



<b>Publication Year</b>	2018
<b>Acceptance in OA@INAF</b>	2021-01-11T15:17:09Z
<b>Title</b>	KiDS-450: cosmological constraints from weak-lensing peak statistics - II: Inference from shear peaks using N-body simulations
<b>Authors</b>	Martinet, Nicolas; Schneider, Peter; Hildebrandt, Hendrik; Shan, HuanYuan; Asgari, Marika; et al.
<b>DOI</b>	10.1093/mnras/stx2793
<b>Handle</b>	<a href="http://hdl.handle.net/20.500.12386/29649">http://hdl.handle.net/20.500.12386/29649</a>
<b>Journal</b>	MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY
<b>Number</b>	474

# KiDS-450: cosmological constraints from weak-lensing peak statistics – II: Inference from shear peaks using $N$ -body simulations

Nicolas Martinet,<sup>1\*</sup> Peter Schneider,<sup>1</sup> Hendrik Hildebrandt,<sup>1</sup> HuanYuan Shan,<sup>1</sup> Marika Asgari,<sup>2</sup> Jörg P. Dietrich,<sup>3,4</sup> Joachim Harnois-Déraps,<sup>2</sup> Thomas Erben,<sup>1</sup> Aniello Grado,<sup>5</sup> Catherine Heymans,<sup>2</sup> Henk Hoekstra,<sup>6</sup> Dominik Klaes,<sup>1</sup> Konrad Kuijken,<sup>6</sup> Julian Merten<sup>7,8</sup> and Reiko Nakajima<sup>1</sup>

<sup>1</sup>Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany

<sup>2</sup>Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

<sup>3</sup>Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, D-81679 Munich, Germany

<sup>4</sup>Excellence Cluster Universe, Boltzmannstr. 2, D-85748 Garching, Germany

<sup>5</sup>Astronomico di Capodimonte, Via Moiariello 16, I-80131 Napoli, Italy

<sup>6</sup>Leiden Observatory, Leiden University, P.O. Box 9513, NL-2300RA Leiden, The Netherlands

<sup>7</sup>Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK

<sup>8</sup>INAF, Osservatorio Astronomico di Bologna, via Pietro Gobetti 93/3, I-40129 Bologna, Italy

Accepted 2017 October 24. Received 2017 October 23; in original form 2017 September 15

## ABSTRACT

We study the statistics of peaks in a weak-lensing reconstructed mass map of the first 450 deg<sup>2</sup> of the Kilo Degree Survey (KiDS-450). The map is computed with aperture masses directly applied to the shear field with an NFW-like compensated filter. We compare the peak statistics in the observations with that of simulations for various cosmologies to constrain the cosmological parameter  $S_8 = \sigma_8 \sqrt{\Omega_m/0.3}$ , which probes the  $(\Omega_m, \sigma_8)$  plane perpendicularly to its main degeneracy. We estimate  $S_8 = 0.750 \pm 0.059$ , using peaks in the signal-to-noise range  $0 \leq S/N \leq 4$ , and accounting for various systematics, such as multiplicative shear bias, mean redshift bias, baryon feedback, intrinsic alignment, and shear–position coupling. These constraints are  $\sim 25$  per cent tighter than the constraints from the high significance peaks alone ( $3 \leq S/N \leq 4$ ) which typically trace single-massive haloes. This demonstrates the gain of information from low-S/N peaks. However, we find that including  $S/N < 0$  peaks does not add further information. Our results are in good agreement with the tomographic shear two-point correlation function measurement in KiDS-450. Combining shear peaks with non-tomographic measurements of the shear two-point correlation functions yields a  $\sim 20$  per cent improvement in the uncertainty on  $S_8$  compared to the shear two-point correlation functions alone, highlighting the great potential of peaks as a cosmological probe.

**Key words:** gravitational lensing; weak – surveys – cosmological parameters – cosmology; observations.

## 1 INTRODUCTION

In a recent study, Hildebrandt et al. (2017) measured the coherent lensing distortions of galaxy images by large-scale structures (LSS) as a function of angular separation in the first 450 deg<sup>2</sup> of the Kilo Degree Survey (hereafter KiDS-450). This cosmic shear analysis yielded an  $S_8 (= \sigma_8 \sqrt{\Omega_m/0.3})$  value that is  $2.3\sigma$  lower than that inferred from *Planck* cosmic microwave background (CMB) measurements (Planck Collaboration XIII 2016). This difference

between low- and high-redshift probes, if it is not due to systematic effects or a statistical fluctuation, may point to new physics. To improve the constraints, we propose to use the statistics of peaks in the weak-lensing (WL) mass map of KiDS-450 in order to infer an additional lensing measurement of  $S_8$ , based on a different statistic than shear two-point correlation functions (hereafter 2PCFs).

The distribution of peak heights in mass maps depends on cosmology. In particular, peaks are sensitive to the matter density  $\Omega_m$  and the amplitude of the matter power spectrum described by  $\sigma_8$  on scales of  $8 h^{-1}$  Mpc, as these parameters impact the mass and the abundance of dark matter (DM) haloes. Peak statistics has been successfully used either to predict achievable cosmological constraints

\* E-mail: nmartinet@astro.uni-bonn.de

(e.g. Dietrich & Hartlap 2010; Kratochvil, Haiman & May 2010; Yang et al. 2011; Hilbert et al. 2012; Marian et al. 2012, 2013; Martinet et al. 2015) or to directly measure them from observations (e.g. Liu et al. 2015a,b; Kacprzak et al. 2016).

In contrast to classical second-order cosmic shear probes, shear peaks are sensitive to the non-Gaussianities in the matter and shear distributions. Commonly, while large peaks correspond to single massive haloes, the lower amplitude peaks are often due to the projection of multiple smaller haloes (Yang et al. 2011; Liu & Haiman 2016) and are also sensitive to cosmology (Jain & Van Waerbeke 2000; Wang, Haiman & May 2009; Dietrich & Hartlap 2010; Kratochvil, Haiman & May 2010). Low-amplitude peaks can also be produced by mass outside collapsed DM haloes or by shape noise contribution, and it is not clear yet which of these three different origins is dominant. Although second-order cosmic shear and peak statistics do not probe the exact same information, they are both sensitive to LSS, and their cosmological constraints are correlated (e.g. Dietrich & Hartlap 2010; Liu et al. 2015a). As these two methods use the same observables but different statistics, comparing their respective constraints is a good test for method-dependent systematics.

Peak statistics have been analysed with various methods. The main differences between studies arise from both measurement and modelling choices. From the measurement point of view, one can choose to reconstruct the WL map in convergence (e.g. Kratochvil, Haiman & May 2010; Yang et al. 2011; Shan et al. 2014; Liu et al. 2015a; Petri, May & Haiman 2016) or shear space through compensated filters (e.g. Kruse & Schneider 1999, 2000; Dietrich & Hartlap 2010; Maturi, Fedeli & Moscardini 2011; Hamana et al. 2012; Martinet et al. 2015; Kacprzak et al. 2016). The shear approach properly deals with the mass sheet degeneracy, which is only approximately handled in the convergence case. See Lin & Kilbinger (2017) for a recent comparison of the cosmological parameter estimates from peaks computed in shear and convergence spaces. Furthermore, working in shear space allows one to include the observational masks, at the cost of computational time, as it requires to drop the Fourier Transform approach.

The modelling of the peak distribution can be done with either simulations or analytical predictions.  $N$ -body simulations capture the non-linear regime of structure formation allowing the use of the full signal-to-noise (S/N) range of peaks. Although most studies rely on simulations, analytical predictions based on the halo mass function offer a promising way to speed up peak studies (Fan, Shan & Liu 2010; Lin & Kilbinger 2015; Shirasaki 2017). In particular, Zorrilla Matilla et al. (2016) showed that the model from Lin & Kilbinger (2015) predicts the mean abundance of high-S/N peaks reasonably well compared to  $N$ -body simulations, but that further development is needed to accurately estimate their variance or to probe the low-S/N tail. Both simulations and analytical predictions need to be adapted to the studied survey to capture the full complexity of the data.

In this paper, we apply aperture masses (Schneider 1996; Bartelmann & Schneider 2001) in shear space. We compare the peak distribution from the KiDS-450 data to the Dietrich & Hartlap (2010) simulations for various cosmologies and infer cosmological constraints on  $S_8$ . We also use mock data from the Scinet Light Cone Simulations (SLICS, Harnois-Déraps et al. 2015) to refine our covariance matrix and estimate the impact of sample variance. Measuring the mass maps for various filter scales, we assess the gain of information from a multiscale analysis. We compare our constraints on  $S_8$  to the KiDS tomographic cosmic shear results (Hildebrandt et al. 2017) in the context of the tension with *Planck*.

Finally, we measure the non-tomographic shear 2PCFs and present joint constraints for peaks and 2PCFs.

This paper is the second in a series of papers on peak statistics in KiDS-450. (Shan et al. 2017, hereafter **Paper I**) conducted an analogous analysis in convergence space, predicting the abundance of high-S/N peaks from an analytical model adapted from Fan et al. (2010). The use of simulations allows us to additionally probe the information contained in the low-S/N peaks, at the cost of only sparsely sampling the  $(\Omega_m, \sigma_8)$  cosmological plane. These two different approaches allow us to derive robust cosmological constraints from the peak statistics of the KiDS-450 survey and represent the largest observational WL peak statistic analyses to date.

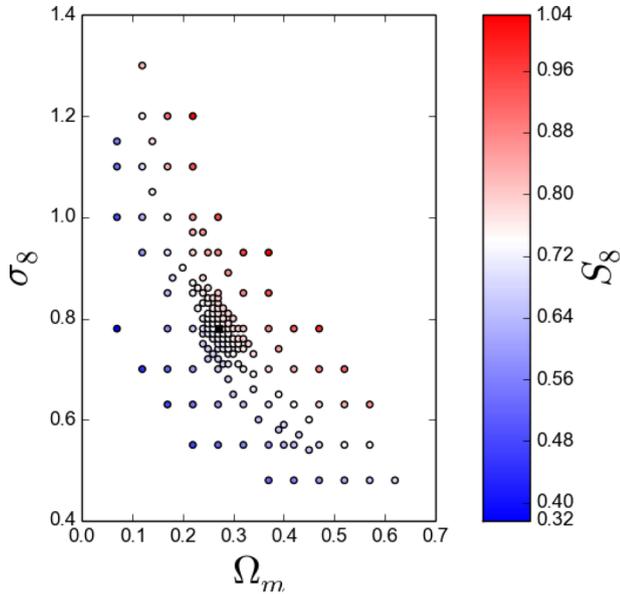
The paper is structured as follows. We describe our observations and simulations in Sections 2 and 3, respectively. We then explain our mass map reconstruction in Section 4 and present the KiDS peak distribution in Section 5. We estimate cosmological constraints in Section 6 and discuss them in Section 7.

## 2 OBSERVATIONS

This analysis is based on the KiDS-450 data release, presented in Hildebrandt et al. (2017) and de Jong et al. (2017), and therefore uses the same input galaxy catalogue. The KiDS survey is also documented in de Jong et al. (2015) and Kuijken et al. (2015) and we refer the reader to these papers for a complete description of the data set and the reduction pipelines. Nevertheless, we summarize the main aspects of the survey and the steps in the reduction that are relevant for this study.

KiDS is a ground-based survey optimized for WL measurements. The KiDS-450 sample is an intermediate release of the ongoing survey that covers  $449.7 \text{ deg}^2$ , split into five patches: three on the equatorial (G9, G12, and G15), and two in the southern sky (G23 and GS). Images are acquired with the OmegaCAM camera on the 2.6 m VLT Survey Telescope, in four optical bands ( $u$ ,  $g$ ,  $r$ , and  $i$ ). WL shape measurements are carried out on the  $r$ -band images which reach a limiting magnitude of 24.9 ( $5\sigma$  in a 2 arcsec aperture) and have a median seeing of 0.66 arcsec. Galaxy shapes are determined with the updated version of the model fitting algorithm *lensfit* (Miller et al. 2007), described in Fenech Conti et al. (2017). Photometric redshifts  $z_B$  are computed with the Bayesian code *BPZ* (Benítez 2000) using the four optical bands and are described in Kuijken et al. (2015). The redshift distribution is estimated from spectroscopically matched galaxies (Hildebrandt et al. 2017). We apply the same redshift cut as for the 2PCFs analysis:  $0.1 < z_B \leq 0.9$ , but do not split the data into different redshift bins. This choice is driven by limitations on the simulation side, and is explained in Section 3.2. The total number of galaxies in our catalogue is  $\sim 12.34$  million after the redshift cuts.

For any shape measurement method, one needs to calibrate the biases in the shear estimates. This is usually decomposed in multiplicative and additive terms in a linear relation between measured and true shear. The multiplicative bias of each galaxy is the same as in Hildebrandt et al. (2017), and is estimated through extensive simulations in Fenech Conti et al. (2017). As suggested in Miller et al. (2013), it is better to correct for multiplicative bias in a global approach to avoid possible correlation between ellipticities and correction factors. This correction is described in Section 4 and applied to each aperture mass in equation (8). We compute the mean additive shear bias as the average weighted ellipticity over all galaxies. The calculation is done independently for each of the five patches, and for each of the two ellipticity components. The values differ from those of Hildebrandt et al. (2017) because they determined it



**Figure 1.** Sampling of the  $(\Omega_m, \sigma_8)$  plane. Each dot represents an  $N$ -body simulation out of which five galaxy catalogues are made. Colours show  $S_8 = \sigma_8 \sqrt{\Omega_m/0.3}$  values. The central black dot corresponds to our fiducial cosmology  $\pi_0 = (0.27, 0.78)$ , which has 35  $N$ -body simulations and therefore 175 pseudo-independent catalogues.

independently for several redshift slices, while we use a single redshift bin. This bias is always lower than  $1.5 \times 10^{-3}$ , and is subtracted from the measured ellipticities.

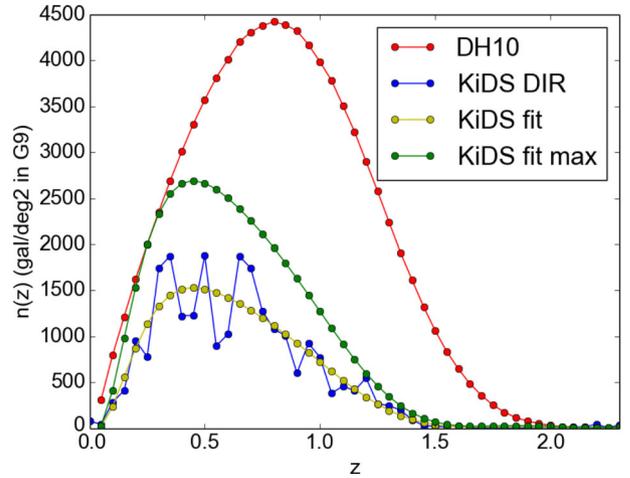
### 3 SIMULATIONS

We derive cosmological constraints by comparing the WL peak distribution of KiDS-450 to that of simulations with varying cosmologies. To that purpose, we use the simulations from Dietrich & Hartlap (2010). In Appendix C, we also use mock catalogues from the SLICS simulations (Harnois-Déraps et al. 2015) to better estimate the covariance matrix, and compare it with the covariance matrix from the Dietrich & Hartlap (2010) simulations that is used for parameter inference.

#### 3.1 Dietrich and Hartlap (2010) simulations

The Dietrich & Hartlap (2010) simulations consist of a set of 192  $N$ -body simulations run with the GADGET-2 software (Springel 2005), with initial conditions generated with the Eisenstein & Hu (1998) transfer function.  $256^3$  DM particles are evolved from  $z = 50$  to 0 in a box of  $200h_{70}^{-1}$  Mpc side length, with particle mass varying between  $9.3 \times 10^9 M_\odot \leq m_p \leq 8.2 \times 10^{10} M_\odot$ , depending on the cosmology. Each simulation spans a  $6 \times 6 \text{ deg}^2$  field of view.

