Starlet ℓ1-norm for weak lensing cosmology

Virginia Ajani, Jean-Luc Starck, Valeria Pettorino

HAL Id: cea-03115961
https://cea.hal.science/cea-03115961

Submitted on 19 Jan 2021

Astronomy
&
Astrophysics

LETTER TO THE EDITOR

Starlet ℓ_1-norm for weak lensing cosmology

Virginia Ajani, Jean-Luc Starck, and Valeria Pettorino

AIM, CEA, CNRS, Université Paris-Saclay, Université de Paris, Sorbonne Paris Cité, 91191 Gif-sur-Yvette, France
e-mail: virginia.ajani@cea.fr

ABSTRACT

We present a new summary statistic for weak lensing observables, higher than second order, suitable for extracting non-Gaussian cosmological information and inferring cosmological parameters. We name this statistic the 'starlet ℓ_1-norm' as it is computed via the sum of the absolute values of the starlet (wavelet) decomposition coefficients of a weak lensing map. In comparison to the state-of-the-art higher-order statistics – weak lensing peak counts and minimum counts, or the combination of the two – the ℓ_1-norm provides a fast multi-scale calculation of the full void and peak distribution, avoiding the problem of defining what a peak is and what a void is: the ℓ_1-norm carries the information encoded in all pixels of the map, not just the ones in local maxima and minima. We show its potential by applying it to the weak lensing convergence maps provided by the `MassiveNus` simulations to get constraints on the sum of neutrino masses, the matter density parameter, and the amplitude of the primordial power spectrum. We find that, in an ideal setting without further systematics, the starlet ℓ_1-norm remarkably outperforms commonly used summary statistics, such as the power spectrum or the combination of peak and void counts, in terms of constraining power, representing a promising new unified framework to simultaneously account for the information encoded in peak counts and voids. We find that the starlet ℓ_1-norm outperforms the power spectrum by 72% on M_ν, 60% on Ω_m, and 75% on A_s for the *Euclid*-like setting considered; it also improves upon the state-of-the-art combination of peaks and voids for a single smoothing scale by 24% on M_ν, 50% on Ω_m, and 24% on A_s.

Key words. cosmological parameters – large-scale structure of Universe – methods: statistical – neutrinos – surveys

1. Introduction

Weak gravitational lensing by the large-scale structure represents a powerful tool for estimating cosmological parameters. Past, present, and future cosmological surveys, such as the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS) (Heymans et al. 2012), the Kilo-Degree Survey (KiDS) (Heymans et al. 2021), the Dark Energy Survey (DES) (Abbott et al. 2019; To et al. 2020), Hyper SuprimeCam (HSC) (Mandelbaum & Hyper Suprime-Cam (HSC) Collaboration 2017), *Euclid* (Laureijs et al. 2011), and the Rubin Observatory (LSST Science Collaboration 2009), consider it as one of the main physical probes for investigating unsolved questions in current cosmology, such as: what the properties of the dark components of the universe are, what the origin of its accelerated expansion is (Riess et al. 1998; Perlmutter et al. 1999), and what the sum of neutrino masses is (Lesgourgues & Pastor 2012). It is very well known that, in the context of weak lensing, second-order statistics as the two-point correlation function or its Fourier transform (the power spectrum) do not capture the non-Gaussian information encoded in the non-linear features of weak lensing data (Weinberg et al. 2013). This has motivated the introduction of several higher-order statistics, such as Minkowski functionals (Kratochvil et al. 2012; Petri et al. 2015; Vicinanza et al. 2019; Marques et al. 2019; Parroni et al. 2020), higher-order moments (Petri et al. 2016; Vicinanza et al. 2018; Peel et al. 2018; Chang et al. 2018; Gatti et al. 2020), the bispectrum (Takada & Jain 2004; Chan & Blot 2017; Coulton et al. 2019), peak counts (Kruse & Schneider 1999; Kratochvil et al. 2010; Dietrich & Hartlap 2010; Maturi et al. 2011; Pires et al. 2012; Hamana et al. 2012; Hilbert et al. 2012; Marian et al.

2012, 2013; Martinet et al. 2015, 2018; Lin & Kilbinger 2015; Giocoli et al. 2018; Peel et al. 2018; Li et al. 2019), the scattering transform (Cheng et al. 2020), wavelet phase harmonic statistics (Allys et al. 2020), and machine learning-based methods (Fluri et al. 2018a; Peel et al. 2019; Gupta et al. 2018; Rasmussen & Williams 2005; Ribli et al. 2019; Shirasaki et al. 2019), to account for non-Gaussian information in cosmological analysis. Focusing on peak counts, it has been shown that this statistic is particularly powerful in breaking degeneracy between the standard model and fifth forces in the dark sector (Peel et al. 2018) as well as in constraining cosmological parameters when employed in a multi-scale setting (Liu et al. 2015; Lin et al. 2016; Fluri et al. 2018b; Ajani et al. 2020; Zürcher et al. 2020). In particular, Ajani et al. (2020) have shown that multi-scale peak counts significantly outperform the weak lensing power spectrum, improving the constraints on the sum of neutrino masses $\sum m_\nu \equiv M_\nu$ by 63% when using a starlet filter; multi-scale peak counts were also shown to be so constraining that the addition of the power spectrum does not further improve constraints. A very interesting feature of multi-scale peaks, when they are obtained using the starlet transform (Starck et al. 2007), is the behaviour of the covariance matrix that tends to encode all information in its diagonal elements.

Another weak lensing probe of large-scale structure is represented by cosmic voids, namely under-dense regions of the large-scale matter field (Colberg et al. 2008; Pisani et al. 2019). Local minima of weak lensing convergence maps, namely pixels with values smaller than their eight neighbouring pixels, have been proposed as tracers of the matter distribution voids to infer cosmological parameters, both in a mono-scale setting (Coulton et al. 2020; Martinet et al. 2021; Davies et al. 2020)

and in a multi-scale setting (Zürcher et al. 2020). More specifically, Coulton et al. (2020) found that lensing minima alone are slightly less constraining than the peaks alone and, in agreement with Martinet et al. (2021) and Zürcher et al. (2020), that the combination of the two statistics produces significantly tighter constraints than the power spectrum.

In this Letter, we propose, for the first time, using the ℓ_1-norm of wavelet coefficients of weak lensing convergence maps. We show that it provides a unified framework for a joint multi-scale peak and void analysis and that it takes into account the information encoded in all pixels of the map.

This Letter is structured as follows: in Sect. 2, we define the ℓ_1-norm. In Sect. 3, we specify the details for the analysis, defining the summary statistics employed and the corresponding choice of filters and binning. We show how we compute the covariance matrix, show how we build the likelihood, show the posterior distributions, and define the estimators used to quantify our results. In Sect. 4, we describe our results, and we draw our conclusions in Sect. 5. We dedicate Appendix A to background definitions in weak lensing and to the details of the simulations. In Appendix B, we discuss features of covariance matrices.

2. Towards the starlet ℓ_1-norm

2.1. Starlet peaks

Multi-scale peak counts can be derived using either a set of Gaussian kernels of different sizes or a wavelet decomposition, such as the starlet transform (Ajani et al. 2020). The starlet transform decomposes a convergence map κ of size $N \times N$ into a set $\mathcal{W} = \{w_1, \ldots, w_{j_{\max}}, c_J\}$ of $J = j_{\max} + 1$ bands of the same size, where $j_{\max}$ is the number of wavelet scales considered, c_J is the coarse scale, namely a very smoothed version of the original image κ, and w_j are the wavelet bands at scale 2^j pixels. A complete description and derivation of the starlet transform algorithm can be found in Starck et al. (2007, 2016). It was also shown in Leonard et al. (2012) that the starlet transform can be interpreted as a fast multi-scale aperture mass decomposition where the aperture mass kernels have a compact support and are compensated functions (namely, the kernel integral is null). Starlet peaks are then derived by considering n bins with bin edges given by the minimum and maximum values of each band in $\mathcal{W}$. An interesting advantage of such an approach is that each wavelet band covers a different frequency range, which leads to an almost diagonal peak count covariance matrix (Ajani et al. 2020). This is not the case when a standard multi-scale Gaussian analysis is applied on the convergence map.

2.2. Starlet extrema

As mentioned in the introduction, cosmic void analysis is an alternative to peak analysis for studying convergence maps, and the combination of the two improves the constraints on cosmological parameters. It is interesting to notice how a starlet decomposition can naturally include a multi-scale void analysis. Instead of extracting maxima (peaks) in each band, we can also extract minima (pixels with values smaller than their eight neighbours), and a joint peak-void multi-scale is therefore obtained by extracting wavelet coefficient extrema (minima + maxima). The starlet decomposition therefore provides a very natural framework for a joint multi-scale peak and void analysis.

2.3. Starlet ℓ_1-norm

A particularity of peak and void statistics is that only a few pixels are considered, while other high-order statistics, such as bispectrum or Minkowski functionals, use all the pixels. In a starlet framework, we should emphasise that starlet peaks have mainly positive values and starlet voids negative values due to the property of the wavelet function. So instead of counting the number of peaks or voids in a given bin i defined by two values, B_i and B_{i+1}, we could take the sum of all wavelet coefficients with an amplitude between B_i and B_{i+1}. If B_i and B_{i+1} are positive, this corresponds to the definition of the set of coefficients $\mathcal{S}_{j,i}$ at scale j and in bin i such that $\mathcal{S}_{j,i} = \{w_{j,k} / B_i < w_{j,k} < B_{i+1}\}$, where k is the pixel index. