D voids

We detect intrinsic voids traced by the galaxy population using the void finder REVOLVER, which is based on the watershed void finder ZOBOV (Neyrinck 2008; Nadathur et al. 2019). The watershed mechanism in ZOBOV uses a Voronoi tessellation to estimate the densities at the particles' positions, is nearly parameter-free and makes no assumptions on the shape of the void. To characterize the size of a void, we use the effective radius  $R_{\text{eff}}$ , which is based on the sum over the volumes  $V_i$  of all the Voronoi cells that constitute a void, i.e.  $R_{\text{eff}} := \frac{3}{4\pi} (\sum_i V_i)^{1/3}$ . We adopt the geometric centroid of the set of cells that constitute an intrinsic void as the centre of that void. This is called the 'barycentre' in the REVOLVER code, but is mathematically the barycentre only if the void is assumed to be filled with a uniform density fluid (Peper & Roukema 2021, Section 1).

### 2.3.2 Photometric void detection

We detect voids from variables in the sky plane that are, in principle, observationally measurable in photometric surveys: the surface overdensity  $\Sigma$ , and three geometric-optics related parameters. We propose the following heuristically derived algorithm for detecting a projected void in a map of the surface overdensity  $\Sigma$  (defined below in Section 2.5.1) or the Sachs (1961) expansion  $\theta$  or shear  $\sigma$  (optical scalars, defined below in Section 2.5.3). Our algorithm for detecting voids from maps of the weak-lensing shear  $\gamma$  is similar, but differs in the ways that are described below in Section 2.5.2.

In contrast to the case for overdense extragalactic objects, we expect the (azimuthally averaged) radial density profile of a void in the sky plane, where the radius is  $r := \sqrt{(x - x_0)^2 + (z - z_0)^2}$  in our  $(x, z)$  grid for an object centre  $(x_0, z_0)$ , to have its lowest values in the centre of the void and a sharp maximum at the void's edge. The Sachs expansion  $\theta$  and shear  $\sigma$  should also have a minimum at the centre of a void and a maximum at the edge. While projection effects for overdense structures are a perennial problem in astronomy (e.g. for determining whether a galaxy group is dynamically real or a chance projection), the projection effects can be expected to be much worse for voids, since voids dominate the volume of the Universe, implying stronger overlaps. In contrast to overdense objects, spectroscopic redshift determination for the rare galaxies in voids is unlikely to be effective in separating a chance projection of voids from a void that is real in three spatial dimensions, and is likely to be a statistically unstable way of dynamically characterizing a void. Moreover, voids, in general, are not perfectly spherical objects, making detection via templates unlikely to be easy. Nevertheless, projection along the line



of sight should provide a modest effect of symmetrization, and by appropriately averaging, we hypothesize that detection is feasible.

We define an isotropized (azimuthally averaged) variable  $\bar{X}$ , i.e. the average on a circle in the  $(x, y \equiv \chi_{\text{OL}}, z)$  grid plane, at radius  $r$  and centred on a pixel  $j$ , i.e.

$$\bar{X}_j(r) = (2\pi)^{-1} \int_0^{2\pi} X(r, \varphi) d\varphi, \quad (2)$$

where  $\varphi$  is the angle around the circle centred on position  $j$  in the grid and  $X$  is either the surface overdensity  $\Sigma$ , or one of the optical scalars  $\theta$  or  $\sigma$  (for the weak lensing shear  $\gamma$ , see equation (15) below). To estimate  $\bar{X}_j(r)$ , we sample the grid values at even intervals around a circle of radius  $r$ , with intervals that give at least one value per  $\text{Mpc } h^{-1}$ , we smooth the values with a third-order Savitzky–Golay filter (Savitzky & Golay 1964), and integrate.

We also define a disc-averaged profile  $\bar{\bar{X}}$  on the disc internal to a given radius (not weighted by the radius) by integrating equation (2) and appropriately normalizing, i.e.

$$\bar{\bar{X}}(r) = \frac{\int_0^r \int_0^{2\pi} X(r', \varphi) d\varphi dr'}{(2\pi)^{-1} \int_0^r \int_0^{2\pi} d\varphi dr'} = \frac{\int_0^r \bar{X}(r') dr'}{r}, \quad (3)$$

and a disc-averaged absolute slope,

$$\left| \bar{\bar{X}}' \right| (r_i) = \frac{\sum_{i' < i} |\text{d}\bar{X}/\text{d}r| (r_{i'})}{i-1}, \quad (4)$$

where the index  $i$  indicates radial discretization in estimating  $\bar{X}$  and  $\bar{\bar{X}}' := \text{d}\bar{\bar{X}}/\text{d}r$ .

Apart from the case of the weak-lensing shear  $\gamma$  (Section 2.5.2), we expect  $\bar{X}$  (and thus  $\bar{\bar{X}}$ ) to increase monotonically from the centre at  $r = 0$  outwards as  $r$  increases, though the projection against other voids, voids' asphericity, and noise will make this monotonicity difficult to detect. For each pixel  $j$  in our sky plane, we define a heuristic selection criterion  $\eta_X$  motivated by the expected monotonicity as follows.

(i) Ignore each pixel  $j$  with  $X(r=0) > X_{\text{median}}$ , where  $X_{\text{median}}$  is the median over all pixels in the sky plane. The motivation is that for any  $X$ , pixels with  $X(r=0) > X_{\text{median}}$  are unlikely to correspond to the centre of a void. The projected or ray-traced variables  $X$  should have their minima at voids' centres. This step should remove many pixels unlikely to be void centres. In principle, this step might remove some pixels that are genuinely the centres of intrinsic voids, especially in the case of voids whose underdensities are shallow intrinsically or in projection against foreground or background overdensities. At the cost of increased computation time, removal of this step might lead to detecting more voids. However, in practice we expect that the gain would be minor, since these types of voids will rarely have the projected void-like profiles that we aim to detect.

(ii) For a given pixel  $j$ , for each radial distance  $r_i$  from the pixel, calculate the circular average  $\bar{X}$  as in equation (2), where for simplicity we omit the subscript  $j$ . We set the interval in  $r_i$  to be smaller than  $1 \text{ Mpc } h^{-1}$  in order to be sensitive to small-scale structure.

(iii) A persistently positive strong positive slope in  $\bar{X}$  is detected as follows. Find  $i_1$ , the first radial position  $i$ , with respect to pixel  $j$ , where all three of the conditions

$$\bar{X}'(r_i) > \left| \bar{\bar{X}}' \right| (r_i) \quad (5)$$

$$\bar{X}(r_i) - \bar{X}(r_{i-1}) > f_{\text{std}} \left\langle \left( \bar{X}(r_{i'}) - \bar{\bar{X}}(r_{i'}) \right)^2 \right\rangle_{i' < i}^{1/2} \quad (6)$$

$$\bar{X}(r_i) > f_{\text{mean}} \bar{\bar{X}}(r_i) \quad (7)$$

**Table 1.** Parameters used in our 2D void detection algorithm in equations (6) and (7).

	$f_{\text{std}}$	$f_{\text{mean}}$
$\Sigma$	0.50	1.50
$\theta$	0.30	1.30
$\sigma$	0.30	1.10

are satisfied over four successive steps  $i-3, i-2, i-1, i$ , where  $\langle \cdot \rangle$  is the mean and  $f_{\text{std}}$  and  $f_{\text{mean}}$  are heuristically chosen fractions. The aim of criteria (5) and (6) is to find a range of radii where the slope  $\bar{X}'$  has a stable and significant increase, i.e. where positive second derivatives  $\bar{X}'' > 0$  are numerically persistent. Criterion (7) aims to also require the slope  $\bar{X}'$  to be sufficiently positive. The values adopted for  $f_{\text{std}}$  and  $f_{\text{mean}}$  are given in Table 1.

(iv) Find  $i_2$ , the first local maximum in  $\bar{X}$  for  $i > i_1$ , i.e. the first local maximum after the persistently strong positive slope condition that determines  $i_1$ .

(v) Define an initial void selection criterion  $\eta_X^0(j) := 1/r_{i_2}$ . The radius  $r_{i_2}$  is the estimated radius of the candidate void.

(vi) Steps (ii)–(v) are carried out for all pixels  $j$  accepted in step (i) (with (iii), (iv) modified in the case of  $\gamma$ ; see Section 2.5.2). In order to cope with the very noisy data, we define a smoothed selection criterion for further use,  $\eta_X(j)$ , as a low-pass triangular filter (weighted mean) of the 25  $\eta_X^0(j)$  values in a  $5 \times 5$  grid of pixels centred on pixel  $j$ .<sup>4</sup> We then find all local minima of the selection criterion  $\eta_X(j)$  over the pixels  $j$  as follows.

(vii) We select a void centred at pixel  $j$  if it dominates its local region in the sense that  $\eta_X(j) < \eta_X(k)$ , where  $k$  indexes pixels in a square grid centred on pixel  $j$  and extending 5 pixels in each of the  $\pm x$  and  $\pm z$  directions. As an extra credibility criterion, selection of a void is only accepted if  $\eta_X(j) < 0.90 \sum_k \eta_X(k) / \sum_k 1$ , where  $k$  indexes all pixels in the map. This algorithm results in a list of projected voids with centres  $j$  and radii  $r_{i_2}(j)$  that represent the largest locally credible voids.

(viii) To avoid cases where a single genuine void is misidentified as two slightly offset voids, we check if two or more centres are closer to one another than

$$\min \{ r_{i_2}(j_1)/4, r_{i_2}(j_2)/4, 10.0 \text{ Mpc } h^{-1} \}. \quad (8)$$

In these cases, we merge these voids into a single void. The new centre and radius of the merged void are defined as the mean of the centres and radii of the unmerged voids.

## 2.4 Matches to intrinsic voids

To quantify whether the photometric detection of voids – 2D voids – successfully finds the intrinsic 3D voids, we first define a heuristically motivated matching criterion to find the best matches, and then compare the set of best matches to an equivalent set of best matches when the list of 2D void parameters is generated randomly (positions) or randomly shuffled (radii). This aims at answering two different questions: (i) given a detected set of 2D voids, are these better than a random set of 2D voids at revealing true 3D voids? (ii) given a set of intrinsic 3D voids, do the detected 2D voids better match these (numerically) real voids better than a random set of 2D voids would? The former question is interesting for observational detection of 3D voids from photometric or other geometric-optics data; the latter is

<sup>4</sup>This step interprets the box as an isolated box, not  $T^2$ , and sets  $\eta_X(j)$  near the borders of the box to a high value to prevent finding minima there.

interesting for using spectroscopically defined 3D voids to motivate searches for gravitational lensing by voids.

#### 2.4.1 Best matches and median sky-plane separation $\mu_{x,z}$

We define the probability of the  $i$ th 2D void being a match to the  $j$ th 3D (watershed) void by first defining the probabilities that the  $x$  and  $z$  positions are close in the  $T^2$  sense and that the radii are logarithmically close. We set a cumulative Gaussian probability that the  $x$  and  $z$  positions for variable  $X \in \{\Sigma, \gamma, \theta, \sigma\}$  are closer to each other than the estimated values,

$$P_{(x,z)_{i,j}}^X = 1 - \operatorname{erf} \frac{d((x,z)_i^X, (x,z)_j^X)}{\sqrt{2}\sigma_{x,z}}, \quad (9)$$

where  $\operatorname{erf}$  is the error function,  $d(\dots)$  is the  $T^2$  minimum distance in the  $x$ - $z$  plane, and  $\sigma_{x,z} = 5.0 \text{ Mpc } h^{-1}$ . Similarly, we set

$$P_{R_{i,j}}^X = 1 - \operatorname{erf} \frac{|\log_{10}(R_i^X/R_j^X)|}{\sqrt{2}\sigma_{\log_{10} R}}, \quad (10)$$

where  $\sigma_{\log_{10} R} = 0.30$ . These are assumed, for simplicity, to be independent probabilities, giving a heuristic overall probability that the  $i$ th 2D void matches the  $j$ th 3D void

$$P_{i,j}^X = P_{(x,z)_{i,j}}^X P_{R_{i,j}}^X. \quad (11)$$

For question (i) (Section 2.4), given a 2D void  $i$ , we find the 3D void  $j$  with the highest matching probability  $P_{i,j}^X$ , for detector variable  $X$ . This does not exclude the possibility that two different 2D voids best identify with the same 3D void. For a set of  $N_{2D}$  detected 2D voids, this gives us a matched set of  $N_{2D}$  objects, which have both 2D and 3D sky position and radius information, presumed to match.

For each object in this set, we calculate the  $T^2$  distance between the 2D and 3D  $(x, z)$  positions and from the distribution of these values, calculate  $\mu_{x,z}(3D|2D)$ , the median distance for a 3D match given a 2D match. In calculating this median, in cases where a single 3D void is the best match for two or more 2D voids, we only consider the match in which  $P_{i,j}^X$  is maximized.

For question (ii), given a 3D void  $i$ , we find the 2D void  $j$  with the highest matching probability  $P_{i,j}$ . Again, this does not exclude the possibility that two different 3D voids best identify with the same 2D void. In practice, since we find fewer 2D voids to 3D voids, there are necessarily cases where multiple 3D voids identify with a single 2D void. For a set of  $N_{3D}$  detected 3D voids, this gives us a matched set of  $N_{3D}$  objects, which have both 2D and 3D sky position and radius information, presumed to match.

For each object in this set, we calculate the  $T^2$  distance between the 2D and 3D  $(x, z)$  positions and infer  $\mu_{x,z}(2D|3D)$ , the median distance for a 2D match given a 3D match, again using the highest  $P_{i,j}^X$  to reduce non-one-to-one matches.

#### 2.4.2 Comparison to matches for random 2D voids

We generate a Monte Carlo simulation of 2D voids by choosing  $N_{2D}$  pairs  $(x, z)$  from a uniform random distribution within the  $x$  and  $z$  ranges of the 2D grid. For both questions (i) and (ii) (separately), for each Monte Carlo simulation, we find matched sets using the same algorithm as above.

To answer question (i) we estimate  $\mu_{x,z}(3D|2D)$  for a given detector  $X$ , for both the original matched set and for each of the simulated matched sets. This yields  $P_{x,z}^X(3D|2D)$ , the frequentist probability that the original  $\mu_{x,z}(3D|2D)$  is less than  $\mu_{x,z}(3D|2D)$  from the Monte Carlo simulations. In other words,  $P_{x,z}^X(3D|2D)$  is the probability that, given the 2D voids, the matches with 3D voids

in the sky plane are no better than those drawn from a Monte Carlo simulation.

Similarly, for question (ii), the frequency with which the original  $\mu_{x,z}(2D|3D)$  is less than the values of  $\mu_{x,z}(2D|3D)$  from the simulations yields  $P_{x,z}^X(3D|2D)$ , the probability that, given the 3D voids, the matches with 2D voids in the sky plane are no better than those drawn from a Monte Carlo simulation of sky positions.

Cases where 3D voids are concentric or approximately overlap in projection will yield only a single 2D void using our algorithm, and are likely to make estimation of the radii difficult. The range of values of the radii are not as conveniently constrained as the  $(x, z)$  centres of the voids. Rather than choosing an arbitrary range for a Monte Carlo simulation, we use a non-parametric method. We define  $P_R^X(3D|2D)$  as the two-sided probability that the Spearman  $\rho$  rank correlation coefficient (Spearman 1904) of the matched set of  $N_{2D}$  values  $R_i^X$  and  $R_{j(i)}^X$  is stronger (positive or negative) than what it would be for a set of paired values where one set is randomly permuted.

Similarly, we define  $P_R^X(2D|3D)$  as the two-sided probability that the Spearman  $\rho$  ranking coefficient of the matched set of  $N_{3D}$  values  $R_i^X$  and  $R_{j(i)}^X$  is stronger (positive or negative) than it would be under random permutations.

### 2.5 Detector variables $\Sigma, \gamma, \theta, \sigma$

Here we define and describe our detector variables  $X \in \{\Sigma, \gamma, \theta, \sigma\}$ , where here we write the generic forms of these variables for simplicity; the more specific forms are given below. These detector variables can, in principle, be derived from a photometric map, given some minimal assumptions, such as a mass-to-light ratio in the case of the surface overdensity  $\Sigma$ , or statistically isotropic distributions of galaxy shape parameters in the case of the other three parameters. We derive each of these from the particle distribution, not from the galaxy distribution. We include  $\Sigma$  since apart from requiring a mass-to-light ratio assumption, it is the simplest to derive from a photometric map.

#### 2.5.1 Surface overdensity $\Sigma$

We calculate the surface overdensity by integrating the overdensity  $\rho - \bar{\rho}$  along the line of sight, neglecting temporal evolution. (Temporal evolution is taken into account with the optical scalar modelling; see Section 2.5.3 below.) Densities are constructed for each particle using a Voronoi tessellation followed by linear interpolation. For a flat model, the surface overdensity in direction  $\hat{n}$  is

$$\Sigma(\hat{n}) = \int_{\chi_{\min}}^{\chi_{\max}} (\rho(\hat{\chi}, \Omega) - \bar{\rho}) d\chi, \quad (12)$$

where  $\chi_{\min} = \chi(z_0) - 0.95L_{\text{box}}/2$  and  $\chi_{\max} = \chi(z_0) + 0.95L_{\text{box}}/2$ . The 0.95 factor neglects the 5 per cent front and back parts of the box to minimize edge effects (Section 2.2). Since we expect the surface overdensity  $\Sigma$  to be negative in a void, we aim to detect it in places where it is physically a surface underdensity, i.e. a projected mass deficit.

**Detection strategy with  $\Sigma$ :** Voids typically have strong underdensities in their interior, so the 2D projection of a void should still show a strong underdensity in the interior after projection. Thus, we search for local minima in  $\Sigma$ .

The projection of foreground and background voids and their walls (in reality, clusters, filaments, and walls) will, to some degree, obscure this search. To the extent that the obscuration can be statistically neglected or removed, the azimuthally averaged radial profile  $\bar{\Sigma}$  (equation 2) should show a slow increase in  $\bar{\Sigma}$  from the centre of a projected void outwards until it nears the (projected)

wall, when a rapid increase should occur, followed by a drop as  $\bar{\Sigma}$  asymptotes to the mean of the environment surrounding the void. Thus, a local maximum in  $\bar{\Sigma}$  at the wall that surrounds the void should be sought. We expect that larger voids should yield clearer signals.

Some pixels, likely containing galaxy clusters or projections of galaxy filaments, were found to be highly overdense, misleading our algorithm's search for void walls because of these overdense pixels' strong influence on  $\bar{\Sigma}$ . To reduce the influence of these extreme pixels, prior to step (i), we truncate  $\Sigma$  values at the 90th percentile of their distribution.

Substructures of overdensities exist inside of voids, similar to the larger scale overdensities of the cosmic web, but traced by dark matter haloes of much lower mass (Gottlöber et al. 2003). These substructures contribute another obscuring factor that should weaken our proposed detection algorithm (Section 2.3) using  $\Sigma$ .

### 2.5.2 Weak-lensing tangential shear $\bar{\gamma}_\perp$

Weak gravitational lensing information is typically extracted from observations by using the distortion of observed images that is induced by cosmological inhomogeneity, with the aim of tracing the spatial distribution of dark matter. We follow the mathematical descriptions of Bartelmann & Schneider (2001), Krause et al. (2013), and Kilbinger (2015). We derive the parameters of this idealized model from the surface overdensity calculated in our cosmological simulation.

We represent the lens plane (Bartelmann & Schneider 2001, fig. 11) with two orthogonal spatial directions with indices  $a$  and  $b$ ; the direction of propagation of the light bundle as it would arrive at the observer from the source if unlensed, represented as a vector in the lens plane,  $\Theta_S$ ; and the direction at which the light bundle reaches the observer after lensing, again a vector in the lens plane,  $\Theta_O$ . Assuming that the rotation of the image vanishes, the Jacobian matrix can be decomposed into the shear  $\gamma$  and the convergence  $\kappa$  (Bartelmann & Schneider 2001, equation 3.11):

$$A = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}. \quad (13)$$

The convergence  $\kappa$  at a generic position in the sky plane can be evaluated as  $\kappa(\hat{n}) = \frac{\Sigma(\hat{n})}{\Sigma_{\text{crit}}}$ , where  $\Sigma_{\text{crit}} = \frac{c^2}{4\pi G} \frac{D_{\text{OS}}}{D_{\text{OL}}D_{\text{LS}}}$  (Bartelmann & Schneider 2001, equation 3.7). The values  $D_{\text{XX}}$  are the angular diameter distances between the observer (O), the lens (L), and the source (S). For simplicity, we assume the observer–lens and lens–source radial comoving distances to be equal, i.e.  $\chi_{\text{OL}} = \chi_{\text{LS}}$ , giving  $\chi_{\text{OS}} = 2\chi_{\text{OL}}$ . We do not attempt void detection with  $\kappa$ , as the result would be equivalent to using  $\Sigma$ .

For our detection strategy we use  $\bar{\Sigma}(r, \hat{n})$ , the isotropized (ring averaged) form of  $\Sigma$  (equation 2) with respect to a given centre  $\hat{n}$  of a possible void, applying a 90 per cent cut in the line-of-sight direction as explained in 2.5.1, and subtract it from the surface overdensity averaged within a disc centred on  $\hat{n}$ , yielding

$$\Delta\Sigma(r, \hat{n}) := \tilde{\Sigma}(r, \hat{n}) - \bar{\Sigma}(r, \hat{n}), \quad (14)$$

where  $\tilde{\Sigma}(r, \hat{n})$  is defined in equation (16). As in Krause et al. (2013, equations 4 and 5) and also derived in Kilbinger (2015, equations 40–47), the mean of the tangential component of the shear internal to a ring at  $r$  can then be evaluated as

$$\bar{\gamma}_\perp(r, \hat{n}) = \frac{\Delta\Sigma(r, \hat{n})}{\Sigma_{\text{crit}}}. \quad (15)$$

The disc average  $\tilde{\Sigma}$  calculated by integrating (2) over the radius, using the usual weighting and now leaving the centre  $\hat{n}$  implicit, is

$$\tilde{\Sigma}(r) = \frac{\int_0^r \int_0^{2\pi} \Sigma(r', \varphi) r' d\varphi dr'}{\int_0^r \int_0^{2\pi} r' d\varphi dr'} = \frac{2\pi \int_0^r \bar{\Sigma}(r') r' dr'}{\pi r^2}. \quad (16)$$

**Detection strategy with  $\bar{\gamma}_\perp$ :** By definition,  $\bar{\gamma}_\perp(r)$  should be close to zero at  $r = 0$ , the centre of a void, and should decrease to a sharp minimum where  $r$  is the radius of the void's (statistical) wall. At greater radii, both the azimuthally averaged surface overdensity  $\bar{\Sigma}$  and the disc-averaged  $\tilde{\Sigma}$  should approach zero, so  $\bar{\gamma}_\perp(r)$  should also increase up to zero. The minimum in  $\bar{\gamma}_\perp(r)$  should reveal the edge of the void.

Since this qualitative behaviour of  $\bar{\gamma}_\perp(r)$  differs from the other detector variables considered, we modify steps (iii) and (iv) of the algorithm of Section 2.3.2 as follows.

Since  $\bar{\gamma}_\perp$  and  $\bar{\gamma}_\perp'$  calculated according to (2) are noisy, we apply extra smoothing, replacing  $\bar{\gamma}_\perp(r_i)$  and  $\bar{\gamma}_\perp'(r_i)$  by  $\langle \bar{\gamma}_\perp(r_i) \rangle_{\{\max(0, i-3), \dots, i+3\}}$  and  $\langle \bar{\gamma}_\perp'(r_i) \rangle_{\{\max(0, i-3), \dots, i+3\}}$ , respectively. This smoothing reduces the role of local fluctuations in the dark matter distribution.

(i) After this smoothing, we search for the radial distance where  $\bar{\gamma}_\perp$  starts dropping sharply, i.e. the index  $i_1$  is the first value  $i$  where  $\bar{\gamma}_\perp(r_i) < \bar{\gamma}_\perp(r_{i-1})$ .

(ii) The radial distance just past the wall is sought as the radial distance where  $\bar{\gamma}_\perp$  increases sharply, i.e. the index  $i_2$  is the first value  $i > i_1$  where  $\bar{\gamma}_\perp(r_i) > \bar{\gamma}_\perp(r_{i-1})$ .

(iii) In addition, to remove choices of a void centre where the best 'wall' found this way has a weak density contrast, we dismiss the candidate detection if  $|\bar{\gamma}_\perp| / \max(|\bar{\gamma}_\perp|) < 0.1$ . For patterns in  $\bar{\gamma}_\perp$  that have almost no significant features, this criterion avoids interpreting a nearly flat curve  $\bar{\gamma}_\perp(r_i)$  as a candidate void.

If both  $r_{i_1}$  and  $r_{i_2}$  are detected, then we continue to step (v) as above (Section 2.3.2). Even if pixel  $j$  is correctly centred on a void's centre, this algorithm for  $\bar{\gamma}_\perp$  can fail to detect  $r_{i_2}$  if the (projected) environment just outside the void's wall includes strong fluctuations. In the case of failure to detect  $r_{i_2}$ , the pixel is considered invalid at step (v) and dropped from further consideration.

### 2.5.3 Optical scalars $\theta$ and $|\sigma|$

We calculate optical scalars following Sasaki (1993). In principle, the optical scalars should model the real Universe more accurately than the weak-lensing scalars described above, since fewer assumptions are required (Clarkson et al. 2012, Section III.B) – the real Universe is inhomogeneous rather than strictly Friedmann–Lemaître–Robertson–Walker (FLRW). Clarkson et al. (2012) give a detailed derivation of the relation between the optical scalars  $\theta$  and  $\sigma$  and the usual weak-lensing scalars, in particular in their equations (41)–(43), with slightly differing conventions. For example, the Sachs shear  $\sigma$  includes information not only from the weak lensing shear  $\gamma$ , but also from the evolution of  $\gamma$  with respect to the affine parameter along a null geodesic. While the differences in practice may not be large, especially in the FLRW context, the advantage of modelling the Sachs optical scalars is to prepare for future work that simultaneously takes into account both structure formation and expansion of the Universe (e.g. Heinesen 2023, and references therein). For simplicity, in the weak lensing case we model the commonly studied tangential weak lensing shear  $\bar{\gamma}_\perp$ , while in the Sachs optical scalar case we model the shear modulus  $|\sigma|$ , as explained below. In the Newtonian approximation, the Ricci tensor can be written as

$$R_{00} \approx 8\pi G \rho \omega^2 \quad (17)$$

and the Weyl tensor components of interest are

$$C_{A0B0} \approx \left( 2\Phi_{;AB} - \delta_{AB}\Phi_{;C}^C \right) \omega^2 \quad (18)$$

$$= \left( 2\Phi_{;\mu\nu} e_A^\mu e_B^\nu - \delta_{AB} \delta^{CD} \Phi_{;\mu\nu} e_C^\mu e_D^\nu \right) \omega^2, \quad (19)$$

(Sasaki 1993, equations 3.22 and 3.21) where  $G$  is the gravitational constant; space and time units are related by  $c = 1$ ;  $\rho$  is density;  $\Phi$  is the gravitational potential;  $\omega = -k_\mu u^\mu = 1 + z_{\text{redsh}}$ , for an observer four-velocity  $u^\mu$ , light propagation one-form  $k_\mu$ , and redshift  $z_{\text{redsh}}$ ;  $\{e_A, e_B\}$  or  $\{e_C, e_D\}$  are a pair of dyad basis vectors; and  $\delta_{AB}$  is the Kronecker delta ( $\delta_{AB} = 1$  if  $A = B$ ,  $\delta_{AB} = 0$  if  $A \neq B$ ). The dyad basis vectors  $e_A, e_B$  span the space-like 2-plane that is orthogonal to the spatial path of the light ray that points from the observer towards the direction of a cell of the grid. We use the Gram–Schmidt method to construct  $e_A$  and  $e_B$ .

The optical scalars – the expansion  $\theta$  (real) and the shear  $\sigma$  (complex) – are related to each other and the Weyl tensor (equation 19) via the coupled pair of differential equations

$$\frac{d}{dv}\theta = -R_{00} - 2|\sigma|^2 - \frac{1}{2}\theta^2 \quad (20)$$

and

$$\frac{d}{dv}\sigma = -(C_{1010} + iC_{1020}) - \sigma\theta. \quad (21)$$

For a visualization of the effect of  $\theta$  and  $\sigma$  on a light bundle we refer to Sasaki (1993, fig. 4), where  $\sigma = \sigma_+ + i\sigma_x$ .

**Detection strategy with  $\theta$  and  $|\sigma|$ :** Both the expansion  $\theta$  and the modulus of the shear,

$$|\sigma| = \sqrt{\text{Re}(\sigma)^2 + \text{Im}(\sigma)^2} \quad (22)$$

should be closely related to the surface overdensity, since integrations along paths approximately (spatially) orthogonal to the lens plane are performed in all three cases. However, these are not exactly analogous. Not only are these distinct physical quantities, but the overdensity integral is performed parallel to the  $y$  axis, while for each pixel in our 2D grid plane, we estimate  $\theta$  and  $|\sigma|$  along a spatially straight path from the observer through the pixel, i.e. only approximately parallel to the  $y$  axis.

In practice, initial numerical exploration shows that  $\theta$  and  $|\sigma|$  behave qualitatively like  $\Sigma$ , in that they start from a low value at the centre of a void and increase to a sharp maximum at a void boundary. Thus, we use the same search algorithm for finding voids in maps of  $\theta$  and  $|\sigma|$  as indicated above in Section 2.3.2, with slightly adjusted parameters (Table 1). While qualitatively similar in numerical terms, the physical meanings of these parameters differ. The optical scalars  $\theta$  and  $|\sigma|$ , if derived from observations, represent the underlying matter distribution with no dependence on observed luminosity and without the simplifying assumptions of the weak-lensing approximation. In particular, the weak-lensing shear  $\bar{\gamma}_\perp$  (equation 15) is an average defined with respect to a hypothesized void centre, while  $\theta$  and  $|\sigma|$  provide maps prior to assumptions about void centres.

## 3 RESULTS

### 3.1 Simulation

We performed an ab initio simulation and detected voids as described above. As indicated in Table 2, we detected  $N_{3D} = 46$  voids in the galaxy population with the watershed mechanism, and smaller numbers of 2D voids using  $\Sigma$ ,  $\bar{\gamma}_\perp$ ,  $\theta$ , and  $|\sigma|$  from the projected density distribution and by ray-tracing through the evolving gravitational potential  $\Phi$ .

**Table 2.** Numbers of intrinsic 3D voids detected with REVOLVER,  $N_{3D}$ , and in the 2D grid,  $N_{2D}^\Sigma$ ,  $N_{2D}^\gamma$ ,  $N_{2D}^\theta$ , and  $N_{2D}^\sigma$ , using the surface overdensity  $\Sigma$ , the weak-lensing shear  $\bar{\gamma}_\perp$ , the Sachs expansion  $\theta$ , and the modulus of the Sachs shear  $|\sigma|$ , respectively.

$N_{3D}$	$N_{2D}^\Sigma$	$N_{2D}^\gamma$	$N_{2D}^\theta$	$N_{2D}^\sigma$
46	28	29	34	39

**Table 3.** Probability that the matches between 3D and 2D voids for detector variable  $X$  are no better than those of randomly generated 2D voids,  $P_{x,z}^X(3D|2D)$  when given 2D voids; or  $P_{x,z}^X(2D|3D)$  when given 3D voids; and probability that the Spearman rank correlation coefficient for the radii of matched 3D and 2D voids for detector variable  $X$  could be that of a set of randomly paired values,  $P_R^X(3D|2D)$  when given a 2D void; and  $P_R^X(2D|3D)$  when given a 3D void. See Section 2.4.2. Plain-text version available at zenodo.8103985/void\_match\_analysis.dat.

$X$	$P_{x,z}(3D 2D)$	$P_{x,z}(2D 3D)$	$P_R(3D 2D)$	$P_R(2D 3D)$
$\Sigma$	0.027	0.0038	0.89	0.94
$\bar{\gamma}_\perp$	0.010	$3.0 \times 10^{-5}$	0.85	1.0
$\theta$	0.00050	$3.0 \times 10^{-5}$	0.25	0.16
$ \sigma $	0.00014	$1.0 \times 10^{-5}$	0.27	0.61

Table 3 shows the probabilities, defined in Section 2.4.2, that quantify the significance of: (i) a detected 2D void revealing the existence of an intrinsic 3D void via its sky plane position or radius,  $P_{x,z}^X(3D|2D)$  or  $P_R^X(3D|2D)$ , respectively, and (ii) an intrinsic 3D void implying that its 2D projection is detectable,  $P_{x,z}^X(2D|3D)$  or  $P_R^X(2D|3D)$ . In each case, these represent the probability that the estimated correspondence between the populations could occur by chance, given prior information on the number of 2D voids (for positions) or non-parametrically (for radii).

Since voids in  $N$ -body simulations are characterized by small numbers of particles, the detection of individual voids, whether in the 3D galaxy distribution or by a 2D detector in variables derived from the particle distribution, is in general numerically sensitive to small changes in machine arithmetic. We performed a small number of independent full-pipeline simulations, retaining the same pseudo-random number seed, to investigate this qualitatively. The re-simulated equivalent of the values listed in Table 3 shows small variations with re-simulation on a given machine, and stronger variation between different machines. We describe our results taking into account our small-scale estimates of their reproducibility, and use the word ‘robust(ly)’ to indicate cross-machine reproducibility. In principle, cosmic variance can be explored by varying the random seeds. In Appendix A, we show the equivalent of Table 3 for a small number of full-pipeline simulations with alternative random number seeds. The specific levels of significance vary somewhat with cosmic variance.

We find  $P_{x,z}^X(2D|3D)$  values (robustly) indicating significant match distributions in all four cases, with  $P_{x,z}^\Sigma(2D|3D) < 0.01$ , and  $P_{x,z}^X(2D|3D) < 0.0001$  for  $X \in \{\bar{\gamma}_\perp, \theta, |\sigma|\}$ . Thus, we find that given the 3D voids found with the watershed algorithm in the galaxy distribution, the sky plane positions of the 2D voids found using the surface overdensity  $\Sigma$  are significantly closer to the former than they would be if the same number of 2D void positions were chosen randomly. In other words, we have a significant response to question (ii) for  $\Sigma$ . This is reassuring, because it shows that despite the projection effects of multiple voids and their aspherical shapes, the centres of the intrinsic 3D voids can be recovered in the 2D  $\Sigma$  distribution.



Moreover, we find that for the weak-lensing tangential shear  $\overline{\gamma}_\perp$ , and for both the Sachs optical scalar expansion  $\theta$  and shear  $|\sigma|$ , the centres of the 2D voids represent the 3D void centres to high significance. Thus, any of the four parameters should be usable to re-detect the void centres known from the 3D voids.

In contrast, if we start with the 2D photometric map and predict the centres of the 3D voids, we only find (Table 3) the weak-lensing tangential shear  $\overline{\gamma}_\perp$ , the Sachs expansion  $\theta$  and the Sachs absolute shear  $|\sigma|$  to significantly and robustly reveal underlying 3D voids, with  $P_{x,z}^X(3D|2D) \lesssim 0.01$  in all three cases. Comparison with  $P_{x,z}^\Sigma(3D|2D)$  in Table 3, for the surface overdensity, shows that discovering a 3D void thanks to its 2D signature is less likely with  $\Sigma$ . In other words, in answering question (i), use of our algorithm with any of the three geometrical optics parameters is more likely to reveal the sky-plane position of the 3D void than using  $\Sigma$ .

These results show that the intrinsic 3D void signal yields detectable void centres with our algorithm in not only the projected (2D) surface overdensity  $\Sigma$ , inferrable from photometric maps with only a mass-to-light ratio assumption, but also in the 2D maps of weak-lensing and Sachs optical shear parameters. If additional information, such as spectroscopic or photometric redshift information, is available, then combining that information with lensing analyses of the data should lead to tighter constraints on the (partly invisible) underdensity distributions, as opposed to using galaxies' sky positions and redshifts alone.

Moreover, in the absence of galaxy redshift information, 2D maps should yield constraints on the mass distribution, at least in the case of  $\overline{\gamma}_\perp$  and  $\theta$ . However, while the void centres are detected, the radii are poorly constrained from either 3D or 2D maps.

We examine these results and caveats more closely in the following sections.

### 3.2 Surface overdensity $\Sigma$

The upper panel of Fig. 1 shows the map of the surface overdensity  $\Sigma$ , together with sky-plane centres of the intrinsic 3D voids of the galaxy distribution and the 2D voids detected via  $\Sigma$  as described in Section 2.5.1. The correspondence between these, formalized in Table 3, can be inspected qualitatively by judging if a 3D void centre (white +) has a 2D void centre (red x) more close to it than a randomly placed point. Of the  $N_{3D} = 46$  intrinsic galaxy voids, only  $N_{2D}^\Sigma = 28$  2D voids are detected (Table 2). The fact that  $N_{2D}^\Sigma < N_{3D}$  is expected, since we did not design our algorithm to distinguish voids that are nearly concentric when projected to the sky plane.

The lower panel of Fig. 1 shows the  $\overline{\Sigma}$  profiles averaged over all the 2D centres, and, independently, averaged over all the 3D centres (projected by ignoring the  $y$  coordinate). Comparison of these two curves (and their standard deviations, hatched) shows how well we might expect our algorithm to perform. The profiles for the 3D centres are those that would be detected if the algorithm were perfect in recovering the intrinsic voids, leaving aside the difference that the 2D detection uses the full dark matter particle distribution, while the 3D detection is galaxy based. It is clear that projection effects are significant: the mean profile (blue curve) does not show the sharp wall typical of voids. It is also clear that we have found profiles in  $\Sigma$  (red curve) that are stronger in contrast from minimum density to highest wall density than those of the intrinsic voids. This suggests that improving the algorithm further based on the motivation of optimizing a typical void-like profile, under the assumption of spherical shapes, would be unlikely to help further: strongly void-like profiles are already well detected.

The top two panels of Fig. 2 show the  $x$  and  $z$  coordinates (spanning the sky plane) of corresponding void centres, where the 2D void centres are those found to best match a given 3D void, as described in Section 2.4. The existence of multiple 3D voids whose best match is a single 2D void is clear in the diagram. We interpret this as illustrating cases where 3D voids are nearly aligned in projection, and thus detected as a single 2D void.

The bottom panel of Fig. 2 shows that void radii are very poorly recovered, and generally underestimated. One factor is clearly the difficulty in distinguishing nearly concentric voids. However, it is also likely that substructure is misidentified as void walls, leading to the underestimates. Overall, the bottom panel of Fig. 2 shows that the radii of our intrinsic population of 3D voids detected with REVOLVER are reduced by about  $5 \text{ Mpc } h^{-1}$ , in an uncorrelated way with a big scatter, when redetected with  $\Sigma$  as 2D voids.

### 3.3 Weak-lensing shear $\overline{\gamma}_\perp$

Using  $\overline{\gamma}_\perp$ , we find  $N_{2D}^\gamma = 29$  2D voids, i.e. roughly two thirds of the number of intrinsic galaxy voids,  $N_{3D} = 46$  (Table 2). Fig. 3 shows that the mean behaviour of a lensing profile in  $\overline{\gamma}_\perp$  using the centres of the intrinsic 3D voids is that  $\overline{\gamma}_\perp$  starts near zero, decreases to negative values in the void, and appears to (in the mean) reach a minimum at the wall radius found by REVOLVER, before increasing to a maximum at a somewhat greater radius. This is reasonable, given the definition for  $\overline{\gamma}_\perp$ . Fig. 3 shows that the 2D voids also have a (mean)  $\overline{\gamma}_\perp$  profile that decreases and then increases to zero, but the increase to zero occurs at lower fractions of the void radius.

Together, these profiles could be interpreted to suggest that applying a systematic correction factor to increase the void radius found when  $\overline{\Sigma}(r, \hat{n}) = \overline{\Sigma}(r, \hat{n})$  (see equation 14) might yield radii that better match those of the 3D voids. The lowest panel of Fig. 4 is qualitatively consistent with this suggestion, as it shows that the 2D voids that are best matched to the 3D voids have radii that are all smaller than the 3D void radii. However, Table 3 shows that correspondence between the radii is insignificant. As in the case of  $\Sigma$ , the projection of nearly concentric intrinsic voids, as well as obscuring effects from more distant overlapping voids, make the use of a single scaling correction for radii poorly motivated, except as a crude statistical correction.

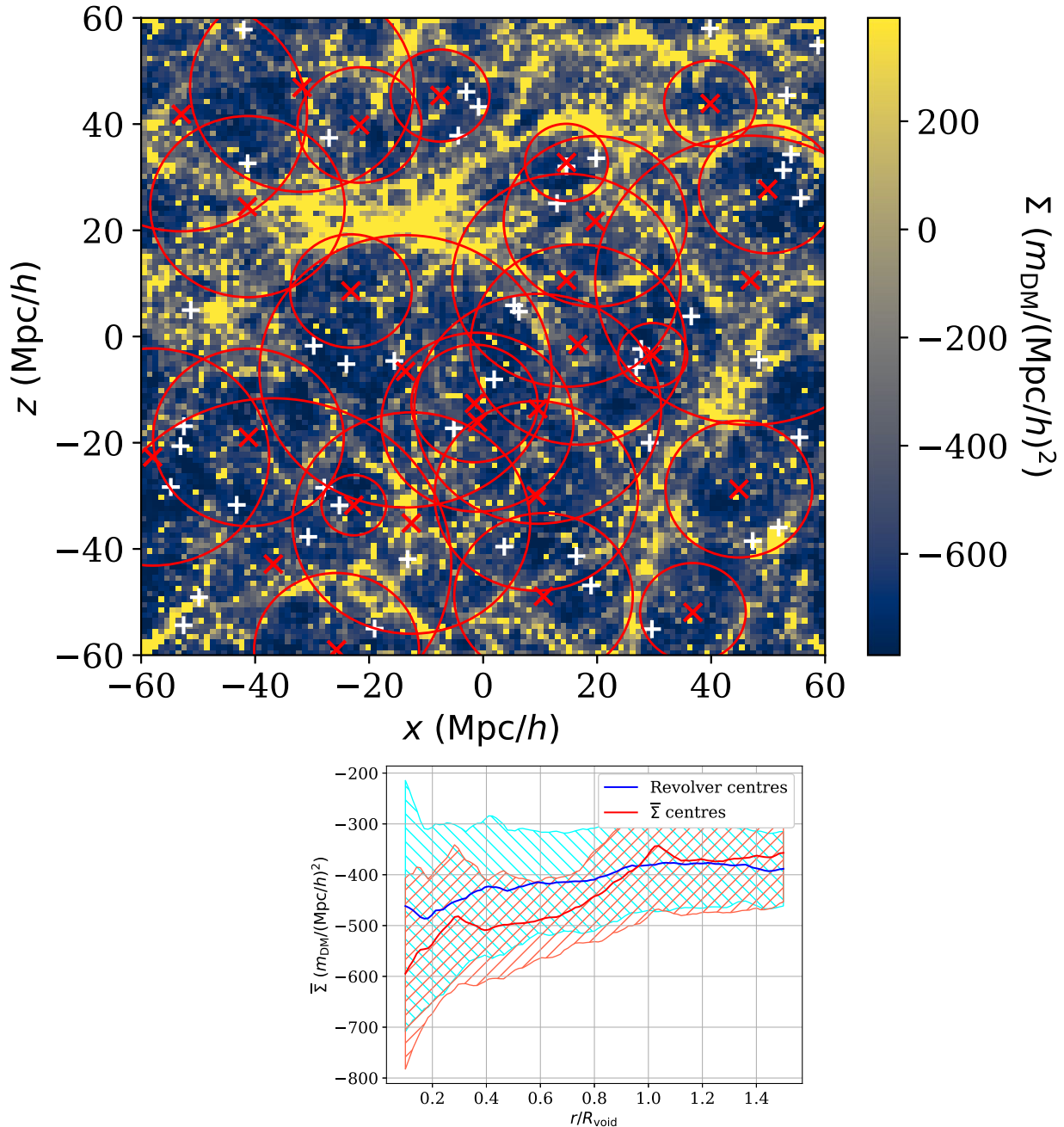
The two upper panels of Fig. 4 show what is quantified in Table 3: the sky-plane positions are recovered non-randomly to high statistical significance. Moreover, for the reverse question (Table 3),  $P_{x,z}^\gamma(3D|2D) \lesssim 0.001$  appears to be robust against re-calculation and machine error, so the use of weak-lensing shear – on its own – to infer the presence of intrinsic 3D galaxy-traced voids appears to be promising.

### 3.4 Optical scalars $\theta$ and $|\sigma|$

#### 3.4.1 Expansion $\theta$

Fig. 5 shows a map of the Sachs expansion  $\theta$  and sky-plane centres of both the intrinsic voids and those detected via  $\theta$ . As indicated in Table 3, given the 3D voids, the best-matched 2D ( $\theta$ ) voids are recovered to high significance via their sky-plane centres. The top two panels of Fig. 6 show the sky-plane matches.

However, we only find  $N_{2D}^\theta = 34$  voids using  $\theta$ , many less than the intrinsic voids. As with  $\Sigma$  and  $\overline{\gamma}_\perp$ , a likely interpretation is projected concentricity of several voids and obscuration by other cosmic web structure. The lower panel of Fig. 5 can be interpreted consistently with this hypothesis: the mean  $\theta$  profile of the full set of intrinsic voids



**Figure 1.** Upper panel: surface overdensity  $\Sigma$  projected along the line of sight. White + symbols represent the  $x, z$  centres of 3D voids found with REVOLVER. The red  $\times$  symbols represent the centres of 2D voids found in the surface overdensity (Section 2.5.1); the red circles represent the walls of these (circular, by definition) voids. Some of the 3D void centres are projected close to one another in the sky plane; our algorithm is not designed to distinguish these as independent voids. Lower panel: radial void profiles of the surface overdensity  $\Sigma$  (equation 2), normalized to the estimated void radius and then averaged, using the set of all (projected) 3D void centres and radii (mean: blue curve; standard deviation: green ‘\’ hatching; ‘Revolver centre’) or using the set of all 2D void centres and radii (mean: red curve; standard deviation: red ‘/’ hatching).

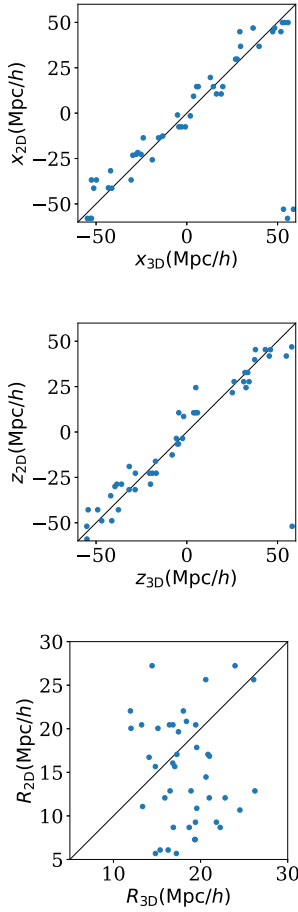
detected with REVOLVER is very weak, which would be consistent with both effects. The profile for 2D voids detected with  $\theta$  is very strong, qualitatively resembling a typical void density profile, with a sharp (mean) wall.

The lowest panel of Fig. 6 shows that the radii are again poorly correlated. Again, this is consistent with the detections using  $\Sigma$  and  $\bar{\gamma}_\perp$ , with the difference that the radii estimated with  $\theta$  expand greatly from the intrinsic voids’ range of around  $15\text{--}25 \text{ Mpc } h^{-1}$  to around  $5\text{--}30 \text{ Mpc } h^{-1}$ . While to some degree these disagreements are likely

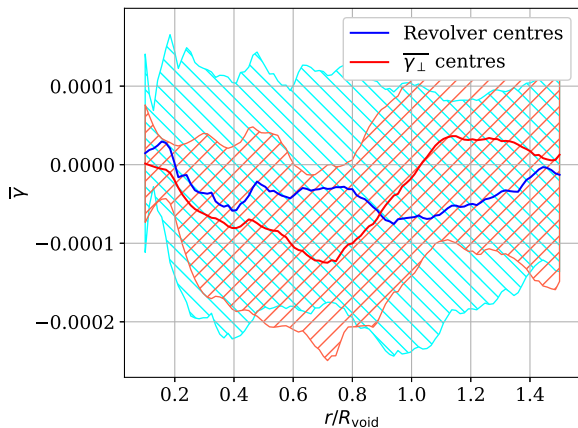
to be induced by the problems of projection, it might also be possible that radii that are gravitationally realistic in terms of the potential  $\Phi$  differ significantly from those traced by the 3D galaxy distribution. This is a question open for further study.

### 3.4.2 Sachs shear $|\sigma|$

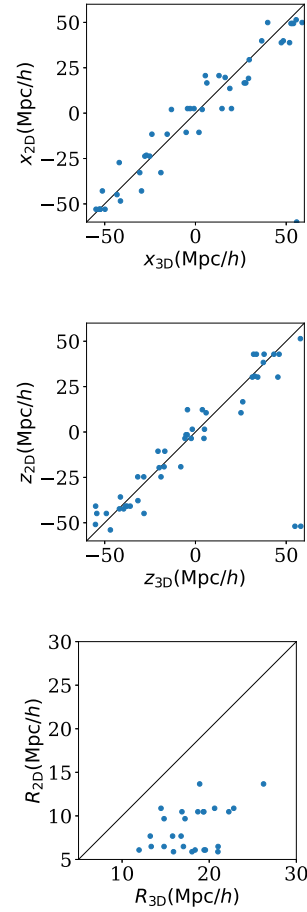
The upper panel of Fig. 7 shows a map of the modulus of the Sachs shear,  $|\sigma|$ . Again, Table 3 shows that given an intrinsic void, the Sachs



**Figure 2.** Top to bottom, respectively: Given a set of intrinsic 3D voids in the galaxy distribution, sky-plane positions  $x$  and  $z$  and radii  $R$  of the voids, and corresponding positions and radii of the 2D voids detected with the surface overdensity  $\Sigma$  that best match these 3D voids. The median  $(x, z)$   $T^2$  distance for the best-matched voids, given a 3D void (Section 2.4), for detections with  $\Sigma$  is  $7.8 \text{ Mpc } h^{-1}$ . Equality is shown by a straight line in all three panels. The radii match poorly, with 2D radii mostly being less than the intrinsic 3D radii. Plain-text data available at [zenodo.8103985/void\\_matches\\_mass\\_def\\_ivn\\_3D.dat](https://zenodo.org/record/8103985/files/void_matches_mass_def_ivn_3D.dat).



**Figure 3.** Radial void profiles of the weak-lensing shear  $\bar{\gamma}_{\perp}$ , as in the lower panel of Fig. 1, for 3D (projected) and 2D ( $\bar{\gamma}_{\perp}$ ) sets of void centres. A map for  $\bar{\gamma}_{\perp}$  is not shown, since the map of weak-lensing mean tangential shear  $\bar{\gamma}_{\perp}(r, \hat{n})$  is redetermined for each possible void centre  $\hat{n}$ .



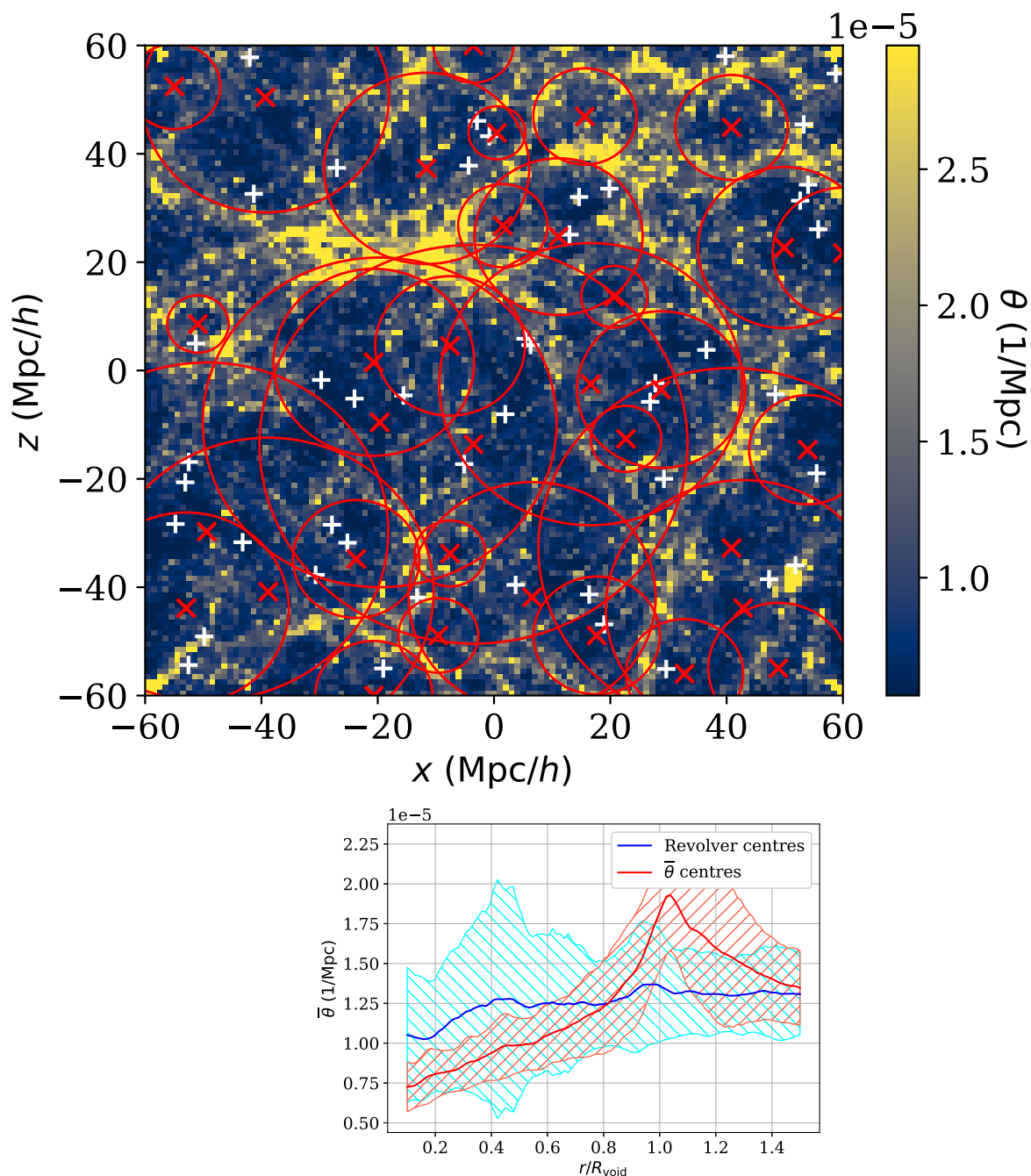
**Figure 4.** Top to bottom, respectively: For each intrinsic 3D void, sky-plane positions  $x$  and  $z$  and radii  $R$  of the best-matched 2D void detected with the weak-lensing shear  $\bar{\gamma}_{\perp}$ , as in Fig. 2. The median  $(x, z)$   $T^2$  distance for the best-matched voids, given a 3D void (Section 2.4), for detections with  $\bar{\gamma}_{\perp}$  is  $6.5 \text{ Mpc } h^{-1}$ . All 2D radii are lower than those of the 3D voids that they correspond to. Plain-text data available at [zenodo.8103985/void\\_matches\\_gamma\\_perp\\_3D.dat](https://zenodo.org/record/8103985/files/void_matches_gamma_perp_3D.dat).

shear detects the voids' positions to high significance using our algorithm. The lower panel of Fig. 7 shows a qualitatively similar result to the use of the expansion  $\theta$ , in the sense that shear profiles for the full set of 3D voids have a weak mean profile, while those for the voids detected in the 2D map of  $|\sigma|$  show a strong void-like profile. Taking into account the good sky-plane matches and poor radial matches shown in Fig. 8, a consistent interpretation is again that the 2D detected profiles are those that bypass both general obscuration and the confusion induced by voids that are nearly concentric in projection.

## 4 DISCUSSION

### 4.1 Void lensing studies when intrinsic voids are known

With the simulation presented here, we have shown that if intrinsic 3D voids are known, then the effects of geometric-optics parameters should be detectable in the sky plane, enabling the study of the role that gravitational lensing plays in the voids. In other words, we have shown a relation between voids in 3D comoving space with their imprints left on maps of the projected and ray-traced variables. Moreover, the lensing patterns induced by a void should provide feedback to better constrain the model of the void itself. As



**Figure 5.** Upper panel: Sachs expansion  $\theta$ , as for Fig. 1, computed using equations (20) and (21), with white + symbols for the  $x, z$  centres of 3D intrinsic galaxy voids and red  $\times$  symbols for the centres of 2D voids detected with  $\theta$ . Lower panel: Radial void profiles of  $\theta$ , as in the lower panel of Fig. 1, for 3D (REVOLVER) and 2D ( $\theta$ ) sets of void centres. A factor of  $10^{-5}$  in the vertical scale is indicated by ‘ $1e-5$ ’ (and similarly in Fig. 7 below).

argued by Sánchez et al. (2017) using  $\Sigma$  (and photometric redshifts to statistically limit the radial depth of the projection), this would confirm that a galaxy-traced void is a genuine underdensity of the dark matter density field. Other weak gravitational lensing detectors, such as the Lyman  $\alpha$  forest (Croft et al. 2018), could also be compared to foreground galaxy-traced voids to check for consistency.

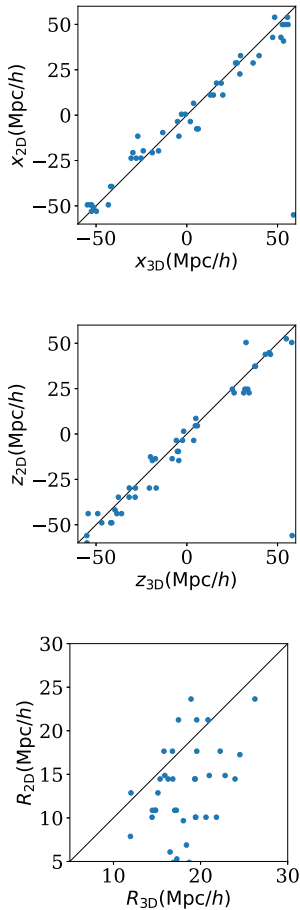
#### 4.2 Blind (redshift-free) searches for voids

Without knowledge of spectroscopic or photometric galaxy redshifts, we currently can justify use of the azimuthally averaged tangential

weak-lensing shear  $\overline{\gamma}_{\perp}$  and of the Sachs expansion  $\theta$  or shear  $|\sigma|$  for analysis of a photometric survey with the intention of inferring the presence of 3D voids, since all three of these robustly yield  $P_{x,z}^X(3D|2D) \lesssim 0.01$  (Table 3). Our calculation would appear to be the first time that the use of Sachs expansion maps has been shown to have the ability to reveal underlying voids. Jeffrey et al. (2021) studied the combined use of the usual weak-lensing convergence and shear in DES photometry over  $4143 \text{ deg}^2$ , which appears to be equivalent to using  $\Sigma$  and  $\overline{\gamma}_{\perp}$ , to reveal underlying voids.

Our algorithms can very likely be improved further. For example, combining all four parameters,  $\Sigma$ ,  $\overline{\gamma}_{\perp}$ ,  $\theta$ , and  $|\sigma|$ , could lead to





**Figure 6.** Top to bottom, respectively: For each intrinsic 3D void, sky-plane positions  $x$  and  $z$  and radii  $R$  of the best-matched 2D void detected with the Sachs expansion  $\theta$ , as in Fig. 2. The median  $(x, z)$   $T^2$  distance for the best-matched voids, given a 3D void (Section 2.4), for detections with  $\theta$  is  $5.8 \text{ Mpc } h^{-1}$ . The 2D radii have a much broader distribution than those of the intrinsic 3D voids, with no obvious correlation. Plain-text data available at [zenodo.8103985/void\\_matches\\_exp\\_given\\_3D.dat](https://zenodo.org/record/8103985/files/void_matches_exp_given_3D.dat).

complementary constraints on whether or not a putative void is real or correctly identified. These would only be partially independent from one another, since the four parameters are related to one another, with  $\theta$  and  $|\sigma|$  taking into account the evolution of the gravitational potential. Deriving the weak-lensing parameters for an initial approximation, and then using the Sachs optical scalars for an analysis to higher accuracy could be one viable strategy. Another extension would be to examine individual pairs of the best-matching 3D and 2D voids from the current algorithm presented here to understand how their match could be improved; or alternatively, examine the worst-matching pairs to understand what obstructs the match and search for ways of avoiding the obstruction.

There are several advantages in detecting voids via their sky-plane effects. Some of the fainter galaxies defining the walls of a void may be too faint to be detected in a given survey. The tracing of dark matter by luminous matter is by a long chain of physical effects: baryonic matter has to be associated with the dark matter, and star formation history and feedback effects need to be modelled. Geometric optics bypasses this causal chain, and should lead to inferences made with fewer assumptions.

### 4.3 Projected void concentricity and obscuring cosmic web structures

Projection of voids to be nearly concentric is expected in our simulation, since we integrate over the full box size of  $L_{\text{box}} = 120 \text{ Mpc } h^{-1}$  and the largest intrinsic voids detected with the watershed algorithm have radii  $R_{\text{eff}} \sim 30 \text{ Mpc } h^{-1}$ . Our algorithm's only strategy that relates to the problem of projected void concentricity is to prefer larger to smaller radii (step (vii) in Section 2.3.2). Figs 2, 4, 6, and 8 show that despite this, the 2D void radii tend to be lower than the intrinsic radii. This empirical result would tend to favour keeping this step unchanged.

In our analysis we make the simplification that the light bundle only experiences gravitational lensing while passing through the simulated volume and that there is no lensing signal generated behind or in front of the simulated volume. Obscuration over the full light cone would clearly be worse than in our model.

Our algorithm already has many parameters. Extending it to allow successive multiple detections of walls could, in principle, lead to a higher rate of detecting the intrinsic voids. Ideally, this should lead to a statistically significant correlation between the intrinsic and photometric void radii; in this work, our correlations in radii are insignificant (Table 3). However, detecting multiple concentric walls would quite likely also lead to false detections.

Strategies for solving the problem of obscuring structures (in the absence of redshift information) are not obvious. Gravitationally dense objects occupy little volume and still suffer from projection effects; voids dominate the volume and thus are strongly affected by projection effects. A Bayesian approach as in Jeffrey et al. (2021) would be worth exploring.

One class of changes to our algorithm worth considering would be to switch equation (3) to include a weighting with the radius  $r'$ . This would make the noise characteristics uniform over the sky area, so that the algorithm is less sensitive to noise in small numbers of pixels. However, for void-like radial profiles, this would require modifying the profile detection aspects of our algorithm to be less sensitive to requiring the void-like shape to be accurate in the void centre.

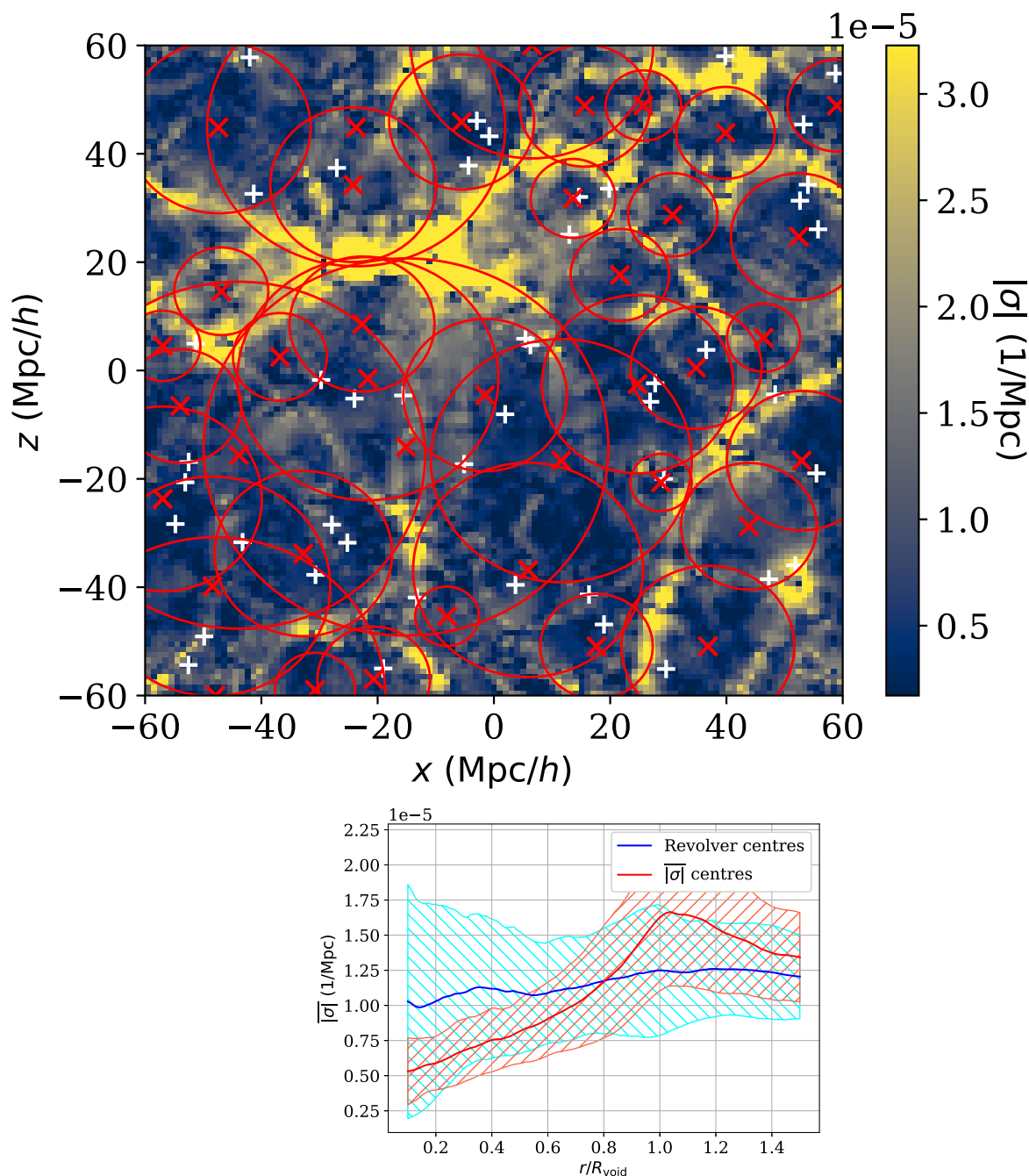
Since our simulation homogenizes the foreground and background of the simulated volume, a real observational survey will include stronger levels of both projected void concentricity and obscuring cosmic web structures.

### 4.4 Other extensions

This work presents an initial proof-of-concept for detecting cosmic voids via geometric optics. The specific results will vary with simulation size in the senses both of length scales and particle resolution, and between the use of different tracer particles. Further work could consider the dependence on the probability of voids being real for a given choice of tracer particle (Neyrinck 2008), and on the choice of tracers themselves (Nadathur & Hotchkiss 2015a).

## 5 CONCLUSION

In this work, we have studied the two questions of whether voids in the cosmic web yield detectable information in projected variables, the surface overdensity  $\Sigma$ , the azimuthal averaged weak lensing shear  $\overline{\gamma}_{\perp}$ , the Sachs expansion  $\theta$ , and the Sachs shear  $|\sigma|$ , and vice versa, whether the sky-plane information can be used to infer the existence of the intrinsic 3D voids. We performed this using a cosmological  $N$ -body simulation starting from initial perturbations generated according to a standard initial power spectrum. We carried

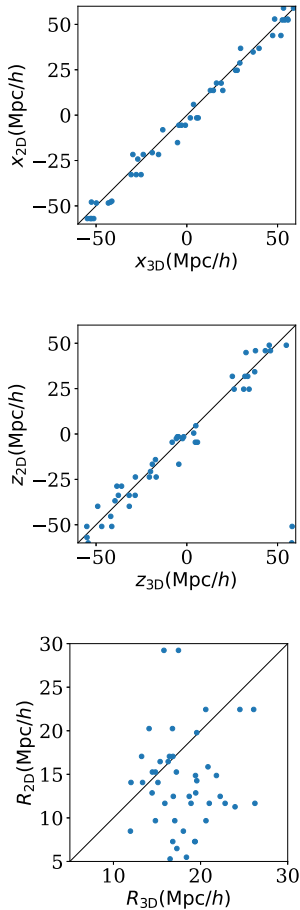


**Figure 7.** Upper panel: Sachs shear  $|\sigma|$ , as for Fig. 1, computed using equations (20) and (21), with white + symbols for the  $x, z$  centres of 3D intrinsic galaxy voids and red  $\times$  symbols for the centres of 2D voids detected with  $|\sigma|$ . Lower panel: Radial void profiles of  $|\sigma|$ , as in the lower panel of Fig. 1, for 3D (REVOLVER) and 2D ( $|\sigma|$ ) sets of void centres.

out the analysis in a fully controlled software environment with full information about the dark matter distribution as well as the luminous matter distribution, which we modelled using galaxies built from a halo merger tree using semi-analytical tools. We detected the intrinsic voids in the 3D comoving spatial distribution of galaxies using a watershed void finder (Section 2.3.1). The void detection in the projected plane (Section 2.3.2) is based on the assumption that the azimuthally averaged profiles of the four detector variables for the voids have shapes with predictable qualitative behaviour. In the case of the surface overdensity  $\Sigma$  and the two Sachs optical scalars  $\theta$  and  $|\sigma|$ , this expected shape is to start from a minimum

at the centre of a void, gradually increase radially outwards, and increase sharply at the void’s wall. The weak-lensing shear  $\bar{\gamma}_\perp$  is expected to start from zero, decrease, and increase to zero just past the void’s wall. Using a heuristically parametrized algorithm for detecting these profiles, adjusted individually for the four detector variables, we found positions and radii of 2D voids traced by these detectors.

We find roughly similar numbers of 2D voids traced by each of the four different detector variables, and in all cases, fewer voids than in the 3D galaxy-traced distribution, as can be seen in Table 2. There are two likely explanations. First, when several intrinsic voids are nearly



**Figure 8.** Top to bottom, respectively: For each intrinsic 3D void, sky-plane positions  $x$  and  $z$  and radii  $R$  of the best-matched 2D void detected with the Sachs shear modulus  $|\sigma|$ , as in Fig. 2. The median  $(x, z)$   $T^2$  distance for the best-matched voids, given a 3D void (Section 2.4), for detections with  $|\sigma|$  is  $4.4 \text{ Mpc } h^{-1}$ . The 2D radii again have a broad distribution, as in Fig. 6. Plain-text data: [zenodo.8103985/void\\_matches\\_sig\\_given\\_3D.dat](https://zenodo.org/record/8103985/files/void_matches_sig_given_3D.dat).

concentric in projection on the sky, our algorithm only detects one of these, since it is not designed to detect multiple walls. Secondly, the foreground and background structures of the cosmic web, i.e. walls, filaments, clusters, and other voids, obscure the signals associated with any single intrinsic void, making detection difficult. The lower panels of Figs 1, 3, 5, and 7 show that the voids detected by us in the projected plane follow the assumed qualitative shapes well, giving confidence that our algorithm works as expected. However, the same panels show that the corresponding mean profiles, using the centres and radial sizes of the 3D intrinsic voids, but the detector variables in the projected plane, are weak.

We interpret these two effects – the detection of fewer 2D voids than those known to exist in the 3D spatial distribution, together with the weak mean profiles of the projected-plane detector variables centred at the intrinsic voids’ locations – as consistent with the undetected voids being (statistically) those that are either the most obscured or are concentric with the detected voids.

Given knowledge of the 3D voids’ centres, we find (Table 3, third column) that the detected 2D voids are significantly closer than random to the 3D voids’ centres in the sky plane, for all four detector variables. In other words, a survey with sufficient spectroscopic or photometric redshift information to detect voids should be usable to infer patterns of gravitational lensing through the voids that should

be measurable using either weak-lensing shear or the Sachs optical scalars (answering question (ii) of Section 2.4 positively).

Conversely, if we only have a photometric survey that is blind, in the sense of having neither spectroscopic nor photometric redshift information, then we have established (Table 3, second column) that the 2D voids detected via weak-lensing tangential shear  $\overline{\gamma}_\perp$ , Sachs expansion  $\theta$  or Sachs absolute shear  $|\sigma|$  significantly reveal the true underlying 3D void population (question (i) in Section 2.4). Use of the surface overdensity  $\Sigma$  provides weaker evidence for revealing the sky-plane positions of the underlying void population.

While these results follow from significant correlations of voids’ locations in the sky plane, we find no significant correlation for the radii. The bottom panels of Figs 2, 4, 6, and 8, show that the 2D void radii tend to be lower than the intrinsic radii. The lack of correlation and the generally lower radii are consistent with the problem of near concentric projection of multiple voids into the sky plane.

While our current results are exploratory, with several caveats as stated above, it does appear that gravitational lensing through individual voids should be observationally detectable. Moreover, weak-lensing tangential shear and Sachs expansion and shear in future blind photometric surveys – such as those provided by the Rubin C. Observatory’s Legacy Survey of Space and Time (LSST; Sheldon et al. 2023) – should reveal the existence of intrinsic 3D voids, yielding predictions that will be falsifiable by spectroscopic followup surveys such as those of the 4-m Multi-Object Spectroscopy Telescope (4MOST; de Jong et al. 2012, 2019; Richard et al. 2019) or the Dark Energy Spectroscopic Instrument (DESI; Levi et al. 2013; Hahn et al. 2022).

## ACKNOWLEDGEMENTS

The authors wish to thank Mariana Jaber, Matteo Cinus, David Bacon, and an anonymous referee for many useful comments and suggestions. This work has been partly supported by the Polish Ministerstwo Nauki i Szkolnictwa Wyższego (MNiSW; now Ministerstwo Nauki i Edukacji) grant DIR/WK/2018/12. This work has been partly supported by the Poznań Supercomputing and Networking Center (PSNC) computational grant 537. Part of this work was supported by Universitas Copernicana Thoruniensis in Futuro under Narodowe Centrum Badań i Rozwoju grant POWR.03.05.00-00-Z302/17. This work has been partly supported by the Narodowe Centrum Nauki (NCN) grant 2022/45/N/ST9/01698.

We gratefully acknowledge the use of the following free-software programs and libraries: 1.23, Boost 1.77.0, Bzip2 1.0.8, converttrees 0.0-522dac5, cosmdist 0.3.12, ctrees 1.01-7c20add, cURL 7.84.0, Dash 0.5.11-057cd65, diff-solver 0.0.1, Discoteq flock 0.4.0, Eigen 3.4.0, Expat 2.4.1, fftw2 2.1.5-4.2, FFTW 3.3.10 (Frigo & Johnson 2005), File 5.42, Fontconfig 2.14.0, FreeType 2.11.0, Git 2.37.5, GNU Autoconf 2.71, GNU Automake 1.16.5, GNU AWK 5.1.1, GNU Bash 5.2-rc2, GNU Binutils 2.39, GNU Bison 3.8.2, GNU Compiler Collection (GCC) 12.1.0, GNU Coreutils 9.1, GNU Difutils 3.8, GNU Findutils 4.9.0, GNU gettext 0.21, GNU gperf 3.1, GNU Grep 3.7, GNU Gzip 1.12, GNU Integer Set Library 0.24, GNU libiconv 1.17, GNU Libtool 2.4.7, GNU libunistring 1.0, GNU M4 1.4.19, GNU Make 4.3, GNU Multiple Precision Arithmetic Library 6.2.1, GNU Multiple Precision Complex library, GNU Multiple Precision Floating-Point Reliably 4.1.0, GNU Nano 6.4, GNU NCURSES 6.3, GNU Patch 2.7.6, GNU Readline 8.2-rc2, GNU Scientific Library 2.7, GNU Sed 4.8, GNU Tar 1.34, GNU Texinfo 6.8, GNU Wget 1.21.2, GNU Which 2.21, GPL

Ghostscript 9.56.1, Help2man, ImageMagick 7.1.0-13, Less 590, Libffi 3.4.2, libICE 1.0.10, Libidn 1.38, Libjpeg 9e, Libpaper 1.1.28, Libpng 1.6.37, libpthread-stubs (Xorg) 0.4, libSM 1.2.3, Libtiff 4.4.0, libXau (Xorg) 1.0.9, libxcb (Xorg) 1.15, libXdmcp (Xorg) 1.1.3, libXext 1.3.4, Libxml2 2.9.12, libXt 1.2.1, LibYAML 0.2.5, Lzip 1.23, mpgratic 0.3.19-4b78328, OpenBLAS 0.3.21, Open MPI 4.1.1, OpenSSL 3.0.5, PatchELF 0.13, Perl 5.36.0, pkg-config 0.29.2, podlators 4.14, Python 3.10.6, ramses-scalav 0.0-482f90f, revolver 0.0-3b15335, rockstar 0.99.9-RC3 + -6d16969, sage 0.0-2be3027, util-Linux 2.38.1, util-macros (Xorg) 1.19.3, X11 library 1.8, XCB-proto (Xorg) 1.15, xorgproto 2022.1, xtrans (Xorg) 1.4.0, XZ Utils 5.4.1 and Zlib 1.2.11. Part of this project used the following PYTHON modules: Astropy 5.1 (Astropy Collaboration 2013, 2018), BeautifulSoup 4.10.0, Cyler 0.11.0, Cython 0.29.24 (Behnel et al. 2011), emcee 3.0.1, Extension-Helpers 0.1, HTML5lib 1.0.1, Jinja2 3.0.3, Kiwisolver 1.0.1, MarkupSafe 2.0.1, Matplotlib 3.3.0 (Hunter 2007), Numpy 1.21.3 (Harris et al. 2020), Packaging 21.3, Pillow 8.4.0, pybind11 2.5.0, PyERFA 2.0.0.1, pyFFTW 0.12.0, PyParsing 3.0.4, python-dateutil 2.8.0, python-installer 0.5.1, PyYAML 6.0, Scipy 1.10.0 (Virtanen et al. 2020), Setuptools 58.3.0, Setuptools-scm 3.3.3, Six 1.16.0, SoupSieve 1.8, Webencodings 0.5.1 and Wheel 0.37.0. The typesetting of this paper was carried out using the following LaTeX and related free-software packages: aegreya 64384 (revision), biber 2.19, biblatex 3.19, bitset 1.3, caption 66580 (revision), courier 61719 (revision), csquotes 5.2n, datetime 2.60, ec 1.0, enumitem 3.9, environ 0.3, etoolbox 2.5k, fancyhdr 4.1, fancyvrb 4.5a, fntcount 3.07, fontaxes 1.0e, fontspec 2.8a, footmisc 6.0d, fp 2.1d, kastrup 15878 (revision), lastpage 2.0a, latexpand 1.7.2, letltxmacro 1.6, listings 1.9, logreq 1.0, mnras 3.1, mweights 53520 (revision), newtx 1.71, pdfcol 1.7, pdfescape 1.15, pdftexcmds 0.33, pgf 3.1.10, pgfplots 1.18.1, preprint 2011, setspace 6.7b, tcolorbox 6.0.3, tex 3.141592653, texgyre 2.501, times 61719 (revision), titlesec 2.14, trimspaces 1.1, txfonts 15878 (revision), ulem 53365 (revision), xcolor 2.14, xkeyval 2.9 and xstring 1.85.

## DATA AVAILABILITY

The full reproducibility package for this paper, including input data and source software CHECKSUMS, is available at zenodo.8103985 and in live<sup>5</sup> and archived<sup>6</sup> GIT repositories. Our main numerical results are available as DOI-identified records as indicated in the captions of Table 3, and Figs 2, 4, 6, and 8. The version of the source package used to produce this paper has the GIT commit hash d8952e7.

## REFERENCES

- Adam G., Bacon R., Courtes G., Georgelin Y., Monnet G., Pecontal E., 1989, *A&A*, 208, L15
- Akhlaghi M., Infante-Sainz R., Roukema B. F., Khellat M., Valls-Gabaud D., Baena-Gallé R., 2021, *Comp. Sci. Eng.*, 23, 82
- Aragón-Calvo M. A., Platen E., van de Weygaert R., Szalay A. S., 2010, *ApJ*, 723, 364
- Astropy Collaboration 2013, *A&A*, 558, A33
- Astropy Collaboration 2018, *AJ*, 156, 123
- Bartelmann M., Schneider P., 2001, *Phys. Rep.*, 340, 291

<sup>5</sup><https://codeberg.org/mpeper/lensing>

<sup>6</sup>swb:1:rev:169cb31e75ba6bcb69dd31741cf53f9c6795f91

- Behnel S., Bradshaw R., Citro C., Dalcin L., Seljebotn D. S., Smith K., 2011, *Comput. Sci. Eng.*, 13, 31
- Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013a, *ApJ*, 762, 109
- Behroozi P. S., Wechsler R. H., Wu H.-Y., Busha M. T., Klypin A. A., Primack J. R., 2013b, *ApJ*, 763, 18
- Bos E. G. P., van de Weygaert R., Dolag K., Pettorino V., 2012, *MNRAS*, 426, 440
- Clarkson C., Ellis G. F. R., Faltenbacher A., Maartens R., Umeh O., Uzan J.-P., 2012, *MNRAS*, 426, 1121
- Croft R. A. C., Romeo A., Metcalf R. B., 2018, *MNRAS*, 477, 1814
- Croton D. J. et al., 2016, *ApJS*, 222, 22
- de Jong R. S. et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. SPIE, Bellingham, p. 84460T
- de Jong R. S. et al., 2019, *The Messenger*, 175, 3
- Frigo M., Johnson S. G., 2005, *IEEE Proc.*, 93, 216
- Gorenstein M. V. et al., 1984, *ApJ*, 287, 538
- Gorenstein M. V., Cohen N. L., Shapiro I. I., Rogers A. E. E., Bonometti R. J., Falco E. E., Bartel N., Marcaide J. M., 1988, *ApJ*, 334, 42
- Gottlöber S., Lokas E. L., Klypin A., Hoffman Y., 2003, *MNRAS*, 344, 715
- Gregory S. A., Thompson L. A., 1978, *ApJ*, 222, 784
- Gruen D. et al., 2016, *MNRAS*, 455, 3367
- Gunn J. E., Gott III J. R., 1972, *ApJ*, 176, 1
- Hahn C. et al., 2023, *AJ*, 165, 253
- Harris C. R. et al., 2020, *Nature*, 585, 357
- Heinesen A., 2023, *Phys. Rev. D*, 107, L101301
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Jeffrey N. et al., 2021, *MNRAS*, 505, 4626
- Jöeveer M., Einasto J., Tago E., 1978, *MNRAS*, 185, 357
- Kauffmann G., Fairall A. P., 1991, *MNRAS*, 248, 313
- Kilbinger M., 2015, *Rep. Prog. Phys.*, 78, 086901
- Krause E., Chang T.-C., Doré O., Umetsu K., 2013, *ApJ*, 762, L20
- Levi M. et al., 2013, preprint ([arXiv:1308.0847](https://arxiv.org/abs/1308.0847))
- Lewis A., Challinor A., 2006, *Phys. Rep.*, 429, 1
- Li B., Zhao G.-B., Koyama K., 2012, *MNRAS*, 421, 3481
- Lilje P. B., Lahav O., 1991, *ApJ*, 374, 29
- Nadathur S., Hotchkiss S., 2014, *MNRAS*, 440, 1248
- Nadathur S., Hotchkiss S., 2015a, *MNRAS*, 454, 889
- Nadathur S., Hotchkiss S., 2015b, *MNRAS*, 454, 2228
- Nadathur S., Carter P. M., Percival W. J., Winther H. A., Bautista J. E., 2019, *Phys. Rev. D*, 100, 023504
- Neyrinck M. C., 2008, *MNRAS*, 386, 2101
- Paczynski B., 1987, *Nature*, 325, 572
- Pan D. C., Vogeley M. S., Hoyle F., Choi Y.-Y., Park C., 2012, *MNRAS*, 421, 926
- Peebles P. J. E., 2001, *ApJ*, 557, 495
- Peper M., Roukema B. F., 2021, *MNRAS*, 505, 1223
- Pisani A., Sutter P. M., Hamaus N., Alizadeh E., Biswas R., Wandelt B. D., Hirata C. M., 2015, *Phys. Rev. D*, 92, 083531
- Prunet S., Pichon C., Aubert D., Pogosyan D., Teyssier R., Gottloeber S., 2008, *ApJS*, 178, 179
- Richard J. et al., 2019, *The Messenger*, 175, 50
- Roukema B. F., Quinn P. J., Peterson B. A., 1993, in Chincarini G. L., Iovino A., Maccacaro T., Maccagni D., eds, ASP Conf. Ser. Vol. 51, Observational Cosmology: an International Symposium, Milano, Italy, 21–25 September 1992. Astron. Soc. Pacific., San Francisco, p. 298
- Roukema B. F., Peterson B. A., Quinn P. J., Rocca-Volmerange B., 1997, *MNRAS*, 292, 835
- Sachs R., 1961, *Proc. R. Soc. Lond. Ser. A*, 264, 309
- Sánchez C. et al., 2017, *MNRAS*, 465, 746
- Sasaki M., 1993, *Prog. Theor. Phys.*, 90, 753
- Savitzky A., Golay M. J. E., 1964, *Anal. Chem.*, 36, 1627
- Schneider P., Ehlers J., Falco E. E., 1992, *Gravitational Lenses*. Springer-Verlag, Berlin
- Sheldon E. S., Becker M. R., Jarvis M., Armstrong R., The LSST Dark Energy Science Collaboration, *Open J. of Astrophysics*, 2023, 6, 17



- Sheth R. K., van de Weygaert R., 2004, *MNRAS*, 350, 517  
 Spearman C. E., 1904, *Am. J. Psychol.*, 15, 72  
 Sussman R. A., Bolejko K., 2012, *Class. Quantum Gravity*, 29, 065018  
 Teysier R., 2002, *A&A*, 385, 337  
 Virtanen P. et al., 2020, *Nat. Methods*, 17, 261  
 Walsh D., Carswell R. F., Weymann R. J., 1979, *Nature*, 279, 381  
 Young P., Gunn J. E., Kristian J., Oke J. B., Westphal J. A., 1980, *ApJ*, 241, 507

## APPENDIX A: COSMIC VARIANCE

While a full calculation of cosmic variance would be computationally heavy, we provide the results of using three alternative choices of random-number-generator seeds in Table A1. This gives a qualitative illustration of cosmic variance. The specific results in significance differ somewhat from those for the simulation of this work (Table 3), especially in the case of using 2D voids to discover the intrinsic 3D voids. Thus, the optimal choice of detector variable for an observational programme will to some degree vary among surveys, depending on the physical realization of primordial fluctuation amplitudes and phases. Using a wide range of detector variables is likely to be more successful in the analysis of real surveys rather than restricting the analysis to only a few variables.

**Table A1.** As for Table 3, probability that the matches between 3D and 2D voids for detector variable  $X$  are no better than those of randomly generated 2D voids, for three examples of alternative random seeds.

$X$	$P_{x,z}(3D 2D)$	$P_{x,z}(2D 3D)$	$P_R(3D 2D)$	$P_R(2D 3D)$
Alternative seed 1				
$\Sigma$	0.013	$2.0 \times 10^{-5}$	0.58	0.33
$\overline{\gamma}_\perp$	0.00035	0.00024	0.14	0.21
$\theta$	0.0018	$5.0 \times 10^{-5}$	0.98	0.83
$ \sigma $	0.027	0.00033	0.69	0.98
Alternative seed 2				
$\Sigma$	0.0059	$8.0 \times 10^{-5}$	0.19	0.42
$\overline{\gamma}_\perp$	0.15	0.026	0.62	0.35
$\theta$	0.017	0.0037	0.0079	0.056
$ \sigma $	0.0031	$4.0 \times 10^{-5}$	0.30	0.66
Alternative seed 3				
$\Sigma$	0.0027	0.0018	0.79	0.67
$\overline{\gamma}_\perp$	$2.0 \times 10^{-5}$	$1.0 \times 10^{-5}$	0.52	0.44
$\theta$	0.12	0.014	0.0056	0.0084
$ \sigma $	0.058	0.0014	0.83	0.38

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