These simulations are run with cosmological parameters  $\pi = (\Omega_m, \sigma_8)$ . Among them, 35 are run with fiducial cosmological parameters  $\pi_0 = (0.27, 0.78)$ , and 158 have  $\Omega_m$  and  $\sigma_8$  spanning a large range of values. One set of simulations was lost due to an archiving issue, and we therefore only use 157 different cosmologies. As seen in Fig. 1, the steps in the  $(\Omega_m, \sigma_8)$  plane are smaller around the fiducial parameters, allowing a better precision on the variation of the WL peak distribution around the cosmological parameter values expected from previous cosmological studies. We also show the variation of  $S_8 = \sigma_8 \sqrt{\Omega_m/0.3}$ , which is the parameter to which we are most sensitive, given the degeneracy between



**Figure 2.** Illustration on patch G9 of the downsampling of the Dietrich & Hartlap (2010) simulations to match KiDS redshift distribution. Red corresponds to the initial Dietrich & Hartlap (2010)  $n(z)$ , blue to the KiDS DIR  $n(z)$ , yellow to the smoothed KiDS DIR  $n(z)$ , and green to the downsampled Dietrich & Hartlap (2010) redshift distribution that matches that of KiDS.

$\Omega_m$  and  $\sigma_8$ . The other parameters that are not probed in this study are fixed to their fiducial values ( $\Omega_b = 0.04$ ,  $n_s = 1.0$ , and  $h_{70} = 1$ ), except  $\Omega_\Lambda$  which varies with  $\Omega_m$  to preserve flatness.

Ray tracing is then performed through each simulation to produce convergence and shear maps, from which a catalogue of galaxies is generated through random position sampling. Random shifting within a simulation snapshot was used to extract five pseudo-independent ray tracings out of a single  $N$ -body run. These mock catalogues mimic the Canada–France–Hawaii Telescope Legacy Survey (CFHTLS) in terms of redshift distribution, galaxy number density, and shape noise.

Further details on the Dietrich & Hartlap (2010) simulations and the creation of the mock catalogues can be found in the corresponding paper.

#### 3.2 Adapting to the KiDS survey

Because the Dietrich & Hartlap (2010) simulations are not tailored for KiDS-450 data, we need to modify their output. In particular, we want to use the same positions, redshift distribution, and shape noise as in the data.

The first step is to modify the redshift distribution of the simulations by subsampling the galaxies in order to match the KiDS redshift distribution. This is possible because the mocks have a much higher galaxy density than KiDS, i.e. 25 versus  $\sim 8.5$  galaxies per square arcminute. We use the DIR redshift distribution detailed in Hildebrandt et al. (2017) which corresponds to the redshift distribution of a magnitude-reweighted sample of spectroscopically matched galaxies in the photometric redshift range  $0.1 < z_B \leq 0.9$ . It was shown that this approach is more precise than using photometric redshifts, and this redshift distribution extends by construction above  $z_{\text{spec}} = 0.9$ . The KiDS galaxy density after applying this redshift cut is  $\sim 7.5$  galaxies per square arcminute. The process is illustrated in Fig. 2. We first fit the KiDS DIR redshift distribution with a polynomial of 12th order chosen to smooth the distribution. We check that this fit does not change the mean redshift of the distribution. However, the Dietrich & Hartlap (2010) mocks contain very few galaxies at  $z > 2$  due to the redshift distribution they adopted. Thus, we reject most galaxies selected in  $0.1 < z_B \leq 0.9$  with  $z_{\text{spec}} \geq 2$ .

This shifts the mean redshift by  $\sim 0.05$  towards a lower value. We then look for the largest multiplicative factor that can be applied to this smoothed distribution in order not to exceed the distribution in the simulation at any  $z$ . Taking the ratio of this last distribution (the green points in Fig. 2) to the  $n(z)$  of the simulations (red points of Fig. 2) gives a weight between 0 and 1 to each redshift bin. We finally downsample the simulation drawing for each galaxy a random number between 0 and 1 and discarding the galaxy if this number is above the weight of the galaxy redshift bin.

We then use a nearest-neighbour approach to assign a simulated reduced shear value at each of the observed positions. One could also use a linear interpolation of the four simulated galaxies closest to the observed one that we try to match. This technique would be more accurate if the simulated galaxies were placed on a grid. However, these galaxies are at random positions, and could lead to combination of shears from source galaxies that are not affected by the same lenses. We therefore assign the shear of the closest neighbour. For each simulation, we then build a catalogue of galaxies whose positions, weights, and intrinsic ellipticities are taken from the observed KiDS catalogue, and shears from the simulation. The KiDS-450 observational masks are also applied when assigning positions. The observed ellipticities are rotated by a random angle before being assigned to our simulated catalogue, allowing us to remove the signal from the observation but retaining its exact shape noise. The shear from the data slightly modifies the amplitude of the intrinsic ellipticity used in the simulations, but this effect is small as the shear amplitude is of the order of a few percents of the ellipticity. We also bias simulated values of the shear by the multiplicative bias measured for the corresponding observed galaxy, so that the bias is consistent between observations and simulations. However, ignoring this bias in the simulation affects the final cosmological constraints by less than  $0.01\sigma$ . Since the peak distribution is very sensitive to shape noise, we make several noise realizations by applying different random rotations to the observed ellipticities. This point is discussed in more detail in Section 5.1.

There are three caveats to this interpolation scheme. The first is that the KiDS data cover  $450 \text{ deg}^2$  while each simulation is only  $6 \times 6 \text{ deg}^2$ . We therefore have to use the same simulation several times to cover the entire observational field, which underestimates the sample variance. The effect of this procedure is studied in Appendix C, making use of the larger SLICS simulations (Harnois-Déraps et al. 2015). The second issue is that the galaxy density of the simulated mock catalogue is not large enough compared to that of the observations to ensure that no simulated shear value is used more than once in the interpolation process. As a consequence, some close galaxies in the matched catalogue will have the same shear (but different intrinsic ellipticities). However, this effect is mitigated by the fact that the separation between closest neighbours is much smaller than the scale of the filter that we are applying in the aperture mass calculation. Quantitatively, the mean separation between closest neighbours is  $0.145 \text{ arcmin}$  with a standard deviation of  $\pm 0.076 \text{ arcmin}$  and the filter's outer and effective radii are  $12.5$  and  $1.875 \text{ arcmin}$ , respectively. Even if a galaxy is attributed a shear from a slightly different position, this difference is not significant as seen by the filter function, leading to the same result as if the shear was estimated at the true galaxy position. This problem would become significant only if we were conducting a tomographic analysis, because the distance to the closest neighbour would become too large. A tomographic approach would thus require to directly build the mock catalogue at the desired positions and redshifts through looking up the values in the shear planes calculated at various redshifts in the simulations. Finally, the downsampling diminishes the

correlations between the lensing mass distribution and the source galaxy distribution, as the observed redshifts are randomly reshuffled. This effect is however accounted for as a systematic bias when evaluating the difference of S/N between simulated and observed peaks (Section 6.2).

The final simulation products consist of 175 catalogues at the fiducial cosmology and 785 catalogues at 157 different cosmologies. These catalogues have their shear values estimated from the Dietrich & Hartlap (2010) simulations, and their positions, weights, and intrinsic ellipticities from the observations. We note that the simulations do not include the full complexity of the observations. In particular, baryon feedback is not captured by these DM only simulations and the lens–source coupling is lost when assigning observed galaxy positions to the mocks. The impact of these effects on cosmological constraints is discussed in terms of systematics in Section 6.2.

### 3.3 Kids SLICS mocks

In Appendix C, we use the SLICS simulations to refine the covariance matrix and study the impact of sample variance on the cosmological constraints. In the rest of the paper, the mocks built from the Dietrich & Hartlap (2010) simulations are used. The SLICS simulations (Harnois-Déraps et al. 2015) consist of 930  $N$ -body simulations with  $1536^3$  particles evolved in a box of  $505 h^{-1} \text{ Mpc}$ , and cover  $10 \times 10 \text{ deg}^2$  in the redshift range  $0 < z < 3$ . Each particle has a mass of  $2.88 \times 10^9 M_{\odot} h^{-1}$ . Every simulation has the same cosmology:  $\Omega_m = 0.2905$ ,  $\Omega_{\Lambda} = 0.7095$ ,  $\Omega_b = 0.0473$ ,  $h = 0.6898$ ,  $\sigma_8 = 0.826$ , and  $n_s = 0.969$ , but different initial conditions.

As described in Hildebrandt et al. (2017), mock galaxy catalogues are drawn from these simulations, estimating the shear at various positions over 18 redshift planes. In addition to several improvements of the simulation quality compared to the Dietrich & Hartlap (2010) simulations, these mocks estimate the shear at the observed galaxy position without resorting to interpolation. This is also in contrast with the mocks used in Hildebrandt et al. (2017) where galaxies are at random positions. We have verified from the Dietrich & Hartlap (2010) simulations that using shear instead of reduced shear does not significantly affect the cosmological constraints derived from our peak estimator. We therefore use shear instead of reduced shear from the SLICS simulations, making the calculation faster.

From this set of simulations, we make 67 independent realizations of the KiDS-450 footprint, using different simulations to tile the space. This means that in contrast to the mocks we build from the Dietrich & Hartlap (2010) simulations which map the full  $450 \text{ deg}^2$  of data with  $36 \text{ deg}^2$  of simulations, these refined mocks better account for sample variance, as  $450 \text{ deg}^2$  of simulations are used to map the  $450 \text{ deg}^2$  of data. Details on the tiling will be available in a forthcoming paper (Harnois-Déraps et al., in preparation).

## 4 APERTURE MASS CALCULATION

Peaks are detected in a map of aperture masses (Schneider 1996; Bartelmann & Schneider 2001). This technique presents several advantages over the classical mass reconstruction from shear. In particular, it avoids the integration over finite area which introduces an unknown constant, due to the so-called mass sheet degeneracy. As the mass sheet degeneracy affects the signal but not the noise, its main effect is to add a random shift, different from field to field, to the S/N distribution of peaks. It also allows one to analytically compute local noise and to deal with masks in a simple fashion.

This led to its extensive use in WL peak analyses (e.g. Dietrich & Hartlap 2010; Marian et al. 2012; Martinet et al. 2015; Kacprzak et al. 2016). In Paper I, the mass map is reconstructed through a shear-convergence inversion (Kaiser & Squires 1993) because it is simpler to model the analytical prediction of peaks in convergence space. However, it is preferable to use the aperture mass statistics as we do in this second paper, to avoid mass sheet degeneracy, and to better handle the masks.

The aperture mass is an integral of the local mass density around position  $\theta_0$ , weighted by a filter function which is compensated in the convergence space:

$$M_{\text{ap}}(\theta_0) = \int d^2\theta U(\theta - \theta_0)\kappa(\theta), \quad (1)$$

where the compensation of the isotropic weight function  $U(\theta)$  is expressed as:

$$\int d\theta \theta U(\theta) = 0. \quad (2)$$

This condition ensures the aperture mass is insensitive to the (linear version of the) mass sheet degeneracy. For any compensated filter in convergence space  $U(\theta)$ , one can compute the equivalent filter  $Q(\theta)$  in shear space, which gives the aperture mass from the tangential shear (Schneider 1996):

$$Q(\theta) = \frac{2}{\theta^2} \int_0^\theta d\theta' \theta' U(\theta') - U(\theta), \quad (3)$$

$$M_{\text{ap}}(\theta_0) = \int d^2\theta Q(\theta - \theta_0)\gamma_t(\theta, \theta_0), \quad (4)$$

where the tangential shear  $\gamma_t(\theta, \theta_0)$  is expressed as a function of both shear components and the angle between the position where the shear is measured and the centre of the aperture  $\phi(\theta, \theta_0)$ :

$$\gamma_t(\theta, \theta_0) = -[\gamma_1(\theta) \cos(2\phi(\theta, \theta_0)) + \gamma_2(\theta) \sin(2\phi(\theta, \theta_0))]. \quad (5)$$

In order to apply aperture masses to observed data, the integration is transformed into a sum over discrete positions where the shear is estimated, i.e. at galaxy positions  $\theta_i$ . The tangential shear is also replaced by the galaxy tangential ellipticity:

$$M_{\text{ap}}(\theta_0) = \frac{1}{n_{\text{gal}}} \sum_i Q(\theta_i - \theta_0)\epsilon_t(\theta_i, \theta_0), \quad (6)$$

where  $n_{\text{gal}}$  is the galaxy density inside the aperture. The masks are easily handled as long as the computation is done in real space, as the masked galaxies can simply be ignored in the computation. However, it significantly increases the computational time compared to Fourier space. We prioritize the exact handling of masks and therefore do the calculation in real space.

Galaxy ellipticity is equal to the reduced shear on average, provided that source galaxies are randomly oriented. This property is the fundamental hypothesis of WL and allows us to replace shear by ellipticity in equation (6), also enabling the analytic computation of the local noise as the standard deviation of the aperture mass in the absence of shear:

$$\sigma(M_{\text{ap}}(\theta_0)) = \frac{1}{\sqrt{2}n_{\text{gal}}} \left( \sum_i |\epsilon(\theta_i)|^2 Q^2(\theta_i - \theta_0) \right)^{1/2}. \quad (7)$$

The sum over the squared ellipticity norm  $(|\epsilon(\theta_i)| = \sqrt{\epsilon_1(\theta_i)^2 + \epsilon_2(\theta_i)^2})$  is sometimes replaced by the 2D dispersion of the ellipticity over the whole survey, and denoted by  $\sigma_\epsilon$ . However, it is more accurate to compute the shape noise at the level of each aperture as it varies from field to field, either for

instrumental or physical reasons, e.g. varying depth, point spread function variations or intrinsic alignments. We define the S/N of each aperture as the ratio of  $M_{\text{ap}}$  and  $\sigma(M_{\text{ap}})$ . Taking *lensfit* shear weights  $w$  into account we can write this S/N as:

$$\frac{S}{N}(\theta_0) = \frac{\sqrt{2} \sum_i Q(\theta_i - \theta_0)w(\theta_i)\epsilon_t(\theta_i, \theta_0)}{\sqrt{\sum_i w(\theta_i)^2 |\epsilon(\theta_i)|^2 Q^2(\theta_i - \theta_0)}} \frac{\sum_i w(\theta_i)}{\sum_i w(\theta_i) [1 + m(\theta_i)]}. \quad (8)$$

As already stated in Section 2, the shear multiplicative bias correction is applied as the average weighted correction over every galaxy multiplicative bias  $m(\theta_i)$  in the aperture. The correction appears as a normalization to  $M_{\text{ap}}(\theta_0)$ :  $\sum_i w(\theta_i) [1 + m(\theta_i)]$ , but does not apply to  $\sigma(M_{\text{ap}}(\theta_0))$  which is only normalized by the sum over the galaxy weights:  $\sum_i w(\theta_i)$ . This is because the multiplicative bias is computed as a shear correction and the aperture mass noise is only sensitive to the intrinsic ellipticities.

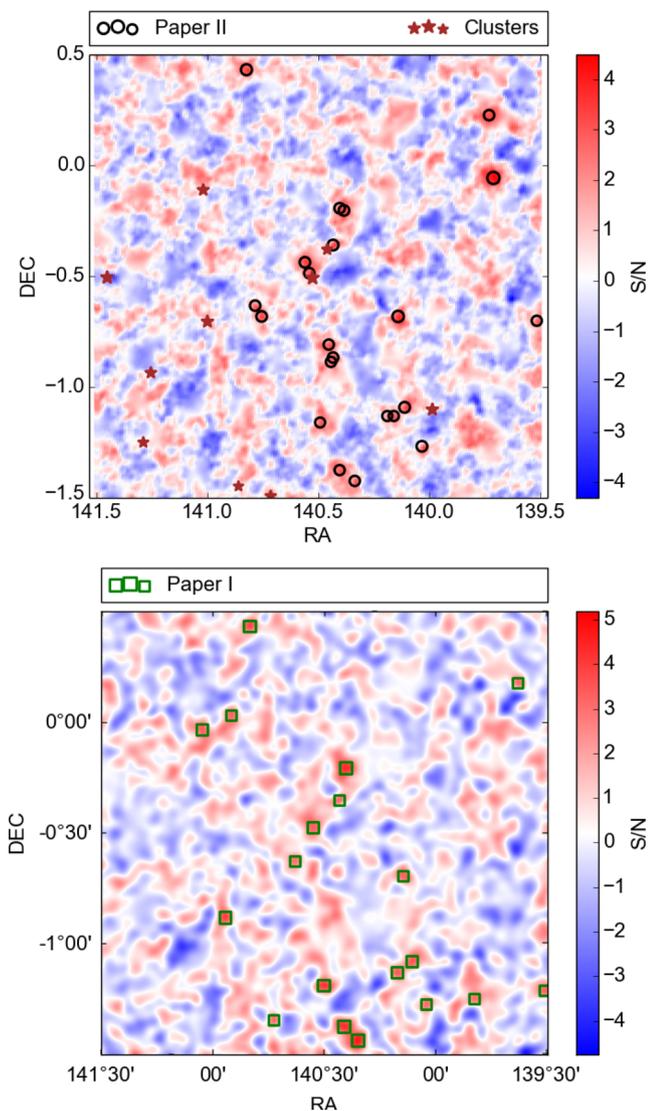
As seen in the equations the aperture mass depends on the filter function  $Q(\theta)$ . As we want to capture the signal from DM haloes, we choose a shape that matches the expected tangential shear signal of a typical halo. While an NFW profile (Navarro, Frenk & White 1997) would work well, we prefer to use an approximation of this profile to speed up the computation, namely the Schirmer et al. (2007) filter function:

$$Q(\theta) = \left[ 1 + \exp\left(6 - 150\frac{\theta}{\theta_{\text{ap}}}\right) + \exp\left(-47 + 50\frac{\theta}{\theta_{\text{ap}}}\right) \right]^{-1} \times \left(\frac{\theta}{x_c\theta_{\text{ap}}}\right)^{-1} \tanh\left(\frac{\theta}{x_c\theta_{\text{ap}}}\right), \quad (9)$$

where  $\theta_{\text{ap}}$  is the radius of the aperture, and  $x_c$  is analogous to the halo concentration in the NFW profile, and is set to  $x_c = 0.15$ , found to be the optimal value for galaxy cluster detection (Hettterscheidt et al. 2005). The first term corresponds to an exponential cut-off at  $\theta \rightarrow 0$  and  $\theta \rightarrow \infty$ . The cut-off at  $\theta \rightarrow 0$  is particularly important to avoid assigning too much weight to galaxies close to the aperture centre where reduced shear values may not be in the WL regime. The size of the filter is also important as it can preferentially select smaller or larger haloes. In this study, we set the fiducial aperture radius to  $\theta_{\text{ap}} = 12.5$  arcmin, which maximizes the number of peaks at  $S/N \geq 3$  in the KiDS data. With the chosen  $x_c$  parameter, this size corresponds to an effective radius  $x_c\theta_{\text{ap}}$  of 1.875 arcmin. This aperture size gives the maximal sensitivity to massive haloes. In Section 7.4, we compute the peak distribution for different filter sizes and discuss correlations between scales and the potential gain of information from a multiscale analysis.

We compute the aperture mass on a grid which covers the KiDS-450 area with a pixel size of 0.59 arcmin. This pixel size is a good trade-off between computational time and accuracy, as decreasing the pixel size further does not lead to the appearance of smaller structures. We discard all pixels closer to the edges of the reconstructed map than the aperture radius to avoid including incomplete apertures. However, we note that these cuts do not significantly affect the cosmological parameter estimates as they are also applied to the simulations which have the same galaxy positions and masks. Maps are made independently for each patch: G9, G12, G15, G23, and GS (see Hildebrandt et al. 2017). Due to the incomplete current tiling of the survey, we also subdivide G12, G15, and GS in 4, 3, and 2 subpatches respectively, to avoid unnecessary computation in empty areas.

As an example of our reconstruction method, we show an aperture mass map for a  $2 \times 2$  deg<sup>2</sup> field in the KiDS-450 footprint in Fig. 3 and the detected peaks with  $S/N \geq 3$ , defined as pixels with greater



**Figure 3.** Top: aperture mass map of a  $2 \times 2 \text{ deg}^2$  field in the KiDS-450 footprint. The S/N of each pixel is colour-coded from blue to red, red corresponding to high- $M_{\text{ap}}$  pixels. Black circles represent peaks in this map with  $S/N \geq 3$ , and brown stars indicate galaxy clusters from Radovich et al. (2017) with redshift  $z \leq 0.5$  and detection level  $\sigma \geq 7$ . Bottom: mass map of the same field computed from the direct shear inversion in Paper I. Green squares correspond to peaks with  $S/N \geq 3$  as detected in Paper I.

S/N than their eight neighbours. For comparison, we also display the mass map and peaks with  $S/N \geq 3$  from Paper I for the same field. This second mass map is computed from a shear inversion method, with a single noise value across the survey. We see that the two maps trace the same LSS, but present slight differences on small scales. There is in particular a higher amount of substructures in the aperture mass map compared to the shear-inverted convergence map. This is probably due to the choice of the smoothing filter, which is an NFW-like 1.875 arcmin filter for aperture mass and a Gaussian 2 arcmin filter in the shear-inversion method. We also find that the peaks from each method do not all overlap due to the differences in the map computation and in the definition of the noise. Although the peak distributions from Paper I and this study are different, the cosmological constraints should be comparable as the modelled peak distributions are computed in a consistent way with the observed distribution for each study.

We also compare our aperture mass map with known galaxy clusters overlapping with the KiDS-450 area (Radovich et al. 2017). These clusters have been detected through a matched filtering technique taking into account the magnitude distribution and density profile. We only retain clusters that are at redshift  $z \leq 0.5$  because higher redshift clusters are unlikely to create a strong shear signal given the mean redshift of the background source population. We also cut out clusters that are detected with less than  $7\sigma$  significance to have a very pure sample. We see from Fig. 3 that there is not a one-to-one correspondence between peaks and clusters. Only a few clusters are associated with peaks, but most clusters coincide with a high-S/N area of the WL mass map. This highlights that even at  $S/N \geq 3$  many peaks are not associated with clusters and contain a significant contribution from projection of low-mass haloes or shape noise contamination. We also note that adding less significant clusters does not qualitatively change these conclusions. Finally, we recall that even if the aperture mass is computed with an NFW filter to match haloes, our method is not optimized to cluster detection. In particular, we are sensitive to the integrated contribution along the line of sight which dilutes the signal from galaxy clusters.

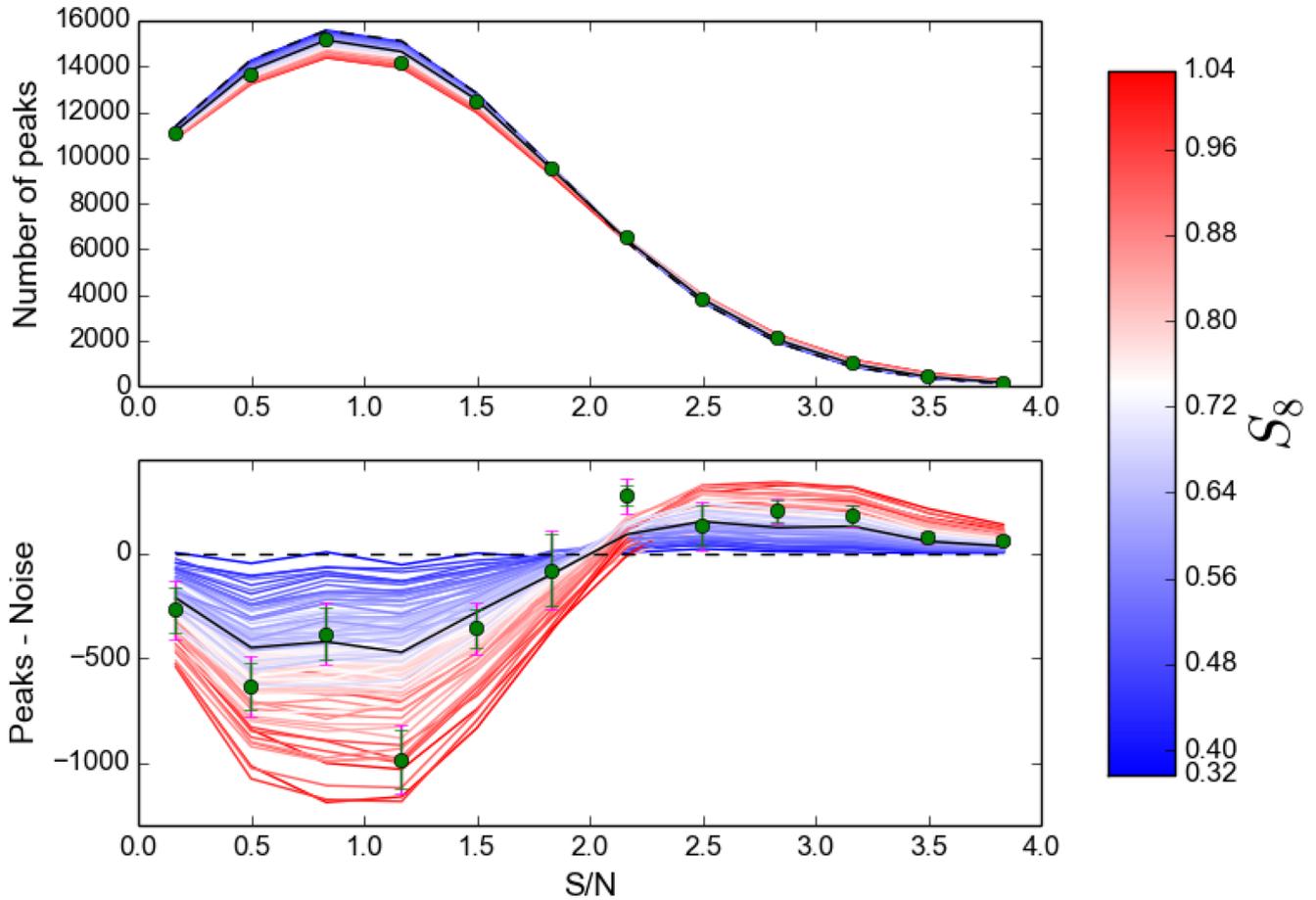
## 5 PEAK DISTRIBUTION

### 5.1 Measurement

Peaks are identified as pixels with S/N higher than their eight neighbours in the aperture mass map, with the pixel scale of 0.59 arcmin. The global strategy is to measure the peak S/N distribution from the observations and the variation of the peak distribution with cosmology from the simulations.

Because we reproduce the same noise in the simulations as that of the observations, we can safely use any part of the peak distribution, including the low-S/N tail. However, the width of the S/N bins and the upper limit of the distribution must be chosen such to ensure that the distribution can be modelled by a multivariate Gaussian when computing the likelihood, i.e. that there is a sufficient number of peaks per bin. We note that this problem can also be dealt with by using the cumulative distribution (e.g. Dietrich & Hartlap 2010) or a varying width to get the same number of peaks per bin (e.g. Martinet et al. 2015). However, we use bins with fixed width because these other two methods would favour the more numerous low-S/N peaks given the chosen range of S/N. The number of bins is also limited by the precision we want to achieve on the covariance matrix. As shown in Taylor & Joachimi (2014), the more degrees of freedom the larger the uncertainty in the covariance. We use 12 bins of S/N equally spaced between 0 and 4, but also try a few other configurations (8 and 16 bins) to ensure that our constraints are insensitive to the bin width for reasonable choices. We refrain from adding peaks with  $S/N \geq 4$  as for these peaks the shear–position coupling becomes significant and can bias the results (Kacprzak et al. 2016). Shear–position coupling, also referred to as boost factor, biases the heights of peaks corresponding to large haloes in the simulations compared to the observations because the redshift distribution of the data is applied to the simulations without prior knowledge of halo positions. This is described in more details in Section 6.2, where we analyse the different systematic biases.

The error bars on the number of peaks displayed in the various figures correspond to the diagonal elements of the covariance matrix estimated from the fiducial cosmology mocks, based on the Dietrich & Hartlap (2010) simulations. For the cosmological analysis, the full covariance matrix is used. We also verify that these error bars are comparable to those computed through bootstrap



**Figure 4.** Peak distribution (top) and differential peak distribution (bottom). The differential distribution corresponds to the peak distribution from which the averaged distribution over five noise-only field is subtracted. Green dots represent KiDS-450 data with error bars from the diagonal elements of the covariance matrix, the black line represents the mean of the fiducial cosmologies, black dashed line the noise only distribution, and coloured lines the various simulations with  $S_8$  increasing from blue to red. Error bars from bootstrap resampling of the data are displayed in magenta.

resampling of the data. To estimate the bootstrap variances, we divide the survey into 50 subpatches with equal number of galaxies. Similarly to Hildebrandt et al. (2017), the definition of the subpatches is based on right ascension cuts as the width in declination is roughly the same at any right ascension in the survey. This division leads to 50 patches which are roughly  $3 \times 3 \text{ deg}^2$ . We then select 50 random patches, with the possibility of selecting the same patch more than once, to create a new peak distribution. Doing so 10 000 times and calculating the dispersion of the peak distribution over them allows us to derive error bars that takes into account sample variance. These error bars are in very good agreement with those of the covariance matrix, as can be seen from Fig. 4 where bootstrap errors are represented in magenta and those from the simulations in green, highlighting that the simulations are a good representation of the data.

Because the peak distribution is dominated by noise, we need to run several realizations of the observed shape noise so that the simulations are not biased to one particular realization of shape noise. For every simulation, we run five random noise realizations. We also build five random noise-only peak distributions from the observations by computing the aperture mass map with all galaxies being randomly rotated. Each of these five realizations is computed with a different random seed, but the seed is the same for all different cosmologies and for the noise-only realization, limiting the impact of random shot noise. The simulations at the fiducial cosmology

all have different random seeds because they are used to estimate the covariance matrix. This allows us to measure differential peak counts, i.e. the peak distribution in the aperture mass map from which we subtract the distribution of noise peaks. We verified that increasing the number of realizations to 20 does not affect the cosmological constraints (less than  $0.3\sigma$  change).

Fig. 4 shows the main results concerning the peak distribution. It displays the peak distribution, the noise distribution, and the differential distribution for the observation and for all simulations. Simulated peak distributions are the mean over the different noise realizations. We see in particular that the peak distribution is dominated by shape noise, but that we can control it by having the same noise in the data and the simulations. Looking at the differential peak distribution, we see a good agreement between the data and the simulations with  $S_8$  slightly higher than the fiducial cosmology. We note also that the simulated peak distributions vary smoothly with cosmology with an increasing number of high-S/N peaks when  $S_8$  is higher. This is expected: an increase in  $\Omega_m$  increases the mass content of the universe and an increase in  $\sigma_8$  increases the clustering of structures, which both lead to more massive haloes and therefore more high-S/N peaks. In the low-S/N regime, we note that the differential peak distribution gets negative. This is because the aperture mass distribution is a convolution between signal which presents a high-S/N tail and random Gaussian noise. The noise acts as a Gaussian smoothing and lowers the amplitude in the convex

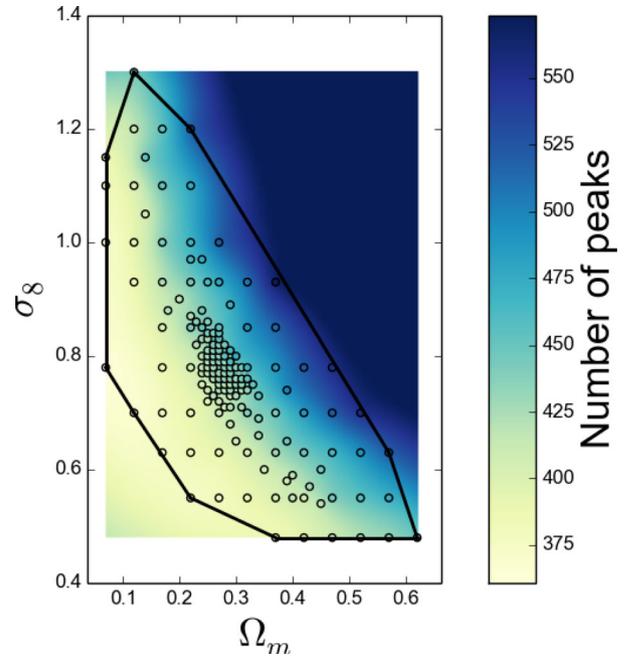
parts of the distribution, while it increases in the concave parts. The peak distribution is thus of higher amplitude than the noise-only distribution at high S/N, of lower amplitude at low S/N, and of equal amplitude when the second derivative of the peak distribution is equal to zero, here around  $S/N = 2$ . This argument is rigorously true only for the aperture mass distribution, but we extend it to the peaks as they are a subsample of the latter distribution. Finally, we note that the observed differential number of peaks deviates by more than  $2\sigma$  from the expectation at  $S/N = 1.33$  and  $2.33$ . Dividing the data in several subareas, we find that only some of the patches are affected by these offsets, but could not find an obvious cause to them. This plot shows that the peak distribution is sensitive to cosmology and that the combination of the KiDS observations and of the simulations we are using does enable us to constrain the cosmological parameter  $S_8$ . Although it would be tempting to use the differential peak counts for extracting cosmology, we prefer to work with the non-subtracted peak distribution to avoid biasing the data which contain only one realization of the noise.

## 5.2 Interpolation

Due to the prohibitive computational cost of simulating, a cosmological grid evenly sampling the  $(\Omega_m, \sigma_8)$  plane, we interpolate the peak distribution at the grid values. We cover a regular grid with step size 0.01 in each direction. Each bin of the data vector is interpolated separately. Peak distributions are averaged over the different ray-tracing and noise realizations before performing the interpolation so that we are not biased by a particular noise realization. We also recall that shot noise is reduced by applying the same random shape noise to all cosmologies (but the fiducial) for each noise realization. We note that it is also possible to directly interpolate the likelihood instead of the peak distribution, but the former method is preferable as it interpolates the expected values of the peak distribution, while the latter also affects the data vector which enters the likelihood.

We use radial basis functions with a multiquadric model which renders well the evolution of the number of peaks with  $\Omega_m$  and  $\sigma_8$  (Liu et al. 2015a). The computation is performed through the `scipy.interpolate.Rbf` PYTHON function (<https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.Rbf.html>).

Because the variation of the peak distribution with cosmological parameters is noisy, we also add some smoothing when interpolating the peak distribution through the ‘smooth’ argument of the `scipy.interpolate.Rbf` function which reduces the number of nodal points in the interpolation process. This improves the rendering of the significance contours in the  $(\Omega_m, \sigma_8)$  plane and we verify that it does not affect the estimated value of  $S_8$ . We also check that the error on the interpolated number of peaks is lower than the Poisson error by comparing the results of the interpolation with the measurements for every available simulation (see Appendix A for details). In principle, we could avoid the smoothing by running simulations for more points in the  $(\Omega_m, \sigma_8)$  plane, but this would be computationally demanding, and unnecessary as we found that the constraints on  $S_8$  do not change for various values of smoothing. This would however improve the cosmological contours in 2D space. Zorrilla Matilla et al. (2016) also found that using a refined grid of  $\sim 8000$  points in the  $(\Omega_m, \sigma_8)$  plane instead of  $\sim 160$  impacts the tails of the 2D contours but not seriously so in the direction perpendicular to the main degeneracy, so that the undersampling does not significantly affect the estimate of  $S_8$ .



**Figure 5.** Interpolation of the number of peaks in the bin  $3.33 \leq S/N \leq 3.66$ . Dots correspond to measured values and the background area to interpolated ones. The black polygon represents the convex hull within which we trust the interpolation. See the text for details.

An example of the interpolated number of peaks in the  $(\Omega_m, \sigma_8)$  space is given in Fig. 5. We see that the interpolation performs reasonably well comparing interpolated values to the nearby measured data points (see Appendix A for the quantitative comparison). However, the extrapolation is very inaccurate. We therefore apply a prior on the likelihood to discard the extrapolation region. This region is defined through a convex hull on the ensemble of points where simulations were run, and is displayed in Fig. 5.

## 6 COSMOLOGICAL CONSTRAINTS

### 6.1 Inferring cosmological parameters

Cosmological parameters are estimated by comparing the observed peak distribution to that of simulations with various cosmologies, in a Bayesian framework.

Our data vector is represented by  $\mathbf{x} \in \mathbb{N}_+^{N_b}$ , the number of peaks in each of the  $N_b$  bins of S/N shown in Fig. 4. Similarly, we define the peak distribution of a simulation with cosmology  $\boldsymbol{\pi} = (\Omega_m, \sigma_8)$  as  $\mathbf{x}_s(\boldsymbol{\pi}) \in \mathbb{R}_+^{N_b}$ .

From Bayes theorem, we can link the probability  $p(\boldsymbol{\pi}|\mathbf{x})$  of one cosmological model given the data vector (i.e. what we want to know) to the probability  $p(\mathbf{x}|\boldsymbol{\pi})$  of the data vector given a cosmology;

$$p(\boldsymbol{\pi}|\mathbf{x}) = \frac{p(\mathbf{x}|\boldsymbol{\pi})p(\boldsymbol{\pi})}{p(\mathbf{x})}. \quad (10)$$

The probability of the data  $p(\mathbf{x})$  is a normalization constant and  $p(\boldsymbol{\pi})$  is a flat prior, with value 1 on the probed range of cosmologies  $\boldsymbol{\pi}$  within the convex hull shown in Fig. 5, and 0 elsewhere.

The peak distribution is assumed to be a multivariate Gaussian distribution. This approximation is valid provided that we have a sufficient number of peaks in each bin, typically a few tens. The main analysis is done with 12 bins evenly spaced between S/N of 0 and 4

but we also check the robustness of our results over two alternative bin widths (0.25 and 0.5). We do not use a Gaussian likelihood but the adapted version of a multivariate  $t$ -distribution presented in Sellentin & Heavens (2016), which still assumes Gaussian distributed data. This likelihood, derived from marginalizing over the true covariance matrix, provides better inference than the traditional Gaussian likelihood with the Hartlap, Simon & Schneider (2007) correction, which only gives an unbiased estimate for the inverse covariance matrix. The likelihood can be written as

$$p(\mathbf{x}|\boldsymbol{\pi}) = c(N_s, N_b) \sqrt{\det \Sigma(\boldsymbol{\pi})} \left[ 1 + \frac{\chi^2(\mathbf{x}, \boldsymbol{\pi})}{N_s - 1} \right]^{-N_s/2}, \quad (11)$$

where  $N_s$  is the number of simulations used to estimate the covariance matrix  $\Sigma$ ,  $c(N_s, N_b)$  is a constant which depends on the number of simulations and the size of the data vector, and  $\chi^2(\mathbf{x}, \boldsymbol{\pi})$  is the  $\chi^2$  function defined in equation (14). We note that this likelihood approaches a Gaussian likelihood when the number of simulations  $N_s$  is large.

With the assumption that the covariance matrix does not depend on  $\boldsymbol{\pi}$ , the numerator of equation (11) is constant and we can write

$$p(\mathbf{x}|\boldsymbol{\pi}) \propto \left[ 1 + \frac{\chi^2}{N_s - 1} \right]^{-N_s/2}, \quad (12)$$

where the covariance matrix is computed from the  $N_s = 175$  simulations of the fiducial cosmology  $\boldsymbol{\pi}_0$ ,

$$\Sigma(\boldsymbol{\pi}_0) = \frac{1}{N_s - 1} \sum_i^{N_s} (\mathbf{x}_{s,i}(\boldsymbol{\pi}_0) - \bar{\mathbf{x}}_s(\boldsymbol{\pi}_0)) (\mathbf{x}_{s,i}(\boldsymbol{\pi}_0) - \bar{\mathbf{x}}_s(\boldsymbol{\pi}_0))^T. \quad (13)$$

The vector  $\mathbf{x}_{s,i}(\boldsymbol{\pi}_0)$  represents the peak distribution of the  $i$ th fiducial simulation and  $\bar{\mathbf{x}}_s(\boldsymbol{\pi}_0)$  is the mean peak distribution over all fiducial simulations.

Because we do not have the computational resources to compute the covariance matrix at each cosmology, we make the assumption that it does not depend on cosmology. Although this approximation does not hold for large variations in the cosmological parameters, Eifler, Schneider & Hartlap (2009) showed that it overestimates the errors on cosmological parameters in the case of second-order cosmic shear (which contains overlapping information with peaks) such that our constraints are conservative. We also note that Zorrilla Matilla et al. (2016) found a  $\sim 15$  per cent improvement in the cosmological parameter forecasts from peaks in simulations when taking into account the cosmological dependence of the covariance matrix, such that the constraints presented in this paper could be further improved with extra simulations at the non-fiducial cosmologies.

The covariance matrix estimates the error correlations in the data. The main sources of errors are galaxy shape noise and sample variance. The first one is probed by applying different random orientations to the intrinsic ellipticities of galaxies, and the second one by using several simulations and several ray tracings through the simulations. Kacprzak et al. (2016) focus on the shape noise contribution by applying many realizations of shape noise to the same simulations. This approach allows them to have a higher number of data vectors in the covariance matrix computation but neglects the contribution from sample variance over that of shape noise. In contrast, we estimate our covariance matrix with  $N_s = 175$  independent data vectors from 35 different simulations with five different ray tracing each, and different shape noise realizations. We compute five covariance matrices with different seeds for shape noise and average the covariances. We use this approach because we find that shape noise and cosmic variance affect the peak distribution at the

same level. The peak distribution of 10 fiducial different simulations with the same shape noise and that of one fiducial simulation with 10 different realizations of the shape noise represents a dispersion of the same order, typically a few to ten percent of the mean value. With this strategy, we estimate an accurate covariance matrix without biasing with non-independent data vectors. The cosmological constraints are almost identical for any individual matrix, but using the average covariance avoids choosing one set of noise realizations over another.

The  $\chi^2$  is defined in equation (14) from comparing the observed data vector  $\mathbf{x}$  to the model  $\mathbf{x}_s(\boldsymbol{\pi})$  estimated from a simulation with cosmological parameters  $\boldsymbol{\pi}$ , using the covariance matrix evaluated at the fiducial cosmology  $\boldsymbol{\pi}_0$ :

$$\chi^2(\mathbf{x}, \boldsymbol{\pi}) = (\mathbf{x} - \mathbf{x}_s(\boldsymbol{\pi}))^T \Sigma^{-1}(\boldsymbol{\pi}_0) (\mathbf{x} - \mathbf{x}_s(\boldsymbol{\pi})) \quad (14)$$

In contrast to the case of 2PCFs, there is no simple analytical prescription for the variation of the peak distribution with cosmology  $\mathbf{x}_s(\boldsymbol{\pi})$ . In fact analytical models exist for the high-S/N peaks, as used in Paper I, but cannot be applied to lower-S/N peaks. For each cosmology, we therefore average the peak distribution over the different realizations of cosmic variance and shape noise, before using them in the  $\chi^2$  computation. We note that the goal here is to have the best knowledge of the expectation value which is different than in the covariance matrix where we want to estimate the variation of the peak distribution with noise. This is also the reason why for the different cosmologies we use the same noise seeds but not for the fiducial ones. Using different seeds for shape noise would increase shot noise between the different cosmologies, requiring to average over a larger number of realizations to extract the cosmological dependence of the peak distribution.

The likelihood  $p(\mathbf{x}|\boldsymbol{\pi})$  is computed at each point of the interpolated grid of parameters, and normalized by the integrated likelihood over the prior support. We then determine the  $1\sigma$  (respectively,  $2\sigma$ ) iso-likelihood contours as the contours enclosing 68 per cent (respectively, 95 per cent) of the total integrated likelihood. For each parameter, we also estimate the most favoured value as the maximum of the likelihood marginalized over the other parameter, and the  $1\sigma$  uncertainty such that it encloses 68 per cent of the marginalized likelihood integrated over the probed parameter range. As the likelihood is computed in the  $(\Omega_m, \sigma_8)$  plane, we apply a change of variables to measure constraints on  $S_8 = \sigma_8 \sqrt{\Omega_m/0.3}$ :

$$p(\Omega_m, S_8|\mathbf{x}) d\Omega_m dS_8 = p(\Omega_m, \sigma_8|\mathbf{x}) d\Omega_m \left. \frac{\partial S_8}{\partial \sigma_8} \right|_{\Omega_m} d\sigma_8. \quad (15)$$

## 6.2 Systematics

Cosmological constraints from shear-peak statistics are affected with several systematics, namely: multiplicative shear bias, mean redshift bias, baryon feedback, intrinsic alignment, and boost factor. Although the impact of these biases on convergence peaks has been discussed in detail in Paper I, they might affect the present analysis differently due to the different methodology and using low-S/N peaks.

Hildebrandt et al. (2017) found that the multiplicative shear bias and mean redshift bias only have a small impact on  $S_8$ , in the case of 2PCFs applied to KiDS-450. In addition, Kacprzak et al. (2016) also found almost no impact on  $S_8$  central values in the case of peak statistics in the Dark Energy Survey – Science Verification (DES-SV) data, and that neglecting these biases leads to 15 per cent tighter constraints with the same definition of  $S_8$  as in our paper.

However, in KiDS-450, we have redshift bias better than  $\Delta z \sim 0.02$  (Kuijken et al. 2015) against  $\sim 0.05$  in DES-SV, and a shear multiplicative bias  $m \lesssim 0.01$  (Fenech Conti et al. 2017) against  $\sim 0.05$  in DES-SV, such that these two biases should be smaller in this study than in Kacprzak et al. (2016). In the case of KiDS-450 2PCFs, the multiplicative shear and photometric redshift biases have negligible effects on the  $S_8$  value and present uncertainties of about 1.7 per cent and 0.8 per cent, respectively. Assuming that these biases impact peak statistics at the same level as they impact the 2PCFs, we can derive conservative constraints by adding a null bias to our  $S_8$  estimate and adding the uncertainty on these biases in quadrature to the statistical error. In principle, it is possible to account for  $m/\Delta z$  biases by modifying their values in the simulations, computing the dependency of the peak distribution on these biases, and then marginalize over it. However this would require lots of computational time for such a small bias as noted in Hildebrandt et al. (2017).

The Dietrich & Hartlap (2010) simulations are DM only, such that they neglect the impact of baryons, which can modify how LSS evolves. Using a set of hydrodynamical simulations, Osato, Shirasaki & Yoshida (2015) measured the impact of baryons on both power spectrum and peak statistics. Their simulations also account for feedback from supernovae and active galactic nuclei. They find a similar bias due to baryonic effects for different ranges of peak amplitudes between  $1 \leq S/N \leq 5$ , and estimate an  $\sim 1.5$  per cent effect on both  $\Omega_m$  and  $\sigma_8$ . This propagates to a  $-2.3$  per cent effect on  $S_8 = \sigma_8 \sqrt{\Omega_m/0.3}$ . We note that we also include peaks in the range  $0 \leq S/N \leq 1$  in our analysis, but Yang et al. (2011) found that for low peaks, the impact of baryons is mitigated, because light rays towards the peaks do not pass near the cores of the haloes along the line of sight, such that our estimate of the baryon bias should be conservative. Our filter function also downweights the central part, i.e. the peak height is not primarily determined by the central portion of the halo if the matched filter and the halo are aligned. This further decreases the impact of baryonic effects. We apply a  $-2.3$  per cent bias to our  $S_8$  estimate in order to correct for baryons. We also add this value in quadrature to the error budget, therefore assuming an uncertainty on the bias as large as the bias itself. This is a conservative approach to account for the fact that we do not accurately know the uncertainty on this bias.

On small scales, IA refers to the radial alignment of satellite galaxies within DM haloes which breaks the fundamental assumption of WL that galaxies are randomly oriented. These alignments are generated by the gravitational potential of high-mass haloes on neighbouring galaxies. This effect is divided into two components: the intrinsic–intrinsic correlations (II), i.e. the alignment of galaxies physically linked together, and gravitational–intrinsic correlations (GI), i.e. the alignment of halo galaxies with the induced shear on background galaxies (Hirata & Seljak 2004). In the case of peaks, the effects of IA can be captured by modelling the alignment of satellite galaxies towards DM halo centres, e.g. the Schneider & Bridle (2010) model. Using this model with the fiducial value for the alignment strength prescribed in Schneider & Bridle (2010), Kacprzak et al. (2016) found a change in the amplitude of shear peaks lower than 5 per cent, applying the same methodology as ours to the DES-SV data. We also note that Sifón et al. (2015) measured the radial alignment of satellite galaxies in a sample of 90 galaxy clusters, securing cluster membership through spectroscopic redshifts, and found negligible alignments. Based on their measurement, they show that the Schneider & Bridle (2010) recommended alignment strength overestimates the IA at small scales (see their fig. 13), such that the effect of IA on the peak distribution is probably much lower than what Kacprzak et al. (2016) found.

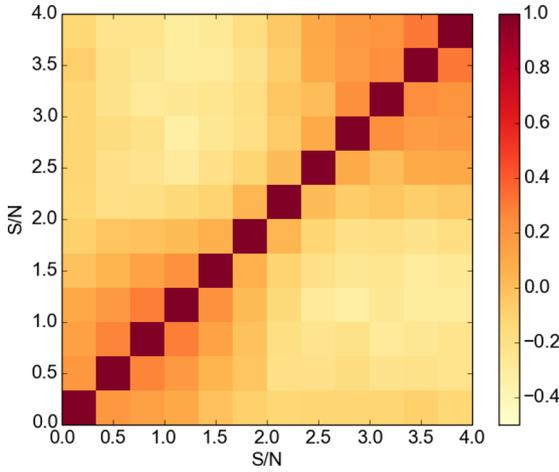
The dilution of the background shear signal due to the inclusion of cluster galaxies is generally compensated for by a radially-dependent boost factor to the shear in cluster lensing studies (e.g. Applegate et al. 2014; Hoekstra et al. 2015; Martinet et al. 2016). In the case of peak statistics, the contamination from cluster galaxies leads to higher peaks in the simulations than in the observations. Around an observed galaxy cluster, the background shear signal is diluted. But in the simulated mocks, where galaxies have the same positions as in the data, at a DM halo position there is no dilution of the shear signal because the distribution of galaxies is imposed by the data. Comparing the radial profile of galaxy density at peak locations in the observation with that of simulations allows one to compute the boost factor in bins of peak S/N. This also accounts for the loss of correlation between the lensing mass and source galaxy distributions when adapting the simulated mocks to the observation (see Section 3.2). With the same peak calculation and simulations as ours, Kacprzak et al. (2016) estimated the variation of the number of peaks per S/N bin due to the boost factor in the DES-SV. They found a variation which is proportional to the S/N of peaks and lower than about 5 per cent for S/N lower than 4, and therefore recommend using bins with S/N lower than this value to avoid large shear dilution effects. In a similar approach but on convergence peaks in KiDS-450, Paper I found a change of about 6 per cent and 10 per cent in the number of peaks in the bin with  $3 < S/N < 3.5$  and  $3.5 < S/N < 4$ , respectively, which correspond to the highest S/N used in this study, and is comparable with the results from DES-SV although the redshift distributions of both surveys are different.

Applying both IA and boost factor corrections, Kacprzak et al. (2016) found a variation of  $S_8$  of 0.01 using shear peaks defined with the same filter as ours, corresponding to a systematics relative bias of  $\sim 1.3$  per cent. We note that IA and the boost factor tend to increase  $S_8$  together. Based on the discussion of the two last paragraphs, we can assume this value to be an upper limit for this systematic bias in the case of KiDS-450. We add the above estimate to our  $S_8$  value and add it in quadrature to the error budget. As noted in Kacprzak et al. (2016), current models correcting for IA and boost factors have a high uncertainty in the case of peak statistics. This highlights a lack of extensive study on the impact of these systematics on peak statistics, and dedicated studies are required to improve these models, which is beyond the scope of this paper.

The biases estimated above are linearly added to our  $S_8$  best estimate and the uncertainties on these biases are added in quadrature to the statistical 68 per cent errors on  $S_8$ . We note that except for the multiplicative shear and mean redshift biases for which we have estimates of the uncertainties, we assumed that the uncertainty on each bias is as large as the bias itself. This allows us to correct for biases in a conservative manner although we lack precise information on the bias uncertainties in the case of baryons, IA, and boost factor. In doing so we also neglect any correlation between the different systematics, except that between the boost factor and IA which are treated together. The joint contribution of every bias leads to a shift of the  $S_8$  value of  $-0.95$  per cent, which is lower than the percent because some biases compensate each other. The total systematic uncertainty is  $\sim 3.2$  per cent and is dominated by baryon feedback. It is added in quadrature to the statistical precision. This value is also similar to the  $\sim 3.6$  per cent systematic uncertainty that was assigned to the 2PCFs analysis (Hildebrandt et al. 2017).

### 6.3 Results

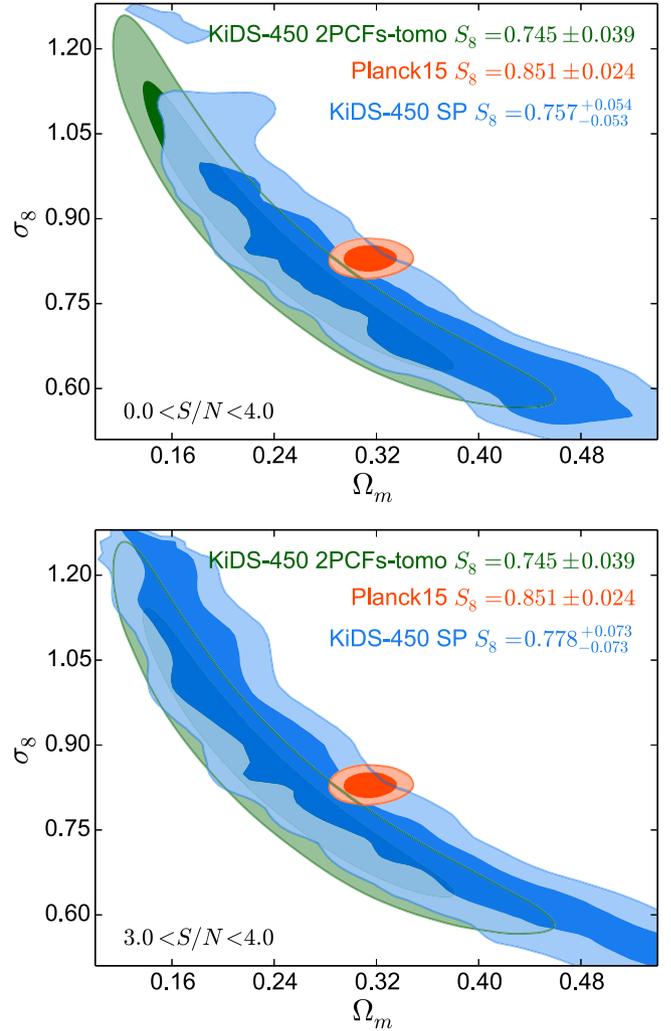
We first show the correlation matrix in Fig. 6. As mentioned earlier, we work with the mean covariance matrix over five realizations



**Figure 6.** Correlation of the covariance matrix for peaks in the range  $0 \leq S/N \leq 4$ .

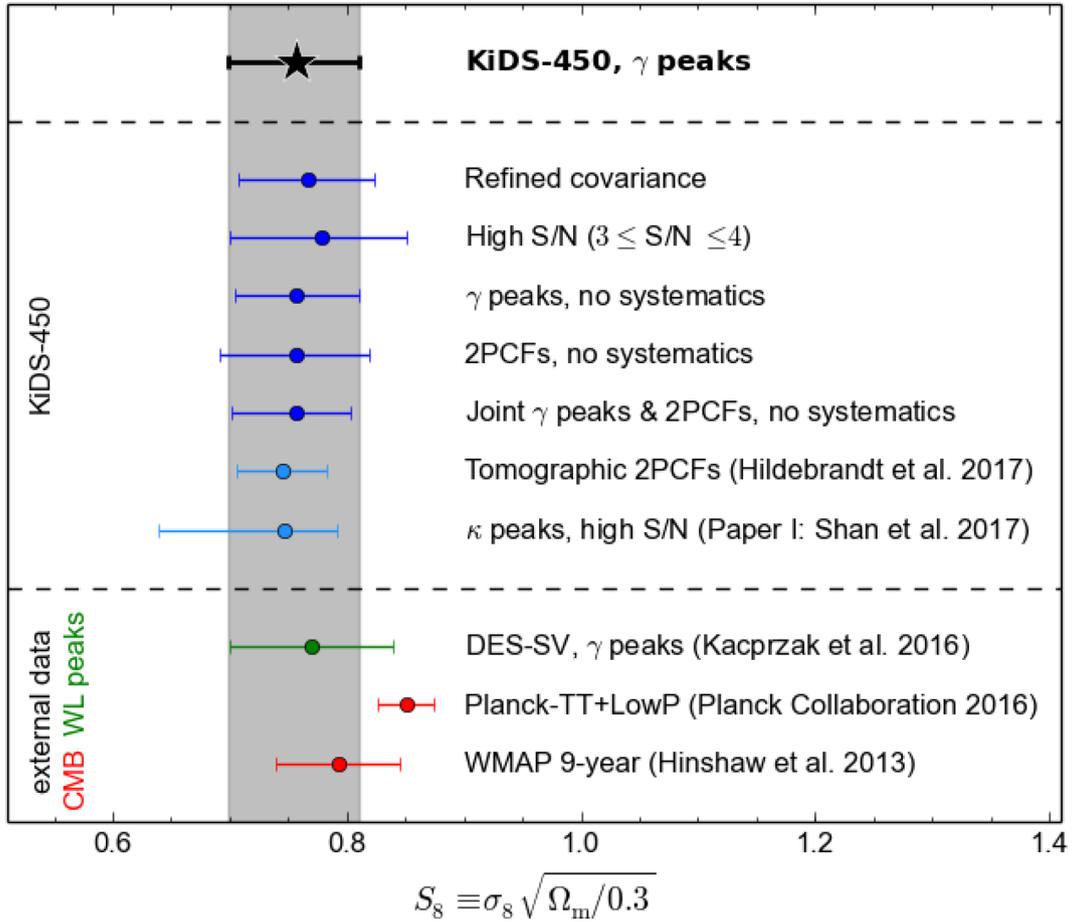
of shape noise, decreasing shot noise in the covariance estimate, although it is still representative of the noise in the data. We note that low- $S/N$  peaks are slightly correlated with one another ( $S/N < 1.66$ ), and high- $S/N$  peaks ( $S/N > 2.66$ ) show even stronger correlations. This is expected as a massive halo tends to correspond to several peaks both due to its large size and its large amount of substructures. However, we see only small correlations between the two regimes of peaks, with close to zero negative off-diagonal terms. This means that the low- and high- $S/N$  peaks probe different information, projections of small structures, and high-mass haloes, respectively. The slight anticorrelation in between the two regimes is also present in the covariance matrix of Zorrilla Matilla et al. (2016) and is due to the fact that when a large halo is detected, projection effects around this halo fade. This is also seen in the peak distribution (Fig. 4) which shows negative differential peak counts in the low- $S/N$  regime and positive ones in the high regime. These ranges of  $S/N$  also roughly correspond to the  $S/N$  where the peak distribution is the most sensitive to cosmology, as seen in Fig. 4.

Cosmological constraints from shear-peak statistics are displayed in Fig. 7, where we show the  $1\sigma$  and  $2\sigma$  contours for the 2D likelihood, and the  $S_8$  best estimate from the marginalized 1D likelihood. We do not present any estimate of  $\Omega_m$  or  $\sigma_8$  because they are highly correlated as shown by the large degeneracy in Fig. 7. We present constraints using the full range of available peak  $S/N$  ( $0 \leq S/N \leq 4$ ), and also using only the high- $S/N$  peaks ( $3 \leq S/N \leq 4$ ). This second plot serves to assess the gain of information from the low- $S/N$  peaks, and also to allow a comparison with peak constraints from analytical predictions as in Paper I. We also note the presence of wiggles in the contours, which are an artefact of the interpolation of the peak distribution with large separation between points in the  $(\Omega_m, \sigma_8)$  plane. These wiggles would disappear if we could use simulations paving more points in the cosmological parameter space. Our best estimates are  $S_8 = 0.757^{+0.054}_{-0.053}$  (68 per cent errors) for the full range of  $S/N$ , and  $S_8 = 0.778^{+0.073}_{-0.073}$  when focusing on high- $S/N$  only. Including the systematics estimated in Section 6.2 yields  $S_8 = 0.750^{+0.059}_{-0.058}$  and  $0.771^{+0.077}_{-0.077}$  for all and high-only  $S/N$ , respectively. The statistical error on  $S_8$  is  $\sim 7.1$  per cent (respectively, 9.4 per cent for high-only peaks) and the systematic uncertainty is  $\sim 3.2$  per cent. Statistical errors therefore dominate systematic ones in the case of KiDS-450. This will no longer be the case for larger surveys and detailed studies are required to better understand, and correct for the systematics affecting shear-peak statistics.



**Figure 7.** Cosmological constraints on  $(\Omega_m, \sigma_8)$ . The  $1\sigma$  and  $2\sigma$  confidence regions are displayed in dark and light areas, respectively. Green contours correspond to KiDS-450 tomographic 2PCFs constraints marginalized over systematics (Hildebrandt et al. 2017), blue to the KiDS-450 shear peak constraints using bins in ranges  $0 \leq S/N \leq 4$  (top), and  $3 \leq S/N \leq 4$  (bottom), and red to *Planck* ‘TT+lowP’ (Planck Collaboration XIII 2016). The top right legend shows the  $S_8$  best value and 68 per cent errors for each study.

In Appendix C, we make use of the refined SLICS simulations to verify that the assumptions made in case of the Dietrich & Hartlap (2010) simulations do not significantly affect the main results of the paper. We find that the refined covariance matrix computed from the SLICS simulations present similar correlations as that of the fiducial mocks, but a higher scatter due to a better inclusion of sample variance. With the refined covariance matrix, we find  $S_8 = 0.760^{+0.061}_{-0.058}$  and  $0.771^{+0.074}_{-0.075}$ , respectively, for  $0 \leq S/N \leq 4$  and  $3 \leq S/N \leq 4$ , and with accounting for systematics. The constraints on  $S_8$  are left almost unchanged by switching between the original and the refined covariance matrix, validating the various approximations made in the Dietrich & Hartlap (2010) mocks (e.g. interpolation and redshift range). We also note that the degeneracy in the  $(\Omega_m, \sigma_8)$  plane does not change. Although the sample variance bias of our simulations has negligible effect on this study, it will become more important for larger area surveys and it might become necessary to



**Figure 8.** Comparison between various constraints on  $S_8$ . The black star represents our best estimate, and the shaded area its error bars. Dark and light blue points correspond to different constraints with peak statistics and 2PCFs, from this paper and from other KiDS-450 papers, respectively. The green point corresponds to a shear peak study for another data set, and red points show constraints from CMB measurements.

use simulations which cover an area which is close to that of the data to account for sample variance.

## 7 DISCUSSION

Fig. 8 summarizes  $S_8$  constraints from this survey and compares them with various other studies. We calculated  $p$ -values as an estimate for the goodness of fit for all the cases considered. The  $p$ -values are calculated for the minimum  $\chi^2$  taking into account the degrees of freedom given by the number of data points minus two free parameters ( $\Omega_m$  and  $\sigma_8$ ). All the values are larger than 0.2, indicating that the models fit the data well.

### 7.1 Information from low- and high-S/N peaks

We first focus on the gain of information from adding the low-S/N peaks. We recall that the large-S/N peaks correspond to single massive haloes, while the low-S/N correspond to alignment of smaller haloes along the line of sight. We find very good agreement between the two regimes, showing that chance alignments and larger haloes are both good tracers of LSS. The constraints shrink by 24 per cent when adding the low-S/N peaks, representing a large gain of information. This highlights the great interest of studying the low-amplitude peaks, which efficiently probe the cosmological information contained in the chance alignments of LSS. This result

is also supported by the study of Shirasaki (2017) who showed that low-amplitude peaks contain significant non-Gaussian information.

In Appendix B, we additionally probe the potential gain from including negative S/N peaks. However, we found that those do not further increase the constraints on  $S_8$  compared to the positive peaks only. This can be explained by the high correlation that we find between the negative and positive peaks.

### 7.2 Comparison with KiDS 2PCFs and Planck

One of the goals of this study is to check whether peak statistics agree with KiDS 2PCFs, in light of the reported mild tension between the latter and *Planck* results.

Peak statistics yield similar constraints on  $S_8$  as 2PCFs. In particular, the degeneracy in the  $(\Omega_m, \sigma_8)$  plane is parallel to that of 2PCFs (Fig. 7), highlighting the strong correlation between the two probes. We note that our estimate of  $S_8$  is in good agreement with the tomographic 2PCFs value reported in Hildebrandt et al. (2017):  $S_8 = 0.745 \pm 0.039$ . We stress, however, that in the case of peaks, we included the different systematics (multiplicative shear bias, mean redshift bias, baryon feedback, intrinsic alignment, and boost factor) as a correction to the best estimate and uncertainties of  $S_8$ , while Hildebrandt et al. (2017) marginalized over the relevant systematics. Hildebrandt et al. (2017) also varied all cosmological parameters, while we can only vary  $\Omega_m$  and  $\sigma_8$  in the case of peaks with the given set of simulations, and therefore underestimate the confidence

regions. Finally, we note that we did not carry out a tomographic analysis of the peak statistics, while the 2PCFs study captures the information from four different source redshift bins. This choice is due to limitations in the available simulation mocks, and explains why constraints are tighter in the case of 2PCFs. Martinet et al. (2015) showed that a tomographic approach can improve constraints from peak statistics by almost a factor of two in the case of *Euclid*-like simulations. The improvement from a tomographic peak analysis has also been noted by Dietrich & Hartlap (2010) using CFHT-like simulations, and by Petri et al. (2016) with Large Synoptic Survey Telescope (LSST)-like simulations.

We find only a slight difference when comparing peak statistics with *Planck* CMB. Our constraints on  $S_8$  present a  $1.6\sigma$  difference with that of *Planck* ( $S_8 = 0.851 \pm 0.024$ , Planck Collaboration XIII 2016), when including systematics. Hildebrandt et al. (2017) reported a tension of  $2.3\sigma$ , however our constraints are 33 per cent weaker in comparison, mainly due to the fact we are not using tomography: we show in Section 7.5 that peak statistics and 2PCFs achieve similar constraints when tomography is not used in both cases. Our best estimate of  $S_8$  is however closer to that of Hildebrandt et al. (2017). Finally, we note that our simulations are run with a Hubble parameter  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$  which is different from the *Planck* estimated value:  $H_0 = 67.3 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . While this could partially explain the difference between the two probes, it is difficult to assess in this paper, as it would require to run extra simulations fixing  $H_0$  to the *Planck* value. We note that our constraints do not change significantly when using the refined covariance matrix (see Appendix C) which is computed with a slightly different Hubble parameter:  $H_0 = 69.0 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . In addition, we found in Paper I that the constraints on  $S_8$  are stable when varying  $H_0$  between  $68$  and  $72 \text{ km s}^{-1} \text{ Mpc}^{-1}$  with the modelled peak function, showing that the observed difference to *Planck* is probably not due to different values of the Hubble constant.

### 7.3 Comparison with other WL peak analyses

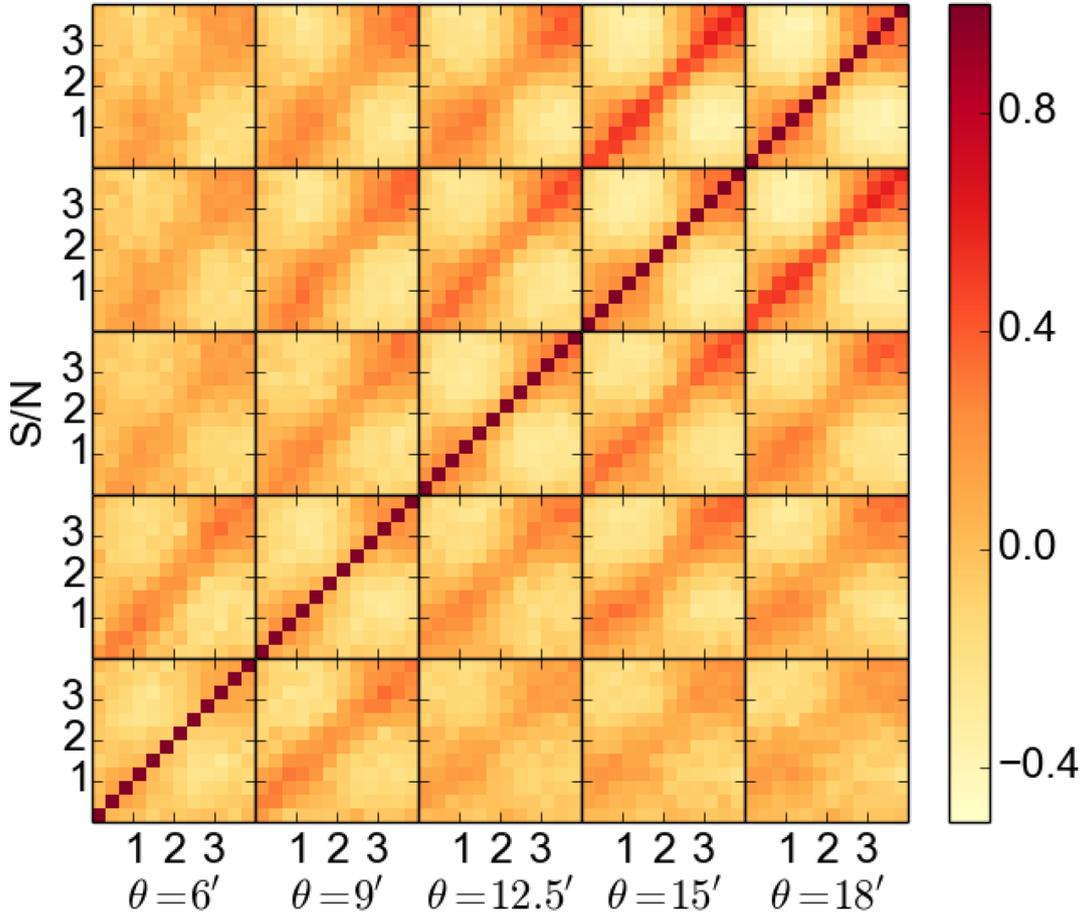
We first compare our constraints to that of Paper I, which uses the same KiDS-450 data set but apply a different approach to peak statistics. Paper I uses convergence peaks instead of shears, and also compares their peak distribution to that from analytical predictions, although calibrating those on the same simulations as the ones we use in this study. Paper I uses a Gaussian filter with a scale of 2 arcmin on the convergence field, which is close to the effective radius of our own filter ( $\theta = 1.875$  arcmin). As the analytical predictions are accurate only in the high-S/N regime where one can assume that peaks are associated with high-mass haloes, Paper I makes use of the high-S/N peaks only ( $3 \leq S/N \leq 5$ ). We therefore only compare our high-S/N ( $3 \leq S/N \leq 4$ ) constraints with the results from Paper I. In this range of S/N Paper I finds  $S_8 = 0.746^{+0.046}_{-0.107}$ . We note that these constraints are in good agreement with ours, with the same constraining power as our high-S/N regime (see Fig. 8). The fact that both studies, while based on completely different approaches, give consistent  $S_8$  estimates is a good assessment of their robustness.

There is, however, a major difference in the orientation of the  $(\Omega_m, \sigma_8)$  degeneracy. In this study, we find a degeneracy similar to that of the 2PCFs even for the high peaks, while in Paper I a flatter degeneracy is seen, closer to that of some cluster studies. The only way for us to reproduce this degeneracy is to use only the highest S/N peaks:  $4 \leq S/N \leq 5$ . This would mean that even in the  $3 \leq S/N \leq 4$  range, there is a large contribution from peaks corresponding to projection effects of LSS and to noise while peaks at  $S/N \geq 4$

mainly correspond to high-mass haloes. The fact that Paper I find a flatter degeneracy even when including peaks at  $3 \leq S/N \leq 4$  could highlight a difference in the definition of S/N which is global in Paper I and local here (the noise is different for each aperture), or show some limits of the model used in Paper I which only accounts for high-mass haloes and not for projection effects. This reasoning holds only under the assumption that the degeneracy is not dominated by other systematic effects. In the literature, we note that different peak studies find different degeneracy orientations. Liu et al. (2015b) find a degeneracy very close to that of Paper I, applying the same predictive model to the Stripe82 data acquired with the CFHT. Using simulations instead of analytical prediction, Liu et al. (2015a) find a degeneracy similar to the convergence power spectrum degeneracy, but also using the full range of S/N convergence peaks. Kacprzak et al. (2016) also find a degeneracy in agreement with that of 2PCFs in DES-SV, using simulations and shear peaks in the full range  $0 \leq S/N \leq 4$ . Quantitatively, when we let the  $\alpha$  parameter free to vary in  $\Sigma_8 = \sigma_8(\Omega_m/0.3)^\alpha$ , we find that the best  $\alpha$  value is between 0.54 and 0.61, and that our constraints on  $\Sigma_8$  are  $\sim 5$  per cent tighter than those with  $\alpha$  fixed to 0.5. This is in good agreement with the best  $\alpha$  values of 0.60 and 0.58 found in Liu et al. (2015a) and Kacprzak et al. (2016), but different from that of Liu et al. (2015b) and Paper I: 0.43 and 0.38, respectively. Simulation-based peak analyses therefore tend to show degeneracy closer to second-order cosmic shear than to clusters but also often make use of a larger range of S/N compared to model-based analyses which are in better agreement with the degeneracy from cluster analyses.

We can also compare our results with those of the DES-SV analysis (Kacprzak et al. 2016). This study is very similar to ours in its methodology. In particular, they use aperture masses with the same filter function in shear space. They also use the same Dietrich & Hartlap (2010) simulations to compute their cosmological constraints. Although they define their fiducial results based on a 20 arcmin filter scale, they also provide  $S_8$  estimates for a 12 arcmin scale, very close to the 12.5 arcmin used in this study. We display their value for  $\theta = 20$  arcmin in Fig. 8 as it corresponds to their fiducial estimate. They find  $S_8 = 0.76 \pm 0.074$  and  $0.73 \pm 0.081$  with a 20 and 12 arcmin filter scale, respectively. These two results are in very good agreement with ours. The error bars are 21 per cent narrower (respectively, 28 per cent for  $\theta = 12$  arcmin) in the case of KiDS due to its larger area ( $450 \text{ deg}^2$  against  $139 \text{ deg}^2$ ) and higher galaxy density (8.53 galaxies per square arcminute against 5.7). We also note that Kacprzak et al. (2016) marginalized over the estimated mean redshift and multiplicative biases. While we do not marginalize over systematics, we include them as an a posteriori correction to our best estimate and uncertainties, also accounting for baryon feedback, boost factor, and IA. Finally, we note that they found similar  $S_8$  when using different aperture sizes. In Section 7.4, we measure the peak statistics for different filter scales and derive constraints from a multiscale analysis.

The cosmological constraints from Liu et al. (2015b) and Liu et al. (2015a) are not presented in Fig. 8 because they use a different definition of  $\Sigma_8$  than ours. While we use  $S_8 = \sigma_8(\Omega_m/0.3)^{0.5}$ , these two other studies, respectively, use  $\Sigma_8 = \sigma_8(\Omega_m/0.27)^{0.43}$  and  $\sigma_8(\Omega_m/0.27)^{0.60}$ . When recomputing our constraints with their definition we find good agreement with both studies. Liu et al. (2015b) find  $\Sigma_8 = 0.82 \pm 0.03$  from the  $\sim 130 \text{ deg}^2$  of the CFHT Stripe 82 Survey with a model-based parameter inference, while we find  $\Sigma_8 = 0.768^{+0.063}_{-0.062}$  and  $0.775^{+0.058}_{-0.057}$  with and without taking systematics into account. Liu et al. (2015a) find  $\Sigma_8 = 0.84^{+0.03}_{-0.04}$  from the  $154 \text{ deg}^2$  of the CFHTLenS in the redshift range  $0.2 < z < 1.3$



**Figure 9.** Correlation of the multiscale covariance matrix for peaks in the range  $0 \leq S/N \leq 4$ . Each square represents the correlation of the peak S/N distribution between two filter scales  $\theta$  in (6, 9, 12.5, 15, and 18 arcmin).

with a simulation-based parameter inference, while we find  $\Sigma_8 = 0.821^{+0.056}_{-0.066}$  and  $0.829^{+0.049}_{-0.060}$  with and without taking systematics into account. We note that their constraints are tighter than ours, while we probe three times as many galaxies as them. It is hard to assess the origin of this difference, and in particular whether it is due to the different surveys or to the different methods. On the first point, we note that in the CFHTLenS they use higher redshift source galaxies that have a stronger lensing signal compared to lower redshift galaxies in KiDS. We are currently at an early stage of peak statistics and it would be valuable to apply different techniques to the same data sets to do robust comparisons between methods. We also note that we understand better the constraints when applying the same method to two different surveys as in the case of the comparison between the present analysis and the DES-SV peak study of Kacprzak et al. (2016).

#### 7.4 Multiscale analysis

We investigate the gain of information from combining the peak statistics of different filter scales. As different filter scales probe different structures, i.e. different halo sizes, combining several scales should yield more precise constraints. However, the information from different scales is correlated: for example a galaxy cluster detected at a smaller scale will be detected at larger scale providing that the scale is not so large that the cluster signal is buried in the noise. See also Marian et al. (2012) for an approach with a single scale of adaptive size. In addition to our fiducial scale of

$\theta = 12.5$  arcmin, we measure the peak distribution for four extra scales leading to the following ensemble of probed scales:  $\theta = (6, 9, 12.5, 15, \text{ and } 18 \text{ arcmin})$ , which respectively correspond to the effective scales  $\theta_{\text{eff}} = (0.9, 1.35, 1.875, 2.25, \text{ and } 2.7 \text{ arcmin})$ , for the filter parameter  $x_c = 0.15$  [see equation (9)]. We recall that the fiducial scale is chosen such as to maximize the number of peaks at  $S/N \geq 3$ . We measure the peak distribution of each scale in the observations and in the simulations, again with five random realizations of shape noise. The multiscale cosmological inference is done in the same way as for the single scale but with a data vector which is the concatenation of the data vectors of the individual scales. The data vector contains 60 elements for the combination of the five scales, which is small enough compared to the number of simulations ( $N_s = 175$ ) to compute accurate constraints with the Sellentin & Heavens (2016) likelihood.

The joint correlation matrix is shown in Fig. 9. This is the mean correlation over five realizations of shape noise, each of which contains the 175 fiducial mocks. We see that the different scales are highly correlated to one another, and also that the closer scales show more correlations. The correlations are larger at large scales, since we increase the scale linearly while the number of galaxies included in the aperture scales with the area, the difference between  $\theta = 6$  and 9 arcmin is therefore larger than the difference between  $\theta = 15$  and 18 arcmin.

We compute the constraints on  $S_8$  from each individual scale and for different combinations of scales. While all estimates are consistent with the fiducial single-scale analysis, the improvement

in precision is at best of  $\sim 10$  per cent. This value is reached when using the two scales  $\theta = 12.5$  and  $15$  arcmin together. Adding extra scales does not improve the constraints further, such that it is not necessary to combine more than two scales. This is supported by the large amount of correlation between scales found in the correlation matrix (see Fig. 9). The constraints might even get less precise when adding extra scales, probably because of anticorrelations between scales. For example, the combination of the five scales yields marginally better constraints than the fiducial scale alone. In addition, we note that the single-scale analyses yield the most precise constraints for  $\theta = 12.5$  arcmin as expected by definition of our fiducial scale. Because multiscale constraints are only mildly better than the single-scale case, we recommend using only one scale to save computation time. We note that Liu et al. (2015a) also conducted a multiscale analysis and found that combining more than two scales does not improve the 2D contours in the  $(\Omega_m, \sigma_8)$  plane further. Their study also seems to show that the multiscale approach is only marginally better in terms of  $S_8$  estimates than the single-scale method.

### 7.5 Peak statistics and 2PCFs joint analysis

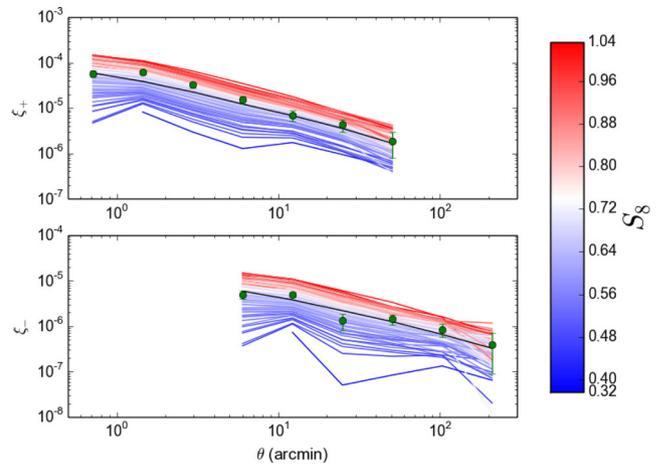
While peaks represent a different statistic than the 2PCFs, they are both sensitive to LSS and therefore probe correlated information. In this section, we use the 2PCFs as the statistics for cosmic shear and find its joint cosmological constraints with peak statistic. The 2PCFs are defined as

$$\xi_{\pm}(\theta = |\theta_a - \theta_b|) = \frac{\sum_{a,b} w_a w_b [\epsilon_t(\theta_a) \epsilon_t(\theta_b) \pm \epsilon_{\times}(\theta_a) \epsilon_{\times}(\theta_b)]}{\sum_{a,b} w_a w_b}, \quad (16)$$

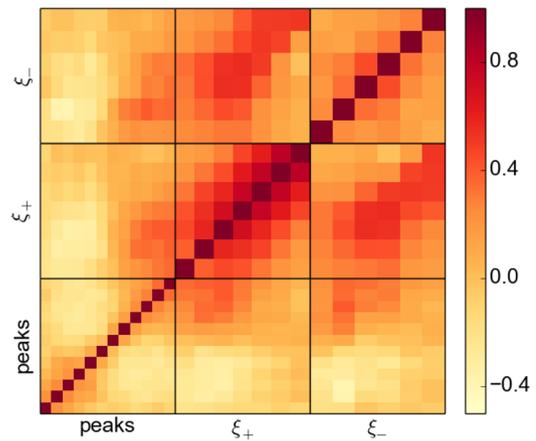
where the sum is over pairs of galaxies  $a$  and  $b$  with separation  $|\theta_a - \theta_b|$  and *lensfit* weights  $w_a$  and  $w_b$ .  $\epsilon_t$  (respectively,  $\epsilon_{\times}$ ) represents the ellipticity component tangential (respectively, perpendicular) to the line between the two galaxies. Shear 2PCFs relate to cosmological parameters through the matter power spectrum (see e.g. Bartelmann & Schneider 2001; Kilbinger 2015).

We measure the  $\xi_+$  and  $\xi_-$  correlation functions using the ATHENA software (Kilbinger, Bonnett & Coupon 2014) with 900 logarithmic bins in the range  $0.5 \text{ arcmin} \leq \theta \leq 300 \text{ arcmin}$  and rebin to nine points logarithmically spaced between the separation limits. In contrast to other KiDS-450 second-order cosmic shear studies (Hildebrandt et al. 2017; Köhlinger et al. 2017; van Uitert et al. 2017) which follow a tomographic approach, we use a single redshift bin in the range  $0.1 \leq z_B \leq 0.9$ . The first reason for this is that we want to compare peaks and 2PCFs with similar approaches, and the second is that in this analysis we get the cosmological constraints from comparing the 2PCFs of the observations with that of the simulations, and the simulated mocks we are using are not suited for the tomographic approach. This analysis is also different from the other KiDS cosmic shear papers, as we derive the constraints from  $N$ -body simulations and not from an analytical prescription. We prefer to use the simulations in this study because we are only interested in the qualitative improvement from the combination of constraints and to ensure that systematics from the simulation approach affect both peaks and 2PCFs measurements. Also, only two cosmological parameters are enabled to vary ( $\Omega_m$  and  $\sigma_8$ ), rather than five or more in the other KiDS-450 cosmic shear studies.

We measure the 2PCFs in the observation and in the Dietrich & Hartlap (2010) simulations which follow the KiDS footprint, with same weights and shape noise amplitude as in the data. We note that the 2PCFs do not depend on galaxy positions such that we can use



**Figure 10.** Shear 2PCFs:  $\xi_+$  (top) and  $\xi_-$  (bottom). Green dots represent KiDS data with error bars from the dispersion in the fiducial cosmologies, black line the mean of the fiducial cosmologies, and coloured lines the various simulations with  $S_8$  increasing from blue to red.



**Figure 11.** Joint correlation matrix for peaks in the range  $0 \leq S/N \leq 4$ ,  $\xi_+$ , and  $\xi_-$ . Each square represents the correlation of one estimator.

the same positions as in the data in the simulations without biasing the 2PCFs. This allows us to measure the 2PCFs on the exact same mocks as for the peaks, which is important to assess the level of correlation between the two probes.

Following Hildebrandt et al. (2017), we use only the seven first bins of  $\xi_+$  and the last six bins of  $\xi_-$ . These are displayed in Fig. 10 as measured in the data (green dots) with error bars from the diagonal elements of the fiducial covariance matrix, in the fiducial simulations (black line for the mean) and in the various cosmologies ranging from low  $S_8$  (blue) to high (red). As for the case of peaks (see Fig. 4), we find a clear dependence on cosmology, with higher  $S_8$  corresponding to higher level of correlation of the shear. The shear correlation in the data is also slightly higher than the fiducial cosmology favouring a moderately higher  $S_8$  value. We note that at the largest scale,  $\xi_-$  presents large error bars. This is because the simulations we use are only  $6 \times 6 \text{ deg}^2$  and we are probing correlations between pairs separated by as much as  $5 \text{ deg}$ , significantly lowering the number of pairs compared to smaller scales.

The joint correlation matrix is given in Fig. 11, for the 175 fiducial mocks, and averaged over five random realizations of shape noise. The peaks span the range  $0 \leq S/N \leq 4$ , and  $\xi_+$  and  $\xi_-$  correspond to the seven and six aforementioned bins, respectively. We see a strong

correlation between the different scales of  $\xi_+$  and also between  $\xi_+$  and  $\xi_-$ . This is also seen in the correlation matrix presented in the appendix of Hildebrandt et al. (2017) for the tomographic case. We also find some correlation between the high-S/N peaks and the small-scale  $\xi_+$  and  $\xi_-$ . This is expected as the peaks are probing LSS on scales of about the size of the filter applied in the aperture mass. However, at other scales we note that the correlations remain low, highlighting that the peak statistics are largely independent from the 2PCFs.

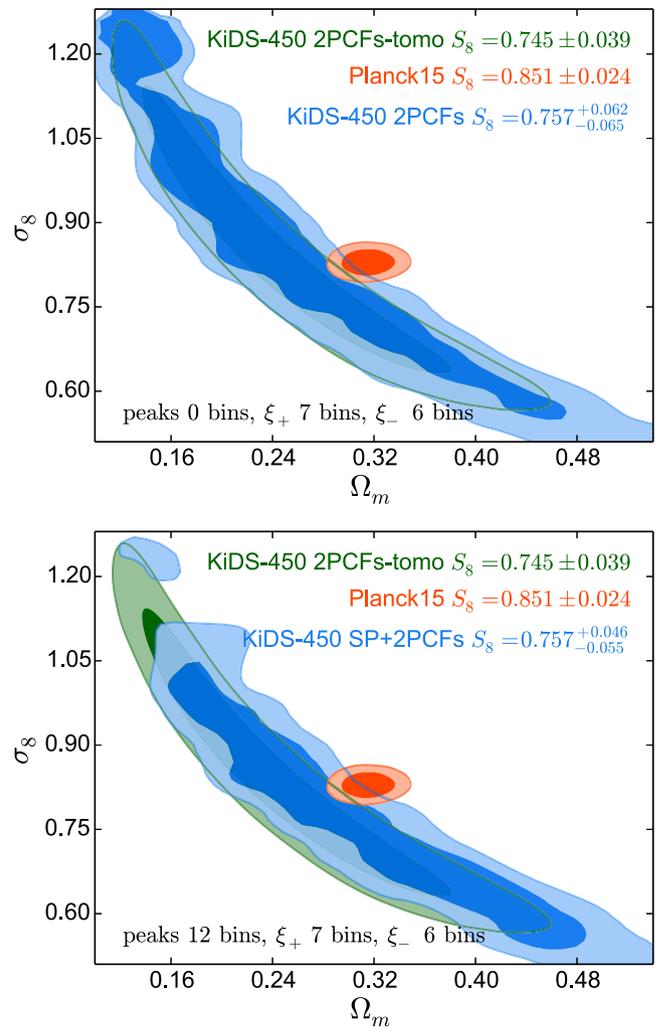
We build new data vectors to estimate cosmological constraints with the same method as in the rest of the paper. We use the Sellentin & Heavens (2016) likelihood with the data, the covariance matrix from the mocks with fiducial cosmologies, and the model from the simulations with various cosmologies interpolated from the mesh at which simulations exist. We probe the constraints from the 2PCFs alone with the concatenation between  $\xi_+$  and  $\xi_-$  as the data vector, and also the joint constraints with the concatenation of peaks between  $0 \leq S/N \leq 4$ ,  $\xi_+$  and  $\xi_-$ . The number of bins of these data vectors are, respectively, 13 and 25, which is still reasonably low compared to the 175 realizations of the fiducial cosmologies used to estimate the covariance matrix.

Fig. 12 shows the constraints for the 2PCFs and the joint constraints. There is a very good agreement between the present non-tomographic 2PCFs constraints and the tomographic constraints of Hildebrandt et al. (2017) both in the  $(\Omega_m, \sigma_8)$  degeneracy and in the  $S_8$  estimate. Our errors are however  $\sim 39$  per cent larger because we do not use the information from the different redshifts. Quantitatively, the constraints from 2PCFs only are  $S_8 = 0.757^{+0.062}_{-0.065}$ , which is the same value as for the peaks but with  $\sim 16$  per cent larger statistical errors. This highlights the very high potential of peak statistics as a cosmological probe compared to the classical WL probe. Furthermore, the combination of both yields an  $\sim 20$  per cent improvement compared to the 2PCFs alone with  $S_8 = 0.757^{+0.046}_{-0.055}$  but no significant improvement from peaks alone. We however stress that this study presents some limitations. First, the likelihood is quite noisy due to sparsity in the probed cosmologies. Second, no systematics are accounted for in this part of the discussion, so only the statistical errors are considered. This is because the impact of systematics on constraints from peak statistics is not known with the same accuracy as that of the 2PCFs, mainly because of the very recent development of peak statistics.

Our results are none the less very promising for peak statistics and call attention to the great interest of developing peaks further in terms of systematics comprehension. Our study also confirm the predictions from simulation-based analyses of the improvement of constraints from joint second-order and higher order cosmic shear over second-order alone (e.g. Dietrich & Hartlap 2010; Hilbert et al. 2012). Finally, we note that in their study of CFHTLenS, Liu et al. (2015a) also found only marginal improvement from adding convergence peaks to the convergence power spectrum compared to peaks alone, and  $\sim 40$  per cent improvement compared to power spectrum alone, taking the full covariance between the two probes into account as we do here.

## 8 CONCLUSION

In this paper, we derive constraints on  $S_8 = \sigma_8 \sqrt{\Omega_m/0.3}$  by comparing the peak statistics in the WL mass map of KiDS-450 to that of simulations with various cosmologies. Mass maps are computed by averaging the tangential shear in a 12.5 arcmin radius aperture with an NFW-like weighting function, which is compensated in the convergence field to avoid the mass sheet degeneracy.



**Figure 12.** Cosmological constraints on  $(\Omega_m, \sigma_8)$  from peaks and 2PCFs. The  $1\sigma$  and  $2\sigma$  confidence regions are displayed in dark and light areas, respectively. Green contours correspond to KiDS-450 tomographic 2PCFs constraints marginalized over systematics (Hildebrandt et al. 2017), blue to the KiDS-450 2PCFs constraints (top), and to the joint peaks in the range  $0 \leq S/N \leq 4$  and 2PCFs constraints (bottom), and red to *Planck* ‘TT+lowP’ (Planck Collaboration XIII 2016). The top right legend shows the  $S_8$  best value and 68 per cent errors for each study.

We find constraints on  $S_8$  in good agreement with those from 2PCFs (Hildebrandt et al. 2017), and also from the independent peak statistics study of Paper I. The latter uses convergence peaks and analytical predictions, focusing on the high-S/N peaks corresponding to high-mass haloes. Our  $S_8$  estimate is  $1.6\sigma$  lower than the value estimated with *Planck* CMB when we account for systematics. We also measure the gain of information when adding the low-S/N peaks, which correspond to projections of low-mass haloes, to the high-S/N peaks, corresponding to high-mass haloes. Quantitatively, the  $S_8$  estimate improves by  $\sim 25$  per cent when adding peaks with S/N lower than 3. We measure the peak distribution with various filter scales finding only a mild ( $\sim 10$  per cent) improvement from combining scales. Refining the covariance matrix to properly account for sample variance only affects cosmological constraints at the level of a few percents, validating the fiducial approach of this paper. Finally, we measure the non-tomographic 2PCFs and find consistent  $S_8$  estimates between peaks and 2PCFs. Combining both probes yields an  $\sim 20$  per cent improvement compared to 2PCFs

alone, highlighting the high potential of peak statistics for future WL surveys.

## ACKNOWLEDGEMENTS

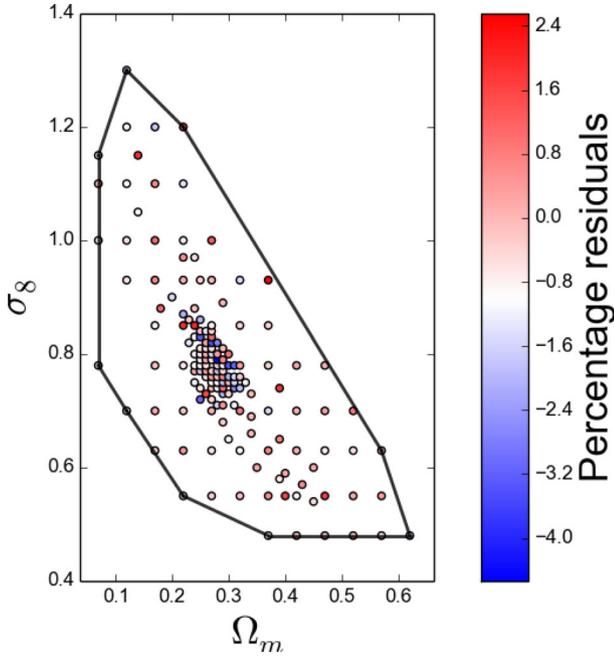
We thank Tomasz Kacprzak, Tim Schrabback, Patrick Simon, and Angus Wright for useful discussions. We additionally thank the referee for interesting comments and suggestions. HHi acknowledges support from the DFG under Emmy Noether grant Hi 1495/2-1. MA and CH acknowledge support from the European Research Council under grant number 647112. HHo acknowledges support from Vici grant 639.043.512, financed by the Netherlands Organization for Scientific Research. This work is supported by the Deutsche Forschungsgemeinschaft in the framework of the TR33 ‘The Dark Universe’ (PS and DK). The research leading to these results has received funding from the European Union’s FP7 and Horizon 2020 research and innovation programs under Marie Skłodowska-Curie grant agreement numbers 627288 and 664931 (JM). JHD acknowledges support from the European Commission under a Marie-Sklodowska-Curie European Fellowship (EU project 656869). KK acknowledges support by the Alexander von Humboldt Foundation. RN acknowledges support from the German Federal Ministry for Economic Affairs and Energy (BMWi) provided via DLR under project number 50QE1103. Computations for the SLICS  $N$ -body simulations were performed in part on the Orcinus supercomputer at the WestGrid HPC consortium ([www.westgrid.ca](http://www.westgrid.ca)), in part on the GPC supercomputer at the SciNet HPC Consortium. SciNet is funded by: the Canada Foundation for Innovation under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund – Research Excellence; and the University of Toronto.

## REFERENCES

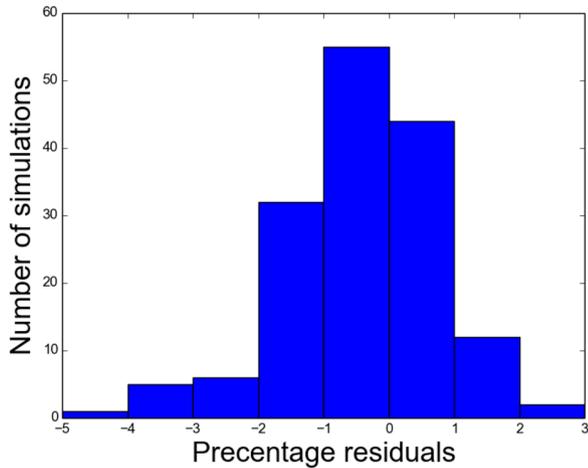
- Applegate D. E. et al., 2014, *MNRAS*, 439, 48  
 Bartelmann M., Schneider P., 2001, *Phys. Rep.*, 340, 291  
 Benítez N., 2000, *ApJ*, 536, 571  
 de Jong J. T. A. et al., 2015, *A&A*, 582, A62  
 de Jong J. T. A. et al., 2017, *A&A*, 604, A134  
 Dietrich J. P., Hartlap J., 2010, *MNRAS*, 402, 1049  
 Eifler T., Schneider P., Hartlap J., 2009, *A&A*, 502, 721  
 Eisenstein D. J., Hu W., 1998, *ApJ*, 496, 605  
 Fan Z., Shan H., Liu J., 2010, *ApJ*, 719, 1408  
 Fenech Conti I., Herbonnet R., Hoekstra H., Merten J., Miller L., Viola M., 2017, *MNRAS*, 467, 1627  
 Hamana T., Oguri M., Shirasaki M., Sato M., 2012, *MNRAS*, 425, 2287  
 Harnois-Déraps J., van Waerbeke L., Viola M., Heymans C., 2015, *MNRAS*, 450, 1212  
 Hartlap J., Simon P., Schneider P., 2007, *A&A*, 464, 399  
 Hettterscheidt M., Erben T., Schneider P., Maoli R., van Waerbeke L., Mellier Y., 2005, *A&A*, 442, 43  
 Hilbert S., Marian L., Smith R. E., Desjacques V., 2012, *MNRAS*, 426, 2870  
 Hildebrandt H. et al., 2017, *MNRAS*, 465, 1454  
 Hirata C. M., Seljak U., 2004, *Phys. Rev. D*, 70, 063526  
 Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, *MNRAS*, 449, 685  
 Jain B., Van Waerbeke L., 2000, *ApJ*, 530, L1  
 Kacprzak T. et al., 2016, *MNRAS*, 463, 3653  
 Kaiser N., Squires G., 1993, *ApJ*, 404, 441  
 Kilbinger M., 2015, *Rep. Prog. Phys.*, 78, 086901  
 Kilbinger M., Bonnett C., Coupon J., 2014, *Astrophysics Source Code Library*, record ascl:1402.026  
 Köhlinger F. et al., 2017, *MNRAS*, 471, 4412  
 Kratochvil J. M., Haiman Z., May M., 2010, *Phys. Rev. D*, 81, 043519  
 Kruse G., Schneider P., 1999, *MNRAS*, 302, 821  
 Kruse G., Schneider P., 2000, *MNRAS*, 318, 321  
 Kuijken K. et al., 2015, *MNRAS*, 454, 3500  
 Lin C.-A., Kilbinger M., 2015, *A&A*, 576, A24  
 Lin C.-A., Kilbinger M., 2017, *A&A*, preprint ([arXiv:1704.00258](https://arxiv.org/abs/1704.00258))  
 Liu J., Haiman Z., 2016, *Phys. Rev. D*, 94, 043533  
 Liu J., Petri A., Haiman Z., Hui L., Kratochvil J. M., May M., 2015a, *Phys. Rev. D*, 91, 063507  
 Liu X. et al., 2015b, *MNRAS*, 450, 2888  
 Marian L., Smith R. E., Hilbert S., Schneider P., 2012, *MNRAS*, 423, 1711  
 Marian L., Smith R. E., Hilbert S., Schneider P., 2013, *MNRAS*, 432, 1338  
 Martinet N., Bartlett J. G., Kiessling A., Sartoris B., 2015, *A&A*, 581, A101  
 Martinet N. et al., 2016, *A&A*, 590, A69  
 Maturi M., Fedeli C., Moscardini L., 2011, *MNRAS*, 416, 2527  
 Miller L., Kitching T. D., Heymans C., Heavens A. F., van Waerbeke L., 2007, *MNRAS*, 382, 315  
 Miller L. et al., 2013, *MNRAS*, 429, 2858  
 Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493  
 Osato K., Shirasaki M., Yoshida N., 2015, *ApJ*, 806, 186  
 Petri A., May M., Haiman Z., 2016, *Phys. Rev. D*, 94, 063534  
 Planck Collaboration XIII, 2016, *A&A*, 594, A13  
 Radovich M. et al., 2017, *A&A*, 598, A107  
 Schirmer M., Erben T., Hettterscheidt M., Schneider P., 2007, *A&A*, 462, 875  
 Schneider P., 1996, *MNRAS*, 283, 837  
 Schneider M. D., Bridle S., 2010, *MNRAS*, 402, 2127  
 Sellentin E., Heavens A. F., 2016, *MNRAS*, 456, L132  
 Shan H. Y. et al., 2014, *MNRAS*, 442, 2534  
 Shan H. et al., 2017, *MNRAS*, preprint ([arXiv:1709.07651](https://arxiv.org/abs/1709.07651)) (Paper I)  
 Shirasaki M., 2017, *MNRAS*, 465, 1974  
 Sifón C., Hoekstra H., Cacciato M., Viola M., Köhlinger F., van der Burg R. F. J., Sand D. J., Graham M. L., 2015, *A&A*, 575, A48  
 Springel V., 2005, *MNRAS*, 364, 1105  
 Taylor A., Joachimi B., 2014, *MNRAS*, 442, 2728  
 van Uitert E. et al., 2017, *MNRAS*, preprint ([arXiv:1706.05004](https://arxiv.org/abs/1706.05004))  
 Wang S., Haiman Z., May M., 2009, *ApJ*, 691, 547  
 Yang X., Kratochvil J. M., Wang S., Lim E. A., Haiman Z., May M., 2011, *Phys. Rev. D*, 84, 043529  
 Zorrilla Matilla J. M., Haiman Z., Hsu D., Gupta A., Petri A., 2016, *Phys. Rev. D*, 94, 083506

## APPENDIX A: EVALUATING THE INTERPOLATION OF THE PEAK DISTRIBUTION

In Section 5.2, we describe how we interpolate the peak distribution from our 157 different points in the  $(\Omega_m, \sigma_8)$  plane to an evenly spaced grid with a step of 0.01 for both parameters. To evaluate the robustness of this interpolation, we compare the interpolated numbers of peaks with the measured ones for every simulation. The relative difference is displayed in Fig. A1 for the same bin of S/N as that of Fig. 5:  $3.33 \leq S/N \leq 3.66$ . The percentage residual is always lower than 5 per cent and ranges from  $-5$  to  $+3$  per cent. The number of peaks in this bin being roughly 450, the Poisson error is about 5 per cent. The error in the interpolation process is therefore always lower than the Poisson error. In addition, we calculate the distribution of the residuals (Fig. A2) and find that it is centred around 0 per cent and that more than 85 per cent of the simulations have a relative error due to the interpolation of less than 2 per cent. This shows that the interpolation process does not add any significant systematic bias.



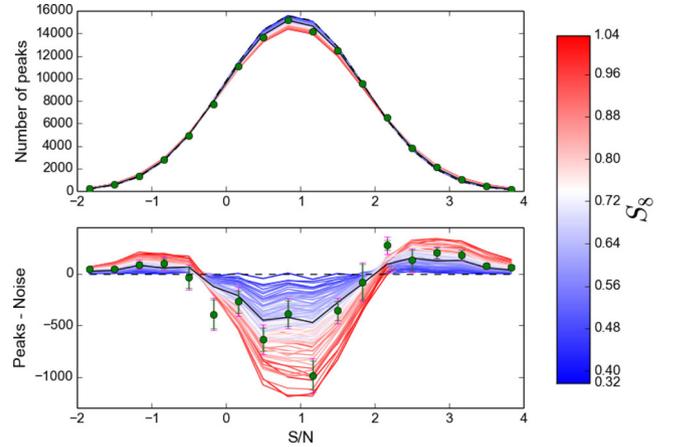
**Figure A1.** Percentage residuals from the interpolation of the number of peaks in the bin  $3.33 \leq S/N \leq 3.66$ . The black polygon represents the convex hull within which we trust the interpolation.



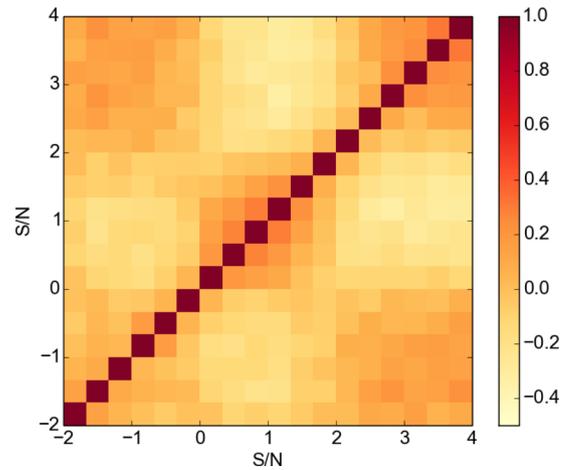
**Figure A2.** Distribution of the percentage residuals from the interpolation of the number of peaks in the bin  $3.33 \leq S/N \leq 3.66$ .

## APPENDIX B: NEGATIVE PEAKS

In this appendix, we study the possible gain of information from using the negative  $S/N$  peaks. We build the data vector with adding all negative bins that pass the Gaussian variable assumption. This gives us a new vector with 20 bins in the range  $-2 \leq S/N \leq 4$ , shown in Fig. B1. We note that the excess peak distribution over the noise peaks in the negative bins presents a similar shape as that of the positive peaks but with about twice smaller values and extending to less negative values. We then compute the constraints on  $S_8$  using the new peak distribution and obtain  $S_8 = 0.767^{+0.050}_{-0.065}$  without taking systematics into account. These constraints are of the same order as those from the positive peak distribution only ( $S_8 = 0.757^{+0.054}_{-0.053}$ ). This can be understood by looking at the correlation matrix from the full peak distribution in Fig. B2. We see that the negative peaks



**Figure B1.** Same as Fig. 4, but for peaks in the range  $-2 \leq S/N \leq 4$ .

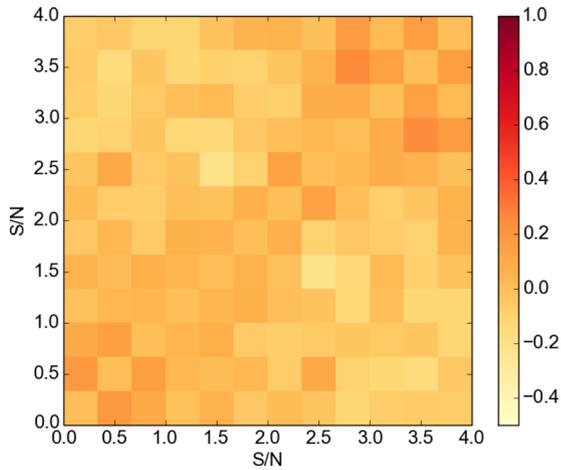


**Figure B2.** Correlation matrix for peaks in the range  $-2 \leq S/N \leq 4$ .

with  $-2 \leq S/N \leq 0$  are highly correlated with those of  $2 \leq S/N \leq 4$ , explaining why we do not gain information by including them.

## APPENDIX C: REFINED COVARIANCE MATRIX

In this section, we use the set of mocks described in Section 3.3 to measure the impact of simulation types on the derived cosmological constraints, and validate part of the analysing process. The SLICS mocks benefit from several improvements compared to those of Dietrich & Hartlap (2010) used in the rest of the paper. First, the SLICS simulations (Harnois-Déraps et al. 2015) have a higher resolution with  $1536^3$  particles against  $256^3$ , extend to  $z = 3$ , and cover  $100 \text{ deg}^2$  instead of  $36 \text{ deg}^2$ . In addition, having access to the full shear planes, we could tailor the simulations specifically for our project and populate the maps with galaxies following the position and  $n(z)$  of the KiDS-450 data, without having to rely on the interpolation scheme that we applied to the Dietrich & Hartlap (2010) mocks. We also extend the redshift range to the full DIR redshift distribution of KiDS-450, while the Dietrich & Hartlap (2010) mocks have almost no galaxies at  $z > 2$  due to the redshift distribution they used, although the simulations also extend to  $z = 3$ . Sample variance is also better included in these refined mocks by using different  $N$ -body simulations to tile the KiDS-450 area, instead of repeating a single simulation across this area.



**Figure C1.** Difference between the fiducial correlation matrix of Fig. 6 and the refined correlation matrix for peaks in the range  $0 \leq S/N \leq 4$ .

We run the same algorithm to identify peaks in these mocks, and derive the cosmological constraints using the covariance matrix from this set of simulations but still using the Dietrich & Hartlap (2010) simulations to compute the model of the peak dependence on cosmology. We use five random realizations of shape noise, the same number as in our fiducial analysis. The number of fiducial mocks is 67 which is lower than the 175 of the main analysis but each of these mocks now better accounts for sample variance. We also

note that the cosmological parameters are slightly different in the fiducial SLICS than in the Dietrich & Hartlap (2010) simulations, but we do not expect a large variation of the covariance matrix with cosmology.

In Fig. C1, we display the difference between the correlation matrix of the first set of simulations shown in Fig. 6 and the refined correlation matrix. The agreement between the two correlation matrices is good, both presenting low correlation between peaks, with somewhat higher correlations at the high-S/N peaks, leading to homogeneous residual correlation. However, the new correlation matrix shows higher scatter than the previous one, and lower correlations, leading to residual correlation of up to 0.2. This can be attributed to the larger area covered by the SLICS simulations and the proper handling of sample variance, which provides us with a more representative population of peaks. We also re-computed the Dietrich & Hartlap (2010) correlation matrix with only 67 simulations finding that the observed differences are not due to the use of different number of simulations.

As described in Section 6.3, the constraints on  $S_8$  using the refined covariance matrix are almost identical to that of the main analysis. This validates the various approximations we made when building the mocks from the Dietrich & Hartlap (2010) simulations.

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