



Universiteit  
Leiden  
The Netherlands

## Constraining cosmology with weak lensing voids

Davies, C.T.; Cautun, M.; Giblin, B.; Li, B.; Harnois-Déraps, J.; Cai, Y.-C.

### Citation

Davies, C. T., Cautun, M., Giblin, B., Li, B., Harnois-Déraps, J., & Cai, Y. -C. (2021). Constraining cosmology with weak lensing voids. *Monthly Notices Of The Royal Astronomical Society*, 507(2), 2267-2282. doi:10.1093/mnras/stab2251

Version: Accepted Manuscript

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/3256339>

**Note:** To cite this publication please use the final published version (if applicable).

# Constraining cosmology with weak lensing voids

Christopher T. Davies,<sup>1\*</sup> Marius Cautun,<sup>2</sup> Benjamin Giblin,<sup>3</sup> Baojiu Li,<sup>1</sup>  
Joachim Harnois-Déraps,<sup>3,4</sup> and Yan-Chuan Cai<sup>3</sup>

<sup>1</sup>*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*

<sup>2</sup>*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*

<sup>3</sup>*Scottish Universities Physics Alliance, Institute for Astronomy, University of Edinburgh, Blackford Hill, Scotland, UK*

<sup>4</sup>*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool, L3 5RF, UK*

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

Upcoming galaxy surveys such as LSST and EUCLID are expected to significantly improve the power of weak lensing as a cosmological probe. In order to maximise the information that can be extracted from these surveys, it is important to explore novel statistics that complement standard weak lensing statistics such as the shear-shear correlation function and peak counts. In this work, we use a recently proposed weak lensing observable — weak lensing voids — to make parameter constraint forecasts for an LSST-like survey. We make use of the *cosmo-SLICS* suite of  $w$ CDM simulations to measure void statistics (abundance and tangential shear) as a function of cosmological parameters. The simulation data is used to train a Gaussian process regression emulator that we use to generate likelihood contours and provide parameter constraints from mock observations. We find that the void abundance is more constraining than the tangential shear profiles, though the combination of the two gives additional constraining power. We forecast that without tomographic decomposition, these void statistics can constrain the matter fluctuation amplitude,  $S_8$  within 0.7% (68% confidence interval), while offering 4.3, 4.7 and 6.9% precision on the matter density parameter,  $\Omega_m$ , the reduced Hubble constant,  $h$ , and the dark energy equation of state parameter,  $w_0$ , respectively. We find that these results are tighter than the constraints given by the shear-shear correlation function with the same observational specifications, indicating that weak lensing void statistics can be a promising cosmological probe potentially complementary with other lensing tests.

**Key words:** gravitational lensing; weak – large-scale structure of universe – cosmology; theory – methods: data analysis

## 1 INTRODUCTION

The standard model of cosmology, in which the dominant matter component consists of cold dark matter (CDM) while the late-time accelerated expansion is driven by a positive cosmological constant,  $\Lambda$ , is highly successful at describing a number of independent observations, which constrain these parameters with a large degree of concordance. Notably, measurement of fluctuations in the Cosmic Microwave Background (CMB) (Planck Collaboration et al. 2018), provides measurements of the present day expansion rate of the universe  $H_0$ , the matter density parameter  $\Omega_m$  and the matter fluctuation amplitude  $\sigma_8$ .

Another promising observational probe that is sensitive to and can be used to constrain many cosmological parameters is gravitational lensing, a phenomenon according to which the light of distant source images is distorted by the gravitational potentials of

the foreground matter. In the strong lensing regime distant galaxies are visibly distorted into large arcs. In the weak lensing (WL) regime, which is the focus of this study, this effect is much smaller, and the weak lensing signal is measured through the correlations in distortions of many source galaxies (Bacon et al. 2000; Kaiser et al. 2000; Van Waerbeke et al. 2000; Wittman et al. 2000). This allows us to probe the total matter distribution of the Universe on the largest scales (see Bartelmann & Schneider 2001; Kilbinger 2015, for reviews), and offers a powerful method to study the clustering of dark matter and its evolution.

Some of the most recent WL observations that supplement the parameter measurements from the CMB include the DES (Troxel et al. 2018)<sup>1</sup>, HSC (Hikage et al. 2019)<sup>2</sup> and KiDS (Asgari et al. 2020)<sup>3</sup> WL surveys. However, all of these surveys measure lower

<sup>1</sup> <https://www.darkenergysurvey.org/>

<sup>2</sup> <https://hsc.mtk.nao.ac.jp/ssp/>

<sup>3</sup> <http://kids.strw.leidenuniv.nl/>

\* E-mail: christopher.t.davies@durham.ac.uk (CTD)

values of  $\sigma_8$  compared to Planck, with a statistically significant disagreement arising in the comparison between the Planck and KiDS constraints. This is one example of the parameter tensions that have arisen in recent years, where different observations point to slightly different values of certain cosmological parameters, implying the presence of either unaccounted for systematics or new physics which are unaccounted for. Another example is the  $H_0$  tension, where multiple observations find that measurements from the early Universe are broadly inconsistent with those of the late universe (Verde et al. 2019), particularly the distance scale measurement of  $H_0$  based on Cepheids by the SH0ES collaboration (Riess et al. 2019).

In order to address these parameter tensions, it is important to measure cosmological parameters as precisely as possible, by maximising the information that can be extracted from a given survey. The standard approach for weak lensing surveys is to measure  $\Lambda$ CDM parameters with two-point statistics such as the shear-shear correlation function or the convergence power spectrum (Schneider et al. 2002; Semboloni et al. 2006; Hoekstra et al. 2006; Fu et al. 2008; Heymans et al. 2012; Kilbinger et al. 2013; Hildebrandt et al. 2017; Troxel et al. 2018; Hikage et al. 2019; Aihara et al. 2019; Asgari et al. 2020). However, two-point statistics do not capture non-Gaussian information, and weak lensing data are highly non-Gaussian due to the non-linear evolution of the Universe. To address this loss, many complimentary statistics have been developed, which encapsulate information beyond two-point statistics. A common and popular example is the abundance of WL peaks, which has been shown to be complimentary to the two-point function and helps break the  $\Omega_m$ - $\sigma_8$  parameter degeneracy (Jain & Van Waerbeke 2000; Pen et al. 2003; Dietrich & Hartlap 2010). Peaks are also shown to outperform the standard methods for constraining the sum of neutrino mass (Li et al. 2019) and  $w_0$  (Martinet et al. 2020). By including complimentary statistics, the measurement errors on cosmological parameters can be reduced, which will help inform the statistical significance of any parameter tensions between multiple observations.

The goal of this paper is to present parameter constraint forecasts for one such complimentary probe, WL voids. Voids are typically identified within the full 3D distribution of matter as regions of low matter density or low tracer density, for which void statistics such as their abundance, radial profiles and shapes contain useful non-Gaussian information (see, e.g., White 1979; Fry 1986; Biswas et al. 2010; Bos et al. 2012; Lavaux & Wandelt 2012). Most studies use galaxy voids, which are identified as underdense regions in the galaxy distribution (e.g., Paz et al. 2013; Sutter et al. 2014; Cautun et al. 2016; Nadathur 2016), where galaxy void statistics are complementary to the galaxy power spectrum and baryonic acoustic oscillations (e.g., Pisani et al. 2015; Hamaus et al. 2016; Nadathur et al. 2019). Recently, void WL profiles have also been shown to be a powerful cosmological probe (see, e.g., Cai et al. 2015; Barreira et al. 2015; Gruen et al. 2016; Barreira et al. 2017; Falck et al. 2018; Baker et al. 2018).

While less explored compared with 3D voids, voids can also be identified in projection, such as in the projected galaxy distribution (e.g. Gruen et al. 2015; Barreira et al. 2017; Sánchez et al. 2017; Cautun et al. 2018) or in a weak lensing map (e.g. Davies et al. 2018; Coulton et al. 2019). Here, we follow the latter approach and define WL voids generally as 2D regions within WL convergence maps that contain low convergence or few to no tracers. In a previous work (Davies et al. 2018), we have shown that the lensing profiles of 2D WL voids identified directly in WL convergence maps can be measured with a larger signal-to-noise ratio (SNR) than those of galaxy voids. This is because WL voids correspond to deeper

underdensities projected along the line of sight than galaxy voids, and hence they have larger tangential shear profiles. This also means that WL voids are better at constraining alternative cosmological models, such as modified gravity models, than galaxy voids (Davies et al. 2019b). Additionally, compared to other WL statistics, WL voids are less affected by baryonic physics (Coulton et al. 2019).

In this paper we use the *cosmo*-SLICS simulation suite (Harnois-Déraps et al. 2019) to identify a particular class of WL voids, the *tunnels*, for a range of cosmological parameters. We use this data to train a Gaussian process regression emulator, which, combined with Markov chain Monte Carlo, allows us to generate likelihood contours and provide forecast parameter constraints for an LSST-like survey.

The tunnel algorithm we use here is one possible choice of WL void finder. In fact, similar to voids identified in the galaxy distribution (e.g. Colberg et al. 2008; Cautun et al. 2018; Baker et al. 2019), there are several void finding methods that have been successfully applied to WL maps. For example, Davies et al. (2020) have carried out a detailed analysis on the impact that varying the WL void definition might have on the resulting WL void statistics. They have found that the ‘tunnel’ void finding algorithm offers a great trade off between maximising the observable tangential shear profile SNR and minimising the impact of observational noise on the void statistics. Therefore, we limit our analysis to only tunnels, and we defer a more detailed study comparing the parameter constraining powers of different void finders to a future work. For galaxy voids, studies have shown that combining different void definitions can lead to improved cosmological constraints (e.g. Paillas et al. 2019).

The layout of the paper is as follows. In Section 2 we outline the relevant theory for WL observations. In Section 3 we describe our mock observational data, emulation and likelihood analysis pipeline and void finding algorithm. In Section 4 we present the WL void statistics used in our analysis and in Section 5 we present our parameter constraint forecasts. Finally we conclude in Section 6. For completeness, we also have two appendices where we study respectively the accuracy of our emulator and the impact of varying the smoothing scale of WL maps.

## 2 THEORY

The lens equation for a gravitationally lensed image is

$$\boldsymbol{\alpha} = \boldsymbol{\beta} - \boldsymbol{\theta}, \quad (1)$$

where  $\boldsymbol{\alpha}$  is the deflection angle between  $\boldsymbol{\beta}$ , the true position of the source on the sky, and  $\boldsymbol{\theta}$ , the observed position of the lensed image. The corresponding Jacobian matrix of the (linear) lens mapping is the deformation matrix  $\mathbf{A}$ ,

$$A_{ij} = \frac{\partial \beta_i}{\partial \theta_j} = \delta_{ij} - \frac{\partial \alpha_i}{\partial \theta_j}. \quad (2)$$

Under the Born approximation and neglecting lens-lens coupling and other second-order effects, the deflection angle can be expressed as the gradient of a 2D lensing potential  $\psi$ ,

$$\boldsymbol{\alpha} = \nabla \psi, \quad (3)$$

where  $\psi$  is given by

$$\psi(\boldsymbol{\theta}, \chi) = \frac{2}{c^2} \int_0^\chi \frac{\chi - \chi'}{\chi \chi'} \Phi(\chi' \boldsymbol{\theta}, \boldsymbol{\theta}) d\chi'. \quad (4)$$

Here,  $\chi$  is the comoving distance from the observer to the source and  $\chi'$  is the comoving distance from the observer to the continuously-distributed lenses, which is also the integration variable.  $\Phi$  is the 3D

lensing potential of the lens, and  $c$  the speed of light.  $\Phi$  is related to the non-relativistic matter density contrast,  $\delta = \rho/\bar{\rho} - 1$ , through the Poisson equation

$$\nabla^2 \Phi = 4\pi G a^2 \bar{\rho} \delta, \quad (5)$$

where  $\rho$  is the matter density of the Universe (with a bar denoting the mean),  $G$  is the gravitational constant and  $a$  is the scale factor.

Eq. (4) shows that the lensing potential is a line-of-sight integral of the matter distribution from the source to the observer. The contribution that matter at distance  $\chi'$  along the line of sight makes to the total lensing potential is weighted by  $(\chi - \chi')/\chi\chi'$  and so depends on its distances from the source and observer.

Eq.(3) allows Eq. (2) to be expressed in terms of  $\psi$

$$A_{ij} = \delta_{ij} - \partial_i \partial_j \psi, \quad (6)$$

where partial derivatives are taken with respect to  $\theta$ . The matrix  $\mathbf{A}$  can be parameterised through the more physically instructive terms convergence,  $\kappa$ , and shear,  $\gamma = \gamma_1 + i\gamma_2$ , as

$$\mathbf{A} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}. \quad (7)$$

This parameterisation allows the convergence and shear to be related to the lensing potential via

$$\kappa \equiv \frac{1}{2} \nabla_{\theta}^2 \psi, \quad (8)$$

and

$$\gamma_1 \equiv \frac{1}{2} (\nabla_{\theta_1} \nabla_{\theta_1} - \nabla_{\theta_2} \nabla_{\theta_2}) \psi, \quad \gamma_2 \equiv \nabla_{\theta_1} \nabla_{\theta_2} \psi, \quad (9)$$

where  $\nabla_{\theta} \equiv (\chi')^{-1} \nabla$ . Eq. (8) can be interpreted as a 2D Poisson equation, and so by substituting Eq. (5) and Eq. (4) into Eq. (8), the convergence can be expressed in terms of the matter overdensity

$$\kappa(\theta, \chi) = \frac{3H_0^2 \Omega_m}{2c^2} \int_0^{\chi} \frac{\chi - \chi'}{\chi} \chi' \frac{\delta(\chi' \theta, \chi')}{a(\chi')} d\chi'. \quad (10)$$

This shows that the observed WL convergence can be interpreted as the projected density perturbation along the line of sight, weighted by the lensing efficiency factor  $(\chi - \chi')\chi'/\chi$ . Here, the lensing efficiency is greatest at  $\chi' = \chi/2$ , when the lens is halfway between the source and the observer.

The above derivation assumes a fixed source plane. However, in real WL observations, the source galaxies do not occupy a single plane at a fixed distance from the observer. The observed catalogue of source galaxies has a probability distribution  $n(\chi)$  that spans over a range of  $\chi$  values, and Eq. (10) must be weighted by this source galaxy distribution in order to obtain  $\kappa(\theta)$  (see, e.g., Kilbinger 2015, for details)

$$\kappa(\theta) = \int_0^{\chi} n(\chi') \kappa(\theta, \chi') d\chi'. \quad (11)$$

In this work, we measure the  $\kappa$  profile,  $\kappa(r)$ , in and around WL voids. However, as  $\kappa(r)$  is not directly observable, it is also useful to relate it to the radial tangential shear profile,  $\gamma_t(r)$ , through

$$\gamma_t(r) = \bar{\kappa}(<r) - \kappa(r), \quad (12)$$

where

$$\bar{\kappa}(<r) = \frac{1}{\pi r^2} \int_0^r 2\pi r' \kappa(r') dr' \quad (13)$$

is the mean enclosed convergence within radius  $r$ . Notice that here and throughout this paper we use  $r$  rather than  $\theta$  to represent the 2D distance from the void centre.

WL observations rely on accurately measuring the shapes of galaxies, and cross correlating the shapes of neighbouring galaxies. However, any correlation in shape due to lensing is dominated by the random shapes and orientations of galaxies, which is the leading source of noises in WL observations, referred to as galaxy shape noise (GSN). Since the lensing signal is weak by definition, when identifying WL peaks (local maxima in the convergence field  $\kappa(\theta)$ ) it is convenient to express the convergence relative to the standard deviation of the corresponding GSN component of the field. This is given by

$$v = \frac{\kappa}{\sigma_{\text{GSN}}} \quad (14)$$

where  $\sigma_{\text{GSN}}$  is the standard deviation of the contributions to the signal from galaxy shape noise.  $\sigma_{\text{GSN}}$  can be calculated by generating mock GSN maps and applying any transformations also applied to the convergence maps, such as smoothing. Mock GSN maps are generated by assigning to pixels random convergence values from a Gaussian distribution with standard deviation

$$\sigma_{\text{pix}}^2 = \frac{\sigma_{\text{int}}^2}{2\theta_{\text{pix}} n_{\text{gal}}}, \quad (15)$$

where  $\theta_{\text{pix}}$  is the width of each pixel,  $\sigma_{\text{int}}$  is the intrinsic ellipticity dispersion of the source galaxies, and  $n_{\text{gal}}$  is the measured source galaxy number density. In this work we use  $\sigma_{\text{int}} = 0.28$  and  $n_{\text{gal}} = 20 \text{ arcmin}^{-2}$  as will be discussed in Section 3.1.

It will also be useful to compare constraints from Eq. 12 to the standard shear two-point correlation function, which is given by

$$\xi_{\pm}(\theta) = \langle \gamma_t \gamma_t \rangle \pm \langle \gamma_{\times} \gamma_{\times} \rangle = \frac{1}{2\pi} \int_0^{\infty} dl l P_{\kappa}(l) J_{0,4}(l\theta), \quad (16)$$

where  $\gamma_t = -\mathbb{R}(\gamma e^{-2i\phi})$  (equivalent to Eq. (12), but presented for completeness),  $\gamma_{\times} = -\mathbb{I}(\gamma e^{-2i\phi})$ ,  $\phi$  is the polar angle of the separation vector  $\theta$ ,  $J_0$  and  $J_4$  are the Bessel functions for  $\xi_+$  and  $\xi_-$  respectively, and  $l$  is the Fourier mode.

### 3 METHODOLOGY

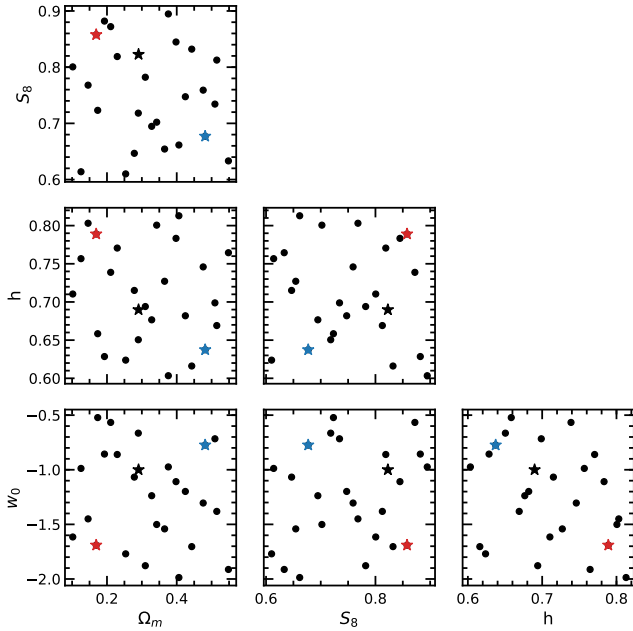
In this section we describe the methodology followed in this work, including the simulations and mock lensing data, likelihood analysis and the weak lensing void (tunnels) finding algorithm.

#### 3.1 Mock Data

In this work we use mock WL convergence maps generated from the *cosmo-SLICS* simulation suite (Harnois-Déraps et al. 2019), which we briefly outline in this subsection.

The *cosmo-SLICS* suite is a set of N-body dark-matter-only simulations run for 26 cosmology nodes in the  $[\Omega_m, S_8, h, w_0]$  parameter space. Here  $\Omega_m$  is the matter density parameter today,  $S_8 = \sigma_8(\Omega_m/0.3)^{0.5}$ ,  $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$  is the reduced Hubble constant, and  $w_0$  is the dark energy equation of state parameter, which is assumed to be a constant. The  $\sigma_8$  parameter is the present-day root-mean-squared matter density perturbation smoothed on  $8h^{-1} \text{ Mpc}$  scales.

The four dimensional parameter space is sampled using a Latin hypercube, which is a sampling algorithm designed to give a high interpolation accuracy for a low node count. The exact cosmological parameter space that is modelled by each simulation node is shown in Fig. 1. At each node, a carefully-designed pair of simulations



**Figure 1.** The four dimensional parameter space ( $[\Omega_m, S_8, h, w_0]$ ) sampled by the *cosmo*-SLICS simulation suite. The fiducial cosmology is indicated by a star with parameter values  $[0.29, 0.82, 0.69, -1.00]$ . We have highlighted two additional nodes with blue  $[0.48, 0.68, 0.64, -0.77]$  and red  $[0.17, 0.86, 0.79, -1.69]$  stars, which are selected as nodes in separate regions of the parameter space, used to exemplify the behavior of WL voids as a function of cosmological parameters.

are run, for which sampling variance is highly suppressed. This is achieved by selecting a pair of initial conditions out of a large number of random realisations, such that the mean matter power spectrum closely matches the ensemble average. The random phases of this pair of initial conditions are used for all cosmology nodes. The simulation volume is a cube with length  $L = 505 h^{-1} \text{Mpc}$ , with  $N = 1536^3$  dark matter particles.

For each node, 50 pseudo-independent light-cones are constructed by resampling projected mass sheets, which are then ray-traced under Born approximation to construct lensing maps and catalogues (see [Harnois-Déraps et al. 2019](#), for full details about the light-cone and catalogue construction).

We use the *cosmo*-SLICS source catalogue down-sampled to match LSST specifications with a source redshift distribution of  $z_s = [0.6, 1.4]$ , which gives a conservative source galaxy number density of  $20 \text{ arcmin}^{-2}$ . From this we generate 50 WL convergence maps for each of the 26 cosmology nodes, with a sky coverage of  $10 \times 10 \text{ deg}^2$  each and pixel grid of dimensions  $3600^2$  ([Giblin et al. 2018](#)). These maps are smoothed with a Gaussian filter with smoothing scale  $\theta_s = 1 \text{ arcmin}$ .

### 3.2 Emulation and likelihood analysis

In this subsection, we outline the procedure used to test the sensitivity of WL void statistics to the cosmological parameters  $\Omega_m$ ,  $S_8$ ,  $h$  and  $w_0$ .

The first step is to measure the WL void statistics from the 50 convergence maps for each of the 26 *cosmo*-SLICS cosmologies shown in Fig. 1. Then, in order to make predictions of the WL void statistics at arbitrary points in the 4D parameter space shown in Fig. 1, we use a Gaussian process (GP) regression emulator from *scikit-learn* ([Pedregosa et al. 2011](#)) to interpolate the void statistics

between nodes. GP regression is a non-parametric Bayesian machine learning algorithm used to make probabilistic predictions that are consistent with the training data (see, e.g., [Habib et al. 2007](#); [Schneider et al. 2008](#), for some of its early applications in cosmology). The emulator requires the training data to sample the parameter space sufficiently, and generally the accuracy of the emulator is limited by the availability of training data. The accuracy of the GP emulator trained on *cosmo*-SLICS was tested extensively and found to yield few per cent accuracy in its predictions of weak lensing two-point correlation functions ([Harnois-Déraps et al. 2019](#)), density split statistics ([Burger et al. 2020](#)) persistent homology statistics ([Heydenreich et al. \(2020\)](#)) and aperture mass statistics ([Martinet et al. 2020](#)). In this work the average void statistics and their standard errors at each set are used as the training data for the emulator. We present results showing the accuracy of the emulator in Appendix A.

Finally, once the emulator has been trained we use Monte Carlo Markov Chain (MCMC) to estimate the posteriors of the parameters for the entire parameter space and produce likelihood contours. We use the *emcee* python package ([Foreman-Mackey et al. 2013](#)) to conduct the MCMC analysis in this work sampling the 4D parameter space as follows.

We employ a Bayesian formalism, in which the likelihood,  $P(\mathbf{p}|\mathbf{d})$ , of the set of cosmological parameters  $\mathbf{p} = [\Omega_m, S_8, h, w_0]$  given a data set  $\mathbf{d}$ , is given, according to Bayes' theorem, by

$$P(\mathbf{p}|\mathbf{d}) = \frac{P(\mathbf{p})P(\mathbf{d}|\mathbf{p})}{P(\mathbf{d})}, \quad (17)$$

where  $P(\mathbf{p})$  is the prior,  $P(\mathbf{d}|\mathbf{p})$  is the likelihood of the data conditional on the parameters, and  $P(\mathbf{d})$  is the normalisation. In our analysis we use flat priors with upper and lower limits respectively for  $\Omega_{rmm}$ :  $[0.10, 0.55]$ ,  $S_8$ :  $[0.61, 0.99]$ ,  $h$ :  $[0.45, 0.90]$ ,  $w_0$ :  $[-1.99, -0.52]$ . We note that these priors slightly extend past the parameter space in Fig. 1, which we have expanded to ensure the likelihood contours presented in this work to not cross the prior boundary.

The log likelihood can be expressed as

$$\log(P(\mathbf{d}|\mathbf{p})) = -\frac{1}{2} [\mathbf{d} - \mu(\mathbf{p})] C^{-1} [\mathbf{d} - \mu(\mathbf{p})], \quad (18)$$

where  $\mu(\mathbf{p})$  is the prediction generated by the emulator for a set of parameters  $\mathbf{p}$ , and  $C^{-1}$  is the inverse covariance matrix. In practice we use the emulator's prediction of the void statistics at the fiducial cosmology as the data  $\mathbf{d}$ . The likelihood returns a 4D probability distribution that indicates how well different regions of the parameter space describe the input data  $\mathbf{d}$ . Note that in Eq. (18) we have assumed that the covariance matrix does not vary with a change in the cosmological parameters.

We calculate the covariance matrices from the 50 WL map realisations from the fiducial cosmology, and divide it by a factor of 180 in order to rescale the covariance matrix from a  $100 \text{ deg}^2$  area to the LSST survey area, which we take as  $18,000 \text{ deg}^2$ . We also multiply the inverse covariance matrix by a factor  $\alpha$ , which accounts for the bias that is present when inverting a noisy covariance matrix ([Anderson 2003](#); [Hartlap et al. 2007](#)), given by:

$$\alpha = \frac{N - N_{\text{bin}} - 2}{N - 1}, \quad (19)$$

where  $N = 50$  is the number of weak lensing maps that have been used to calculate the covariance matrix and  $N_{\text{bin}}$  is the number of bins for which the statistic is computed. We note however that [Sellentin & Heavens \(2016\)](#) present an alternative approach to robustly account for the uncertainty in the estimated covariance, via a student- $t$  likelihood distribution.

### 3.3 The tunnel algorithm

To identify WL voids, we use the tunnel algorithm initially proposed in Cautun et al. (2018), which identifies the largest circles in a 2D tracer catalogue that are empty of tracers. We choose to use this void finding algorithm since, compared with several other common 2D void finders, it gives void lensing profiles with high SNR, whilst also being least affected by the observational noises associated with weak lensing measurements, such as galaxy shape noise (Davies et al. 2020). The tunnel algorithm requires an input tracer catalogue to identify voids. For the identification of WL voids, we use WL peaks as tracers of the underlying convergence field (this avoids the necessity to have a synthetic 2D galaxy map for this analysis). Here we define WL peaks as local maxima in the WL convergence map as in Eq. (14).

To identify tunnels, we first construct a Delaunay triangulation of the tracers (WL peaks). This produces a unique tessellation of the map with triangles, where each vertex is a tracer and the tessellated triangles enclose no tracers. From each triangle, a corresponding circumcircle can be defined, which is a circle that is directly on top of its Delaunay triangle with all vertices of the latter residing on the circumcircle's circumference. This tessellation is unique, and by definition gives circles that do not enclose any tracers. To avoid identifying the same regions as voids multiple times, we discard any circumcircles whose centers reside inside a larger circumcircle. The resulting list represents our tunnel catalogue, where each tunnel is characterized in terms of the centre and radius of its corresponding circumcircle.

The WL peak catalogues that may be used to identify tunnels contain peaks with a range of amplitudes (or heights)  $\nu$ . WL peaks of different amplitudes trace different components of the WL map, where the peaks with low or negative amplitudes trace underdense regions of the map, and those with high amplitudes trace overdense regions. Furthermore, peaks with low amplitudes are more susceptible to either being created or contaminated by GSN. It is therefore convenient to generate multiple sub-catalogues of a given WL peak catalogue, by retaining only the peaks with amplitudes larger than a given  $\nu$  value. Varying the  $\nu$  thresholds allows us to study how the tunnels respond to tracer catalogues with different properties. In this work we use WL peak catalogues with amplitudes of  $\nu > 1, 2$  and 3 to identify tunnels, and will also use these  $\nu$  values to denote the corresponding tunnel catalogues.

In Fig. 2 we show a visualisation of tunnels identified from catalogues of WL peaks with amplitudes  $\nu > 2$ . The figure shows WL maps, WL peaks and tunnels for the fiducial cosmology (left), and two sample cosmologies, blue (middle) and red (right) to exemplify the impact of changing cosmological parameters. Here it can be seen in the bottom left part of the panels that the red cosmology, which has the highest  $S_8$  value of the three reference cosmologies, contains more overdense (orange) regions than the other two cosmologies. The increased overdensity in the bottom left of the red-cosmology panel leads to more WL peaks, which generates more small tunnels and fewer large tunnels relative to the other two cosmologies. This highlights how changing the cosmological parameters changes the structure observed in WL maps and the corresponding WL void properties.

## 4 WEAK LENSING VOID STATISTICS

In this section we present the weak lensing void statistics used in this analysis, showing their abundance in Section 4.1, and the tangential shear profiles in Section 4.2.

**Table 1.** Forecast of percentage uncertainties obtained from various WL void statistics for an LSS-like survey. The first block of 4 rows show 68% CL while the bottom 4 rows show 95% CL. In each block, the results shown in the first three lines are quoted from the tightest contours in each figure in Section 5 (see first column for more details). In the last line of each block, ‘ $\gamma$ -2PCF’ stands for the parameter constraints using the cosmic shear two-point correlation functions for the same maps as used for the cosmic void statistics.

Statistic	$\Omega_m$	$S_8$	$h$	$w_0$
<b>68% confidence limits</b>				
$dn/dR_\nu$ (combined)	4.3%	0.7%	4.7%	7.0%
$\gamma_t$ (combined)	6.8%	1.3%	6.3%	12.5%
$dn/dR_\nu$ and $\gamma_t$ ( $\nu > 1$ )	4.9%	0.8%	5.0%	6.9%
$\gamma$ -2PCF	4.5%	0.8%	4.1%	7.8%
<b>95% confidence limits</b>				
$dn/dR_\nu$ (combined)	8.3%	1.4%	8.9%	13.5%
$\gamma_t$ (combined)	13.3%	2.6%	12.3%	25.1%
$dn/dR_\nu$ and $\gamma_t$ ( $\nu > 1$ )	9.7%	1.6%	10.0%	13.9%
$\gamma$ -2PCF	12.6%	2.2%	11.5%	23.0%

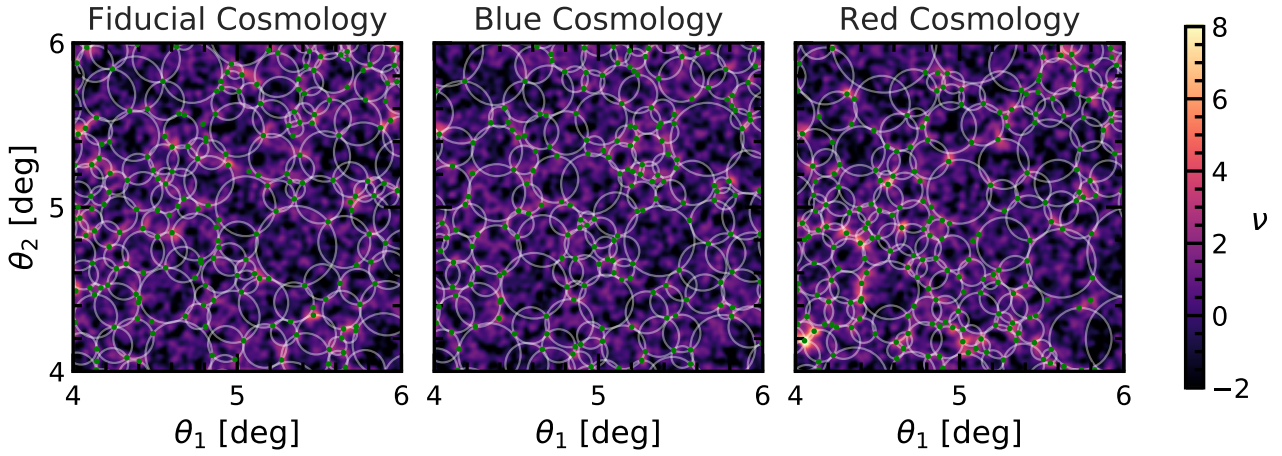
### 4.1 Void abundance

Fig. 3 shows the differential void abundance per unit area as a function of void radius. The three panels correspond to voids identified in three WL peak catalogues, with peak heights  $\nu > 1, 2$  and 3. Void abundances for each of the nodes in Fig. 1 are plotted in grey, the fiducial cosmology in black and two sample cosmologies in colour (blue and red – corresponding to the two cosmologies in the middle and right panels of Fig. 2).

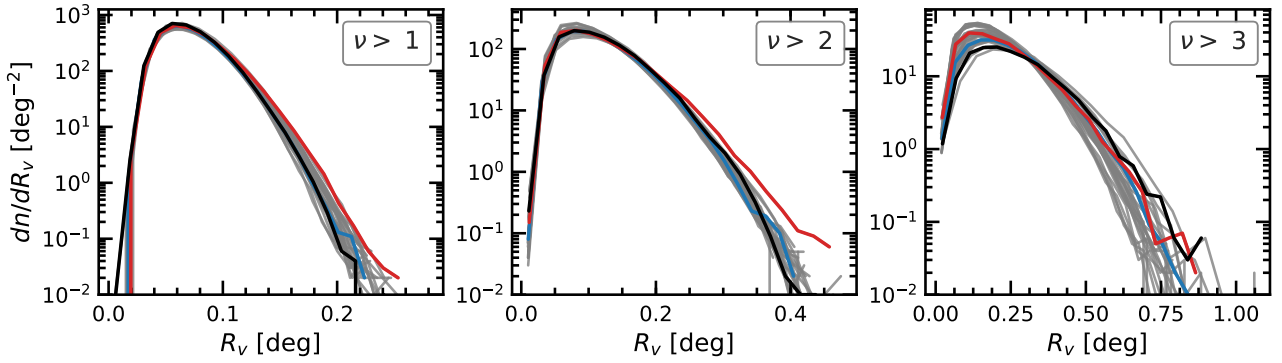
The figure shows that as the  $\nu$  threshold increases, the total number of WL voids decreases (given by the area under the curves), and the average size of the voids increases. The spread in the void abundances over all cosmologies is largest for the  $\nu > 3$  catalogue. However, the data is also noisier in this catalogue, because there are fewer peaks with  $\nu > 3$  and subsequently fewer tunnels.

The red cosmology produces more large voids for the  $\nu > 1$  and 2 catalogues than the fiducial and blue cosmologies. However, the same behaviour is not seen for the  $\nu > 3$  catalogue, which may indicate that the sensitivity of the void abundance to specific cosmological parameters changes as fewer tracers are used to identify WL voids. The red cosmology has the largest  $S_8$  and smallest  $\Omega_m$  compared to the fiducial and blue cosmologies. Increasing  $S_8$  or  $\Omega_m$  increases the clustering of matter which leads to a wider range of WL void sizes, as we have seen in Fig. 2: this is because the enhanced clustering creates more peaks with  $\nu > 1$  or 2 in dense regions, reducing the void sizes there, and at the same time reduces the amplitudes of some low peaks in underdense regions, increasing void sizes there. On the other hand, for the  $\nu > 3$  catalogue, the peaks are sparser in all three cosmologies (hence voids are larger), and the fact that the red cosmology has more peaks at  $\nu > 3$  again restricts the sizes of its voids, this time affecting the largest ones.

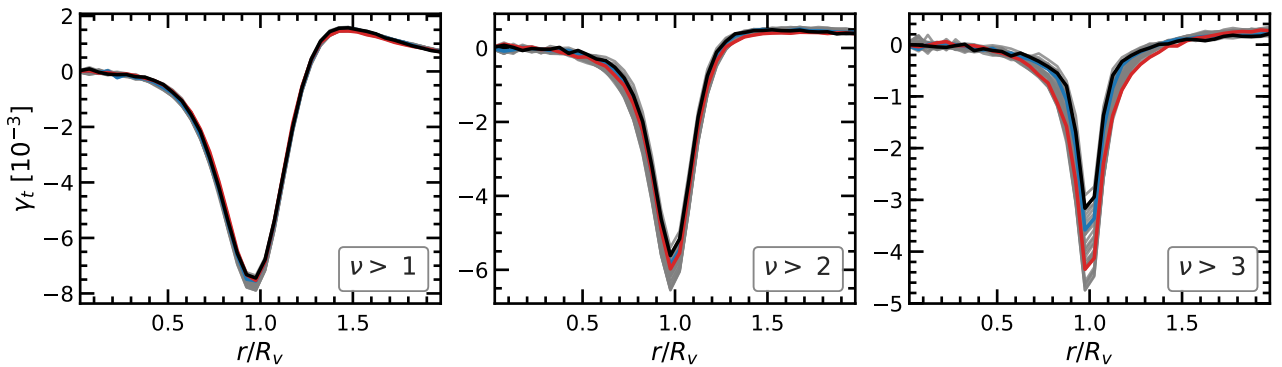
For  $\nu > 1$  the fiducial cosmology produces the fewest large voids compared to the red and blue cosmologies, however for  $\nu > 3$  it produces the most large voids. The change in relative behaviour between the fiducial, red and blue cosmologies as the  $\nu$  threshold increases, indicates that void abundances measured from different WL peak catalogues contain complementary information to each other. We will see this point more clearly later when looking at the constraints from void abundances.



**Figure 2.** (Colour Online) A visualisation of WL peaks (green points) used to identify the tunnels (white circles) in the WL convergence maps (colour map). The left panel shows tunnels for the fiducial cosmology, while the middle and right panels show tunnels for the blue and red cosmologies highlighted in Fig. 1 respectively. The colour bar on the right indicates the convergence field in units of  $\nu = \kappa/\sigma_{\text{GSN}}$ .



**Figure 3.** (Colour Online) The differential void abundance as a function of void radius  $R_V$ . The three panels correspond to voids identified in different WL peak catalogues, with peak heights  $\nu > 1, 2$  and  $3$  (from left to right). The void abundances for all cosmologies in Fig. 1 are plotted in grey. Results for the fiducial (black), red (red) and blue (blue) cosmologies are over-plotted in colour.

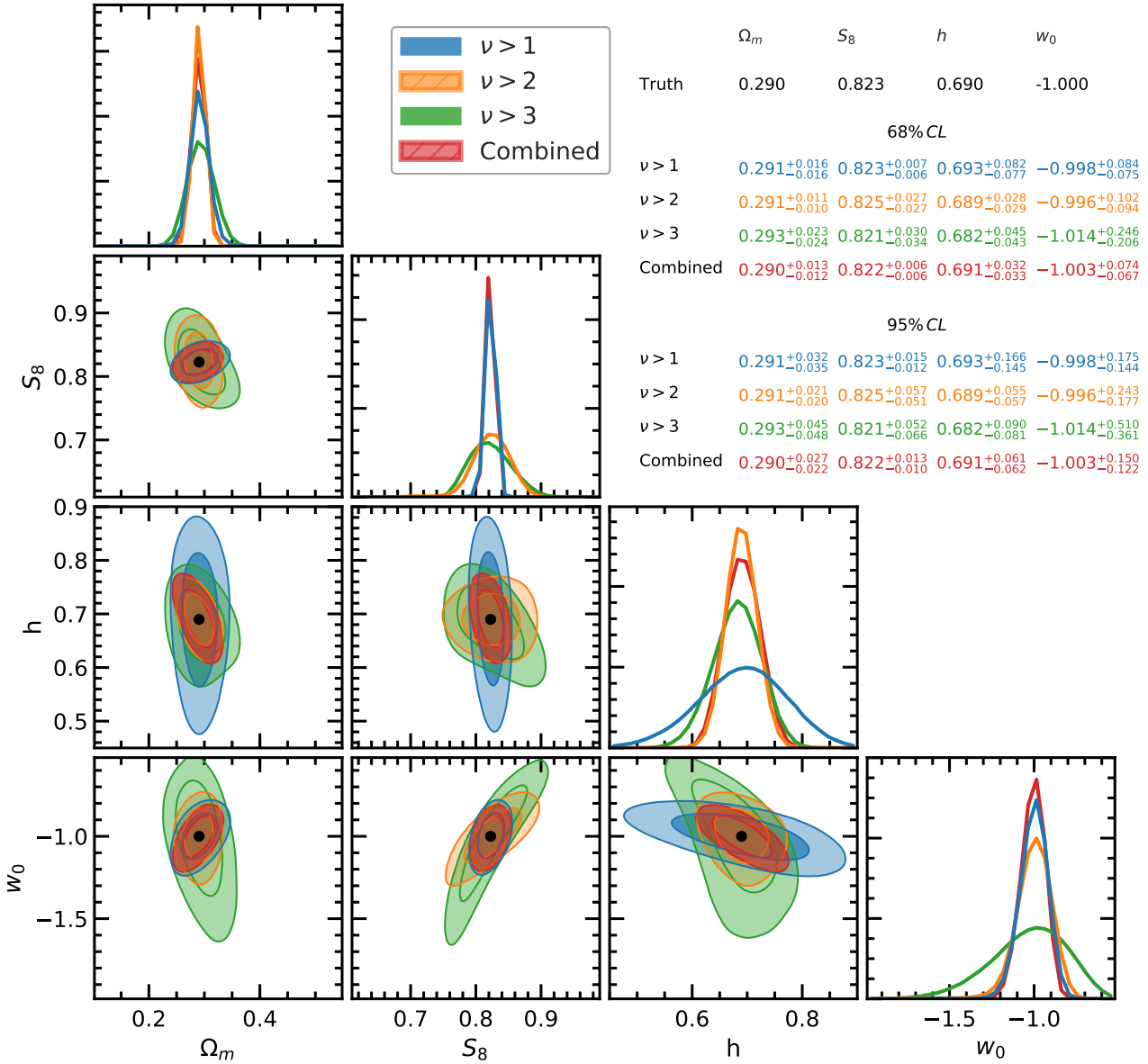


**Figure 4.** (Colour Online) The tangential shear profiles as a function of re-scaled distance to void centre,  $r/R_V$ . The three panels correspond to voids identified in three WL peak catalogues with peak heights  $\nu > 1, 2$  and  $3$  (from left to right). The tangential shear profiles for all cosmologies in Fig. 1 are plotted in grey. Results for the fiducial cosmology (black), red (red) and blue (blue) cosmologies are over-plotted in colour.

#### 4.2 Lensing tangential shear profiles

Fig. 4 shows the tangential shear profiles for WL voids, where the panels (from left to right) show WL voids identified in the  $\nu > 1, 2$  and  $3$  catalogues. Tangential shear profiles for all cosmologies are plotted in grey, with the fiducial and two highlighted cosmologies

plotted in colour as in Fig. 3. The tangential shear profiles are calculated by first measuring the convergence profiles in annuli centered on the void center (pixels are interpolated for small annuli), where the number of annuli used is the lensing profile bin number. The annuli are then stacked as a function of relative angular size



**Figure 5.** (Colour Online) Constraint forecasts on cosmological parameters measured from void abundances. Contours are shown for WL voids identified in WL peak catalogues with  $\nu > 1$  (blue),  $\nu > 2$  (orange),  $\nu > 3$  (green) and the combination of all three catalogues (red). The true cosmological parameter values used to generate the data are indicated by the black point. The diagonal panels show the 1D marginalised probability distribution, and remaining panels show the marginalised 2D probability contours enclosing the 68% and 95% confidence intervals. The table in the top right shows true parameter values (top) and the inferred parameter values for the different peak catalogues with 68% (upper section) and 95% (lower section) confidence limits.

( $r/R_V$ ), weighted by their corresponding void area. Using Eq. 12, this is then converted to the tangential shear profiles.

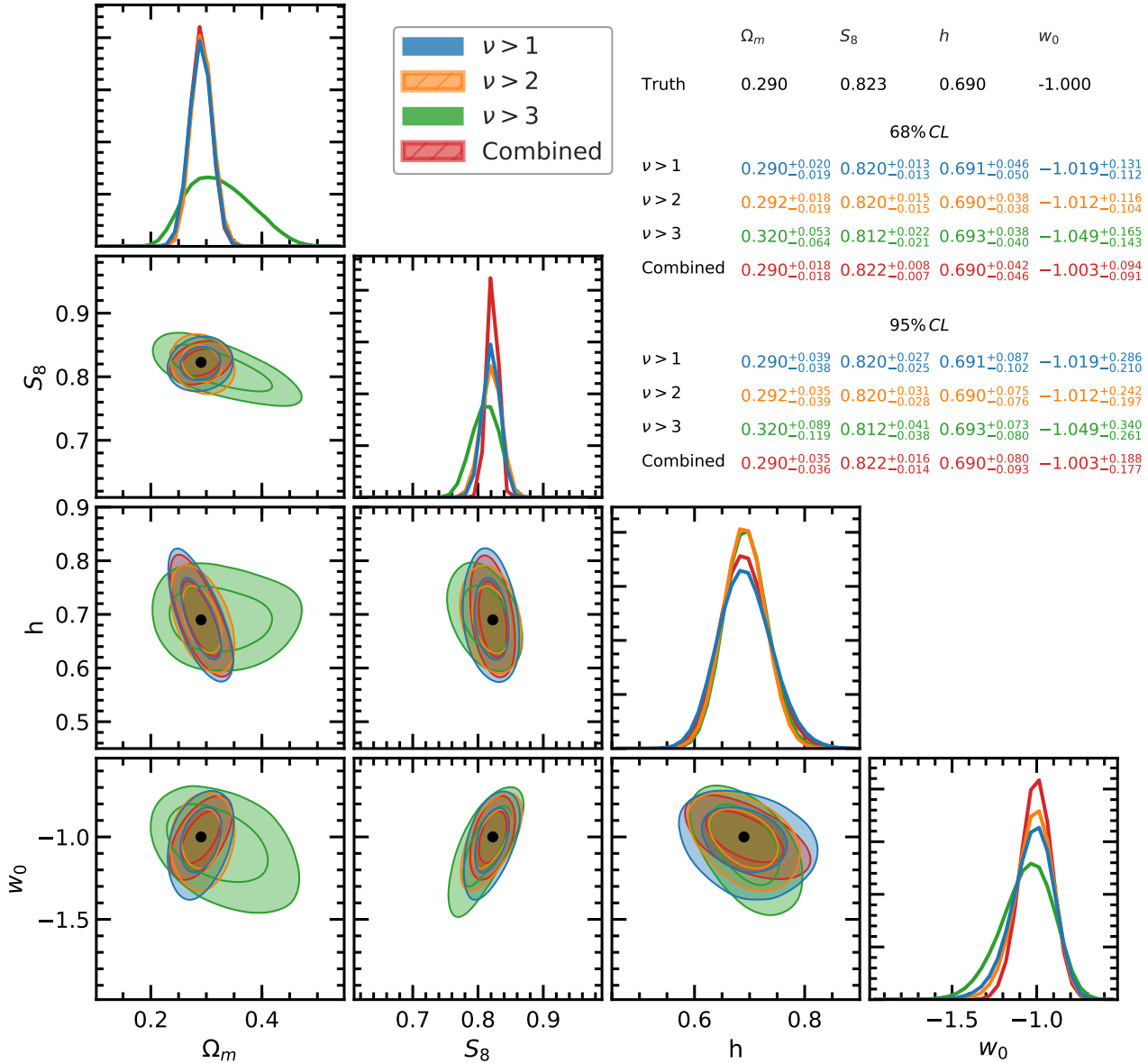
The tangential shear profiles plotted here are negative, which indicates that the WL voids behave like concave lenses and their interiors correspond to underdense regions. The figure shows that as the  $\nu$  threshold increases, the depth of the tangential shear profiles at  $r/R_V = 1$  decreases, but meanwhile the spread in the amplitude of the tangential shear profiles, as well as the spread in the width of the peak around  $r/R_V = 1$ , amongst all cosmologies, increases. Note that the peaks of the tangential shear profiles appear to be narrower as the  $\nu$  threshold increases, but this is an artificial consequence of the fact that these plots are made against  $r/R_V$  with the void radius  $R_V$  larger for larger  $\nu$  thresholds.

For the  $\nu > 1$  catalogue, the fiducial, red and blue cosmologies

all lie on top of each other. For  $\nu > 2$ , the red cosmology (with the largest  $S_8$  value) has a deeper tangential shear profile compared to the other two cosmologies. For  $\nu > 3$  the difference in amplitude increases further between the three reference cosmologies, with the fiducial cosmology having the lowest (absolute) amplitude, however the general trend between the three reference cosmologies is the same for all  $\nu$  thresholds. Part of this can again be attributed to the high  $S_8$  value in the red cosmology, which enhances the clustering of matter, resulting in low-density regions becoming more underdense. However, the three reference cosmologies have very distinct values for the other three parameters, in particular  $\Omega_m$ , which means that an intuitive and yet complete explanation of their relative behaviours is difficult to gauge by eye.

The observation that, although the spread in shapes amongst





**Figure 6.** (Colour Online) The same as Fig. 5 but for the tangential shear profiles. See the caption in Fig. 5 for more details

all cosmologies increases with the  $\nu$  threshold, the general order in which they appear does not change, indicates that there may not be much complementary information between tangential shear profiles measured from different  $\nu$  thresholds. Also, while the differences between the different cosmologies are larger for the  $\nu > 3$  catalogue, the fact that there are relatively fewer voids in this catalogue means that its constraining power is not necessarily stronger than the other two catalogues, as we will see shortly.

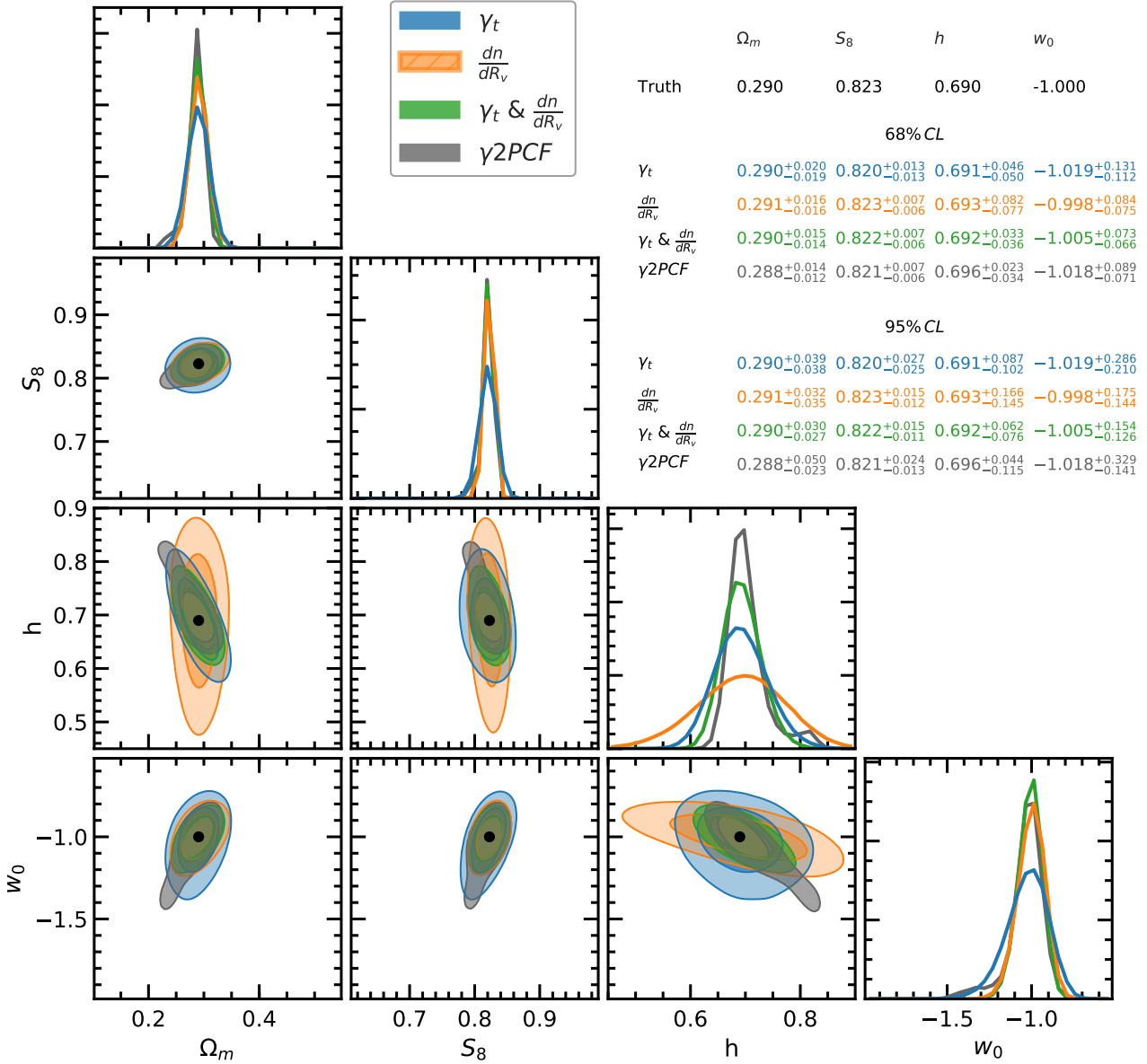
## 5 PARAMETER CONSTRAINTS FORECAST

In this section we present parameter constraint forecasts for an LSST-like survey from the void abundances and tangential shear profiles of WL voids, as well as their combinations.

### 5.1 Void abundance constrains

Fig. 5 shows the likelihood contours for measuring the four cosmological parameters with WL void abundance. The diagonal panels of the figure show the 1D marginalized likelihood distribution and remaining panels show the 2D marginalized likelihood contours. For each inference case, the inner and outer contours indicate respectively the 68% and 95% confidence limits (CL). As mentioned above, we use the fiducial cosmology as our ‘observed’ data set, which is indicated by the black point. The figure shows results for three  $\nu$  thresholds with  $\nu > 1$  (blue),  $\nu > 2$  (orange) and  $\nu > 3$  (green). We also show results for the combination of all three catalogues (red). The table in the top right of the figure indicates the estimated cosmological parameters with their corresponding 68% (top) and 95% (bottom) CL, for each of the contours.

All of the contours presented in this figure are centered on the true (fiducial) cosmology, indicating that the WL void abundances are able to measure cosmological parameters without any bias. The



**Figure 7.** (Colour Online) The same as Fig. 5 but for the tangential shear profiles (blue), void abundance (orange), and combination of the two (green). Results are shown for  $\nu > 1$  WL peaks. See the caption in Fig. 5 for more details. Note that, for comparison, we have added the contours from the shear-shear two-point correlation function (without tomography) extracted from the same maps in grey colour, and the corresponding constraints on the parameters are also listed in the table in grey.

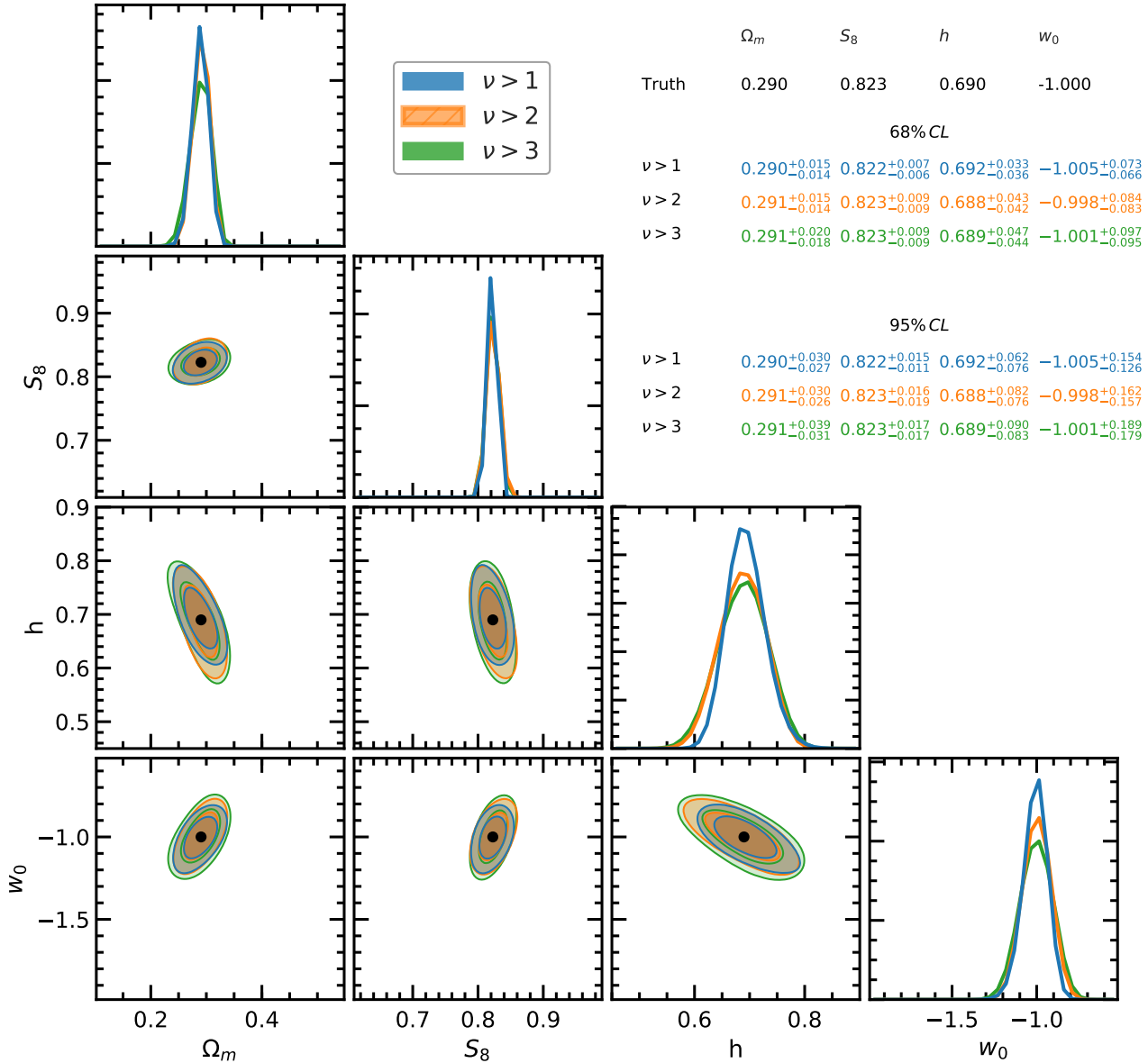
tightest contours are for the  $\Omega_m - S_8$  plane, which is expected since these are the cosmological parameters to which WL analysis is the most sensitive.

For nearly every combination of parameters, the three contours for the  $\nu > 1, 2$  and 3 void catalogues occupy different parts of the marginalised 2D parameter space, or have different degeneracy directions, where most of the overlapping occurs around the true values. As suggested by the behaviour of the three reference cosmologies discussed in Section 4.1, this indicates that void abundances measured from different catalogues contain complementary information to each other. We therefore also show parameter likelihood contours for the combination of the WL void abundances from the three catalogues in red.

The ‘Combined’ contours are smaller than any of the individual contours, for all combinations of parameters. This shows that

parameter measurements from the WL void abundances are significantly improved when multiple catalogues are used. The 68% and 95% CL percentage accuracy that the combined WL void abundance is able to measure the parameters with is shown in Table 1. We note that when combining multiple catalogues, we had to reduce the number of bins used for each catalogue to avoid having a data vector with more bins than the number of WL map realisations that we had available for calculating the covariance matrix. For the individual catalogues we use 20 bins, and when combining catalogues we use 10 bins per catalogue giving 30 in total. This coarser binning per catalogue means that the red contours presented here may be larger than the ‘true’ contours. We plan to run further simulations and make more mock WL maps in the future, to improve the constraint here.

Theoretically, the abundance of tunnels identified from a WL



**Figure 8.** (Colour Online) The same as Fig. 5 but for the combination of the tangential shear profiles and the void abundance. Results are shown for the three WL Peak catalogues with  $\nu > 1$  (blue),  $\nu > 2$  (orange),  $\nu > 3$  (green). See the caption in Fig. 5 for more details.

peak catalogue depends not only on the number of peaks, but also on their clustering pattern. We therefore expect that the information contained in WL void abundance and peak correlation functions may have substantial overlapping. The latter probe has been studied in details by Davies et al. (2019a), with certain scaling properties observed. While it is beyond the scope of the current work, we will conduct a similar analysis by forecasting the parameter constraining power by WL peak two-point correlation functions in a followup study.

## 5.2 Tangential shear constraints

Fig. 6 shows likelihood contours for the four cosmological parameters from tangential shear profiles. The colours of the contours correspond to the same void catalogues as in Fig. 5. Again, the contours are smallest in the  $\Omega_m$ - $S_8$  plane, and each contour is roughly

centered on the true parameter values indicated by the black point, showing that the tangential shear profiles are able to measure these parameters without any bias. The figure shows that the contours from  $\nu > 1$  and  $\nu > 2$  are similar in size, and the  $\nu > 3$  contours are significantly larger and in most cases entirely enclose the other contours. All of the contours in this figure, unlike in the case of void abundances, occupy similar regions of the parameter space, or have similar degeneracy directions. This confirms our conclusion based on observation of Fig. 4, namely the tangential shear profiles from different peak catalogues do not offer much complementarity.

As in Section 5.1, we combine the tangential shear profiles from all three catalogues to generate ‘Combined’ likelihood contours. Note that for individual catalogues the tangential shear profiles are calculated with 30 bins each, and when combined, individual catalogues are plotted with 10 bins each, giving 30 in total. Again, this is to avoid having more data bins than WL map realisations used to calculate the covariance matrix. We expect this approach

might give contours that are slightly larger than the truth, due to the coarser binning. By combining catalogues, we find only a small improvement in contour size relative to the  $\nu > 1$  catalogue, which again suggests that there is very little complementary information between the different  $\nu$  catalogues for the tangential shear profiles.

The strongest constraints from the tangential shear profiles are for the combined contour. We summarise the 68% and 95% CL for the  $\gamma_t$  combined case in Table 1.

### 5.3 Constraints by combining void abundance and tangential shear

In this section we present parameter constraint forecasts for the combination of the WL void abundance and tangential shear profiles.

Fig. 7 shows contours for the tangential shear profiles (blue) WL void abundance (orange) and the combination of the two (green) for the  $\nu > 1$  catalogue. We also include the shear-shear two-point correlation function constraints as a comparison. We follow Asgari et al. (2020) and sample the 2PCF using 9 logarithmically-spaced angular separation bins from 0.5 to 300 arcmin, and use only the  $\xi_+$  correlation function. Note that the WL void abundance and tangential shear profile contours shown in this figure are also presented as the blue contours in Fig. 5 and Fig. 6 respectively. We include these contours in this figure to illustrate how the contours behave when combining the two probes.

For all combinations of parameters, the WL void abundance contours and the tangential shear contours occupy similar regions of the parameter space and have similar degeneracy directions. However there is a large difference in size between the two for all panels that include  $h$ .

The combined contours are smaller than the contours from the tangential shear profiles alone, which shows that there is a benefit of combining these two probes. When comparing the tangential shear contours and the combined contours, the parameter with the greatest relative increase in accuracy is  $S_8$ . This is shown in the table in the top right of the figure, where it can be seen that the combined errors are roughly half the tangential shear errors, although this seems to be mainly due to the fact that void abundance alone can give much stronger constraint on  $S_8$ . Finally, the parameter with the greatest absolute increase in accuracy enabled by this combination is  $w_0$ , which suggests that using different probes to strengthen constraints could be particularly important for dark energy properties.

Note that, as in the previous subsections, we have had to reduce the number of bins with which the two statistics are calculated when creating the combined contours, to avoid having a data vector with more bins than the number of realisations used for the covariance matrix calculation. For the combined contours in this figure, we use the void abundance calculated with 10 bins, and the tangential shear profiles calculated with 20 bins.

We show the percentage errors (at 68% and 95% CL) for the combination of the void abundance and tangential shear profiles, as well as those of the shear two-point correlation function ( $\gamma$ -2PCF), in Table 1. The grey contours in Fig. 7 show that the shear two-point correlation function provides similar constraints to those from the combination of the void abundance and tangential shear profiles, for all parameters at the 68% CL. Where constraints on  $\Omega_m$  and  $h$  are slightly better for the shear two-point correlation function. And constraints on  $w_0$  are tighter for the combined void abundance and tangential shear. However, for the 95% CL, constraints are better for the combined void abundance and tangential shear profiles for all parameters, relative to the shear two-point correlation function.

As shown by Table 1, the 68% CL constraints for the combined

void abundance with three  $\nu$  thresholds are tighter than those of the shear two-point correlation function for  $\Omega_m$ ,  $S_8$  and  $w_0$ . For the 95% CL constraints, the difference is even larger, where better constraints come from the combined void abundance for all parameters. The difference is most significant for  $w_0$ , where the cause of this behaviour is due to the tails in the contours for the shear two-point correlation function in the 2D planes which are most prominent for  $w_0$ . We expect that combining void abundance and  $\gamma$ -2PCF will further tighten the constraints, though this will be left for a future study.

Fig. 8 shows contours for the WL void abundance and tangential shear profiles combined, for the three catalogues  $\nu > 1$  (blue), 2 (orange) and 3 (green) respectively. The smallest contours are for the  $\nu > 1$  catalogue (which has been shown in green colour in Fig. 7) and increasing the  $\nu$  threshold increases the contour size, where the larger contours almost entirely enclose the smaller contours. This is because the number of voids decreases with  $\nu$  increasing, meaning that the statistical uncertainties increase and the constraining power is weakened.

Nevertheless, it is interesting to note that in Fig. 6 the tangential shear contours for the  $\nu > 3$  catalogue are large. The same is also true in Fig. 5 with the WL void abundance for the same catalogue. The resulting contour when the two statistics are combined however is significantly smaller, as shown by the green contour in Fig. 8. So even for this catalogue where individual constraints are poor, their combination is highly beneficial.

## 6 DISCUSSION AND CONCLUSIONS

In this paper we have tested the sensitivity of the WL void abundances and tangential shear profiles to four cosmological parameters:  $\Omega_m$ ,  $S_8$ ,  $h$  and  $w_0$ . To this end, we have trained a Gaussian Process emulator with 26 cosmologies sampled in this 4D parameter space using Latin hypercube, which can be used to predict these two void statistics for arbitrary cosmologies (within the range spanned by the training cosmologies). We have investigated the impact of changing the number of WL peaks used as tracers to identify voids, and ran Markov Chain Monte Carlo samplings from our mock weak lensing data to forecast the accuracies at which these four parameters can be constrained by a future, LSST-like, lensing survey, using different combinations of the above WL void statistics.

The results from Fig. 5 show that the WL void abundance combined over all catalogues gives the tightest parameter constraints, where the greatest sensitivity is to the  $S_8$  parameter. This is because the abundances of WL voids identified from WL peak catalogues at different  $\nu$  thresholds have different dependencies and degeneracy directions in the studied parameter space. We suspect that there is a close interlink between the void abundance and the peak two-point correlation function, but will defer a detailed study of the latter to a follow-up work. For now, we conclude that complementary information is contained in the abundances of voids from different WL peak catalogues, a fact that should be utilised in order to maximise the use and scientific return of future lensing data.

WL void tangential shear profiles, in contrast, provide less tight constraints on the same cosmological parameters, and the results from different peak catalogues do not seem to be complementary to each other. In particular, for low- $\nu$  peak catalogues such as  $\nu > 1$  (Fig. 6), there is little degeneracy between  $\Omega_m$  and  $S_8$ ; this is because  $S_8$  is designed to break the degeneracy between  $\Omega_m$  and  $\sigma_8$  for standard WL analysis, e.g., shear two-point correlation function, and the low- $\nu$  peaks have little bias with respect to the underlying

convergence field, so that their tangential shear profiles follow more closely the parameter dependency of the shear two-point correlation functions. WL void abundances, on the other hand, can have further degeneracies between  $\Omega_m$  and  $S_8$  (as seen in Fig. 5), indicating that they have different degeneracy directions between  $\Omega_m$  and  $\sigma_8$  compared with the shear two-point function, and therefore can lead to additional constraints to the latter.

Nevertheless, we highlight that the above conclusions only apply to the 4D parameters space that we have focused on in this work, and may change if additional physics is introduced into the cosmological model. In [Davies et al. \(2019b\)](#) we found that the tangential shear profiles are able to distinguish between modified gravity models with a larger signal-to-noise ratio than the void abundance. This suggests that there may be other cosmological parameters not studied here, such as those governing modified gravity laws, to which the tangential shear profile is more sensitive than the WL void abundance. We leave an exploration of this possibility to future works.

Finally, we have found that combining void abundance and tangential shear is another way to obtain tighter parameter constraints. Even for the  $\nu > 3$  catalogues, for which these two void statistics give poor individual constraints, significant improvement has been found with their synergy.

Overall, we find that weak lensing voids can be a promising cosmological probe to constrain models. The cosmological parameter to which the WL void statistics are most sensitive is  $S_8$ , which can be measured at the sub percent level (68% CL). We also find that  $\Omega_m$  can be measured to within  $\approx 4\%$ ,  $h$  to within  $\approx 5\%$  and  $w_0$  within  $\approx 7\%$  (all 68% CL).

As a comparison, we find that parameter constraints from the combined void abundances are comparable to those from the shear two-point correlation function (which were obtained from the same WL maps) at the 68% CL, however the constraints are tighter for the combined void abundance at the 95% CL. We note that we are unable to combine multiple catalogues without reducing the number of bins so the results presented here are upper limits. Further combinations such as combining the void abundance and tangential shear profiles over all catalogues, will show the maximum constraining power of WL voids, where it would also be possible to include further peak catalogues with higher or lower  $\nu$  thresholds. However, combining so many statistics into a single data vector requires a larger number of simulations than those currently available to us, so we leave such a study to a future project. We also note that constraints from the shear two-point correlation function can be improved by using tomography ([Martinet et al. 2020](#)), and it is therefore also important to test how tomography can improve the constraints from WL void statistics in the future.

Throughout this study, we have adopted a Gaussian smoothing of  $\theta_s = 1$  arcmin. It may also be interesting to study how the parameter constraints depend on the smoothing scale used to smooth the WL convergence maps. We know that using larger smoothing scales increases the size of the WL voids and reduces their total number ([Davies et al. 2020](#)). A larger number of WL map realisations will then be required in order to accurately measure WL void statistics for larger smoothing scales, so we leave such a study to future work. Nevertheless, we have performed a test by using a larger smoothing scale,  $\theta_s = 2$  arcmin, and in [Appendix B](#) we give a brief summary of the resulting parameter constraints. We can see that the results are similar to what we have found for a 1 arcmin smoothing, cf. [Fig. 7](#).

We have mentioned several potential follow-up works above, including a more in-depth study based on a larger number of simulations. Having a greater number of WL map realisations will also allow us to use larger data vectors in our likelihood analysis. This

would allow us to calculate likelihood contours for the tangential shear profiles combined over all catalogues with a greater number of bins, which may decrease the errors on the parameter estimates. This is because coarse binning of the statistics increases the contour size. It will further enable us to combine both the void abundances and their tangential shear profiles for all peak catalogues in a single analysis, which again we have not been able to do in this work.

It will also be important to develop an understanding of how the void function is affected by systematics including intrinsic alignments, baryonic feedback, and masking (which can bias statistics measured from convergence maps, e.g. see [Giblin et al. \(2018\)](#)), which we leave to future study.

In [Davies et al. \(2020\)](#) we studied the differences in WL void statistics between WL voids identified from different void finders. We found that the tunnel algorithm offered one of the best compromises between high signal-to-noise ratio and small impact from galaxy shape noise in the tangential shear profiles. However, it will also be interesting to assess the constraining power of WL voids identified using other void finders such as the watershed algorithm. The aim is to have a fully comprehensive study of the many different and unexplored ways to use future high-quality weak lensing data to maximise our ability to test cosmological models and constrain cosmological parameters.

## ACKNOWLEDGEMENTS

We thank Ludovic Van Waerbeke for the code used in the weak lensing convergence map construction.

CTD is funded by a UK Science and Technology Facilities Council (STFC) PhD studentship through grant ST/R504725/1. MC is supported by the EU Horizon 2020 research and innovation programme under a Marie Skłodowska-Curie grant agreement 794474 (DancingGalaxies). JHD acknowledges support from an STFC Ernest Rutherford Fellowship (project reference ST/S004858/1). BG acknowledges the support of the Royal Society through an Enhancement Award (RGF/EA/181006). YC acknowledges the support of the Royal Society through a University Research Fellowship and an Enhancement Award. BL is supported by an ERC Starting Grant, ERC-StG-PUNCA-716532, and additionally supported by the STFC Consolidated Grants [ST/P000541/1, ST/T000244/1].

Computations for the N-body simulations were enabled by Compute Ontario ([www.computeontario.ca](http://www.computeontario.ca)), Westgrid ([www.westgrid.ca](http://www.westgrid.ca)) and Compute Canada ([www.computecanada.ca](http://www.computecanada.ca)). The SLICS numerical simulations can be found at <http://slics.roe.ac.uk/>, while the cosmo-SLICS can be made available upon request. This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility ([www.dirac.ac.uk](http://www.dirac.ac.uk)). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

## REFERENCES

- Aihara H., et al., 2019, *PASJ*, **71**, 114
- Anderson T. W., 2003, *An introduction to multivariate statistical analysis*. Wiley-Interscience
- Asgari M., et al., 2020, arXiv e-prints, [p. arXiv:2007.15633](https://arxiv.org/abs/2007.15633)
- Bacon D. J., Refregier A. R., Ellis R. S., 2000, *Mon. Not. Roy. Astron. Soc.*, **318**, 625

- Baker T., Clampitt J., Jain B., Trodden M., 2018, *Phys. Rev. D*, 98, 023511
- Baker T., et al., 2019, preprint ([arXiv:1908.03430](https://arxiv.org/abs/1908.03430))
- Barreira A., Cautun M., Li B., Baugh C. M., Pascoli S., 2015, *J. Cosmo. Astropart. Phys.*, 2015, 028
- Barreira A., Bose S., Li B., Llinares C., 2017, *J. Cosmo. Astropart. Phys.*, 02, 031
- Bartelmann M., Schneider P., 2001, *Phys. Rept.*, 340, 291
- Biswas R., Alizadeh E., Wandelt B. D., 2010, *Phys. Rev. D*, 82, 023002
- Bos E. G. P., van de Weygaert R., Dolag K., Pettorino V., 2012, *Mon. Not. Roy. Astron. Soc.*, 426, 440
- Burger P., Schneider P., Demchenko V., Harnois-Déraps J., Heymans C., Hildebrandt H., Unruh S., 2020, *A&A*, 642, A161
- Cai Y.-C., Padilla N., Li B., 2015, *Mon. Not. Roy. Astron. Soc.*, 451, 1036
- Cautun M., Cai Y.-C., Frenk C. S., 2016, *Mon. Not. Roy. Astron. Soc.*, 457, 2540
- Cautun M., Paillas E., Cai Y.-C., Bose S., Armijo J., Li B., Padilla N., 2018, *Mon. Not. Roy. Astron. Soc.*, 476, 3195
- Colberg J. M., et al., 2008, *Mon. Not. Roy. Astron. Soc.*, 387, 933
- Coulton W. R., Liu J., McCarthy I. G., Osato K., 2019, preprint, ([arXiv:1910.04171](https://arxiv.org/abs/1910.04171))
- Davies C. T., Cautun M., Li B., 2018, *Mon. Not. Roy. Astron. Soc.*, 480, L101
- Davies C. T., Cautun M., Li B., 2019a, *Mon. Not. Roy. Astron. Soc.*, 488, 5833
- Davies C. T., Cautun M., Li B., 2019b, *Mon. Not. Roy. Astron. Soc.*, 490, 4907
- Davies C. T., Paillas E., Cautun M., Li B., 2020, arXiv e-prints, [p. arXiv:2004.11387](https://arxiv.org/abs/2004.11387)
- Dietrich J. P., Hartlap J., 2010, *Mon. Not. Roy. Astron. Soc.*, 402, 1049
- Falck B., Koyama K., Zhao G.-B., Cautun M., 2018, *Mon. Not. Roy. Astron. Soc.*, 475, 3262
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- Fry J. N., 1986, *Astrophys. J.*, 306, 358
- Fu L., et al., 2008, *A&A*, 479, 9
- Giblin B., et al., 2018, *Mon. Not. Roy. Astron. Soc.*, 480, 5529
- Gruen D., et al., 2015, *Mon. Not. Roy. Astron. Soc.*, 455, 3367
- Gruen D., et al., 2016, *Mon. Not. Roy. Astron. Soc.*, 455, 3367
- Habib S., Heitmann K., Higdon D., Nakhleh C., Williams B., 2007, *Phys. Rev. D*, 76, 083503
- Hamaus N., Pisani A., Sutter P. M., Lavaux G., Escoffier S., Wandelt B. D., Weller J., 2016, *Phys. Rev. Lett.*, 117, 091302
- Harnois-Déraps J., Giblin B., Joachimi B., 2019, *A&A*, 631, A160
- Hartlap J., Simon P., Schneider P., 2007, *A&A*, 464, 399
- Heydenreich S., Brück B., Harnois-Déraps J., 2020, arXiv e-prints, [p. arXiv:2007.13724](https://arxiv.org/abs/2007.13724)
- Heymans C., et al., 2012, *Mon. Not. Roy. Astron. Soc.*, 427, 146
- Hikage C., et al., 2019, *Pub. Astron. Soc. Japan*, 71
- Hildebrandt H., et al., 2017, *Mon. Not. Roy. Astron. Soc.*, 465, 1454
- Hoekstra H., et al., 2006, *Astrophys. J.*, 647, 116
- Jain B., Van Waerbeke L., 2000, *Astrophys. J.*, 530, L1
- Kaiser N., Wilson G., Luppino G. A., 2000, preprint, ([arXiv:astro-ph/0003338](https://arxiv.org/abs/astro-ph/0003338))
- Kilbinger M., 2015, *Rept. Prog. Phys.*, 78, 086901
- Kilbinger M., et al., 2013, *Mon. Not. Roy. Astron. Soc.*, 430, 2200
- Lavaux G., Wandelt B. D., 2012, *Astrophys. J.*, 754, 109
- Li Z., Liu J., Matilla J. M. Z., Coulton W. R., 2019, *Phys. Rev. D*, 99, 063527
- Liu J., Petri A., Haiman Z., Hui L., Kratochvil J. M., May M., 2015, *Phys. Rev. D*, 91, 063507
- Martinet N., Harnois-Déraps J., Jullo E., Schneider P., 2020, arXiv e-prints, [p. arXiv:2010.07376](https://arxiv.org/abs/2010.07376)
- Nadathur S., 2016, *Mon. Not. Roy. Astron. Soc.*, 461, 358
- Nadathur S., Carter P. M., Percival W. J., Winther H. A., Bautista J. E., 2019, *Phys. Rev. D*, 100, 023504
- Paillas E., Cautun M., Li B., Cai Y.-C., Padilla N., Armijo J., Bose S., 2019, *Mon. Not. Roy. Astron. Soc.*, 484, 1149
- Paz D., Lares M., Ceccarelli L., Padilla N., Lambas D. G., 2013, *Mon. Not. Roy. Astron. Soc.*, 436, 3480
- Pedregosa F., et al., 2011, *Journal of Machine Learning Research*, 12, 2825
- Pen U.-L., Zhang T., van Waerbeke L., Mellier Y., Zhang P., Dubinski J., 2003, *Astrophys. J.*, 592, 664
- Pisani A., Sutter P. M., Hamaus N., Alizadeh E., Biswas R., Wandelt B. D., Hirata C. M., 2015, *Phys. Rev. D*, 92, 083531
- Planck Collaboration et al., 2018, arXiv e-prints,
- Riess A. G., Casertano S., Yuan W., Macri L. M., Scolnic D., 2019, *Astrophys. J.*, 876, 85
- Sánchez C., et al., 2017, *Mon. Not. Roy. Astron. Soc.*, 465, 746
- Schneider P., van Waerbeke L., Kilbinger M., Mellier Y., 2002, *A&A*, 396, 1
- Schneider M. D., Knox L., Habib S., Heitmann K., Higdon D., Nakhleh C., 2008, *Phys. Rev. D*, 78, 063529
- Sellentin E., Heavens A. F., 2016, *Mon. Not. Roy. Astron. Soc.*, 456, L132
- Semboloni E., et al., 2006, *A&A*, 452, 51
- Sutter P. M., Lavaux G., Wandelt B. D., Weinberg D. H., Warren M. S., Pisani A., 2014, *Mon. Not. Roy. Astron. Soc.*, 442, 3127
- Troxel M. A., et al., 2018, *Phys. Rev. D*, 98, 043528
- Van Waerbeke L., et al., 2000, *A&A*, 358, 30
- Verde L., Treu T., Riess A. G., 2019, *Nature Astronomy*, 3, 891
- White S. D. M., 1979, *Mon. Not. Roy. Astron. Soc.*, 186, 145
- Wittman D. M., Tyson J. A., Kirkman D., Dell'Antonio I., Bernstein G., 2000, *Nature*, 405, 143

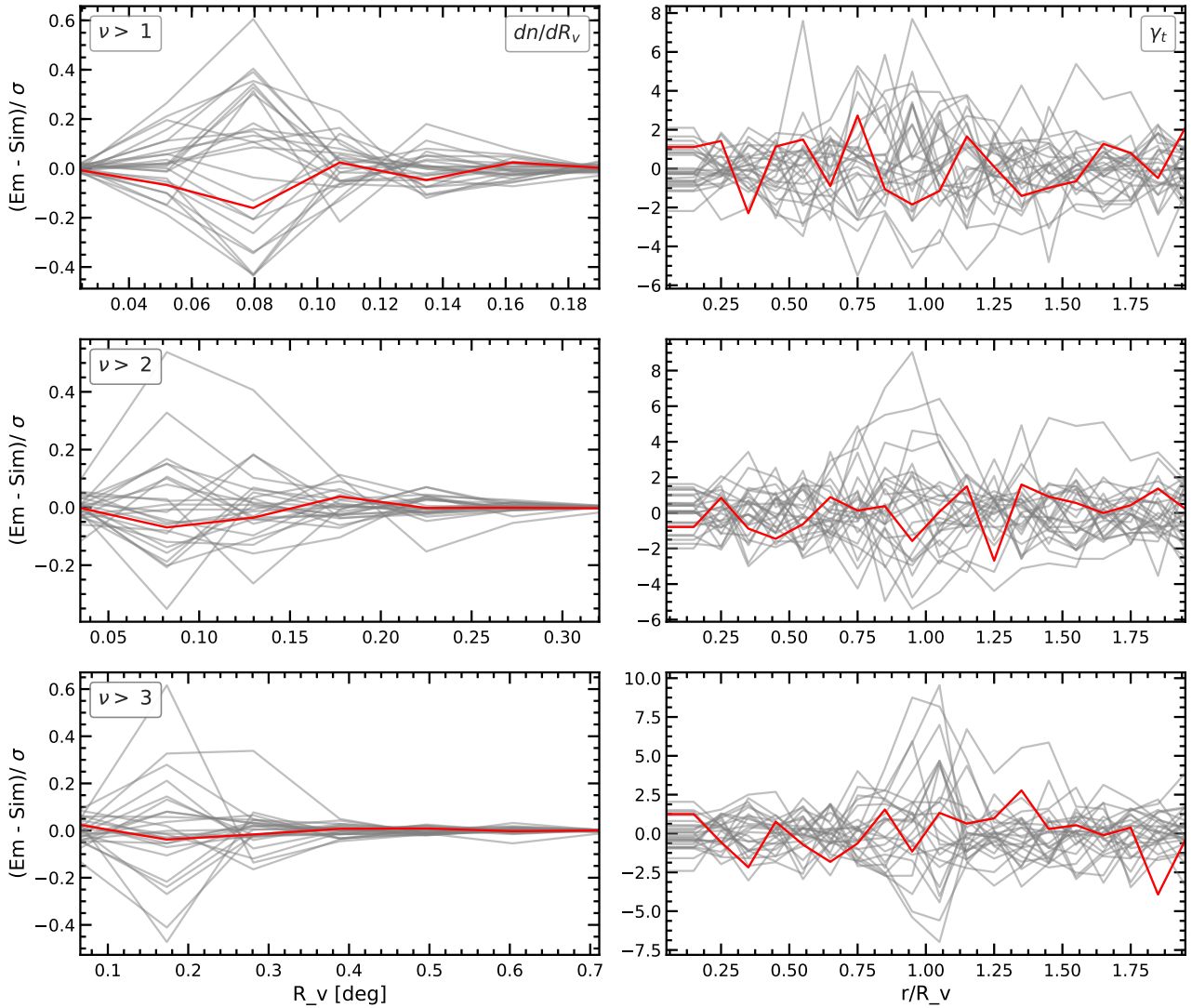
## APPENDIX A: ACCURACY OF THE EMULATOR

In order to test the accuracy of the GP emulator used to interpolate statistics between the cosmological parameter nodes in Fig. 1, we perform a cross validation test, which is outlined as follows. First, we remove one node from the training set of simulated data, and train the emulator with the remaining 25 cosmologies. An emulator prediction for the missing node is then calculated. The result is compared to the simulated version, by taking the difference between the two and dividing it by the standard error of the simulated data for that node. The above steps are then repeated 25 more times, by removing a different node from the training set at each iteration, which results in measurements of the emulator accuracy at each node. We note that the above procedure provides an upper limit for the emulator accuracy, since the emulator accuracy increases as more training data is used, and the cross validation measurements uses training data with one less node than the training data used in the main analysis.

Fig. A1 shows the cross validation test performed for the WL void abundance (left column) and the tangential shear profiles (right column). Results are shown for the catalogues with  $\nu > 1, 2$  and 3 in the top, middle and bottom rows respectively. The cross validation test at each node is plotted in grey, with the fiducial cosmology plotted in red. We highlight the fiducial cosmology because we use it as our mock observed data when generating likelihood contours. This makes it the most important region of the parameter space to emulate accurately.

The figure shows that the emulator accuracy does not vary greatly as a function of the  $\nu$  threshold. We find that the emulator is able to accurately predict the WL void abundance to within  $\approx 0.1\sigma$ , and within  $\approx 2\sigma$  for the tangential shear profiles

Regions towards the center of the 4D parameter space will be emulated more accurately than those at the boundary, since there is less training data for the GP emulator to train from at the edges of the parameter space. This is what creates the large spread amongst the grey curves in each panel, where curves towards the center of the 4D parameter space are more accurate, as shown by the fiducial cosmology. We are currently developing a suite of simulations to sample areas of the *cosmo*-SLICS parameter space more densely,



**Figure A1.** (Colour Online) The cross validation of the emulator accuracy. One node is removed from the training set and the emulation and simulation of the removed nodes are compared relative to its standard error. This is repeated for each of the 26 nodes, which gives an upper limit on the emulator accuracy. The left and right columns show results for the WL void abundance and tangential shear profiles respectively.

which will help to further improve the accuracy of the emulator by providing more training cosmologies that more densely sample the parameters space through Latin hypercubes or other node design schemes.

## APPENDIX B: THE IMPACT OF THE MAP SMOOTHING SCALE

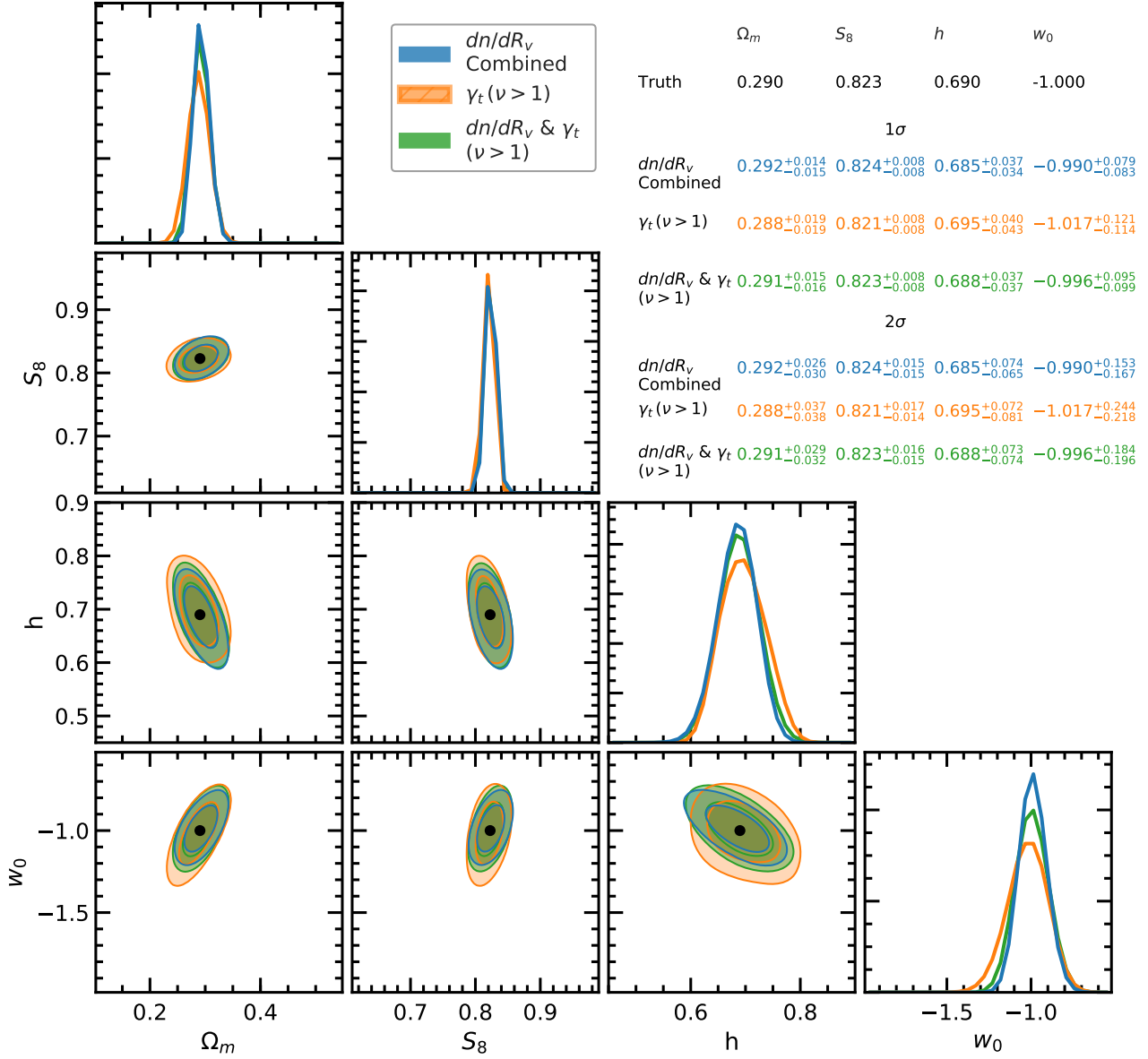
The analyses carried out in this work used smoothed WL convergence maps, which is required to suppress GSN. However this introduces an additional free parameter in the analysis – the smoothing scale applied to the maps, where we use a Gaussian smoothing of 1 arcmin in the main body of this work. In [Davies et al. \(2020\)](#) we studied how varying the smoothing scale impacts the resulting WL void statistics, and [Liu et al. \(2015\)](#) have shown that parameter constraints from WL peaks can be improved when multiple smoothing scales are combined. It is therefore useful to also show results for a different smoothing scale.

The likelihood contours for the statistics presented in Table 1

are shown in Fig. B1, but for a smoothing scale of 2 arcmin. These contours behave in a similar way to the case of 1 arcmin smoothing, with the tightest constraints coming from the combined WL void abundance. Overall these constraints are only slightly poorer than for the smaller smoothing scale.

It is possible to create constraints from combining multiple smoothing scales. However, due to the limited number of WL map realisations (which ultimately limits the length of our data vector to less than 50 bins), we leave this analysis to a future work.

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



**Figure B1.** (Colour Online) Likelihood contours for the statistics presented in Table 1, with WL void statistics identified in WL convergence maps smoothed over a 2 arcmin scale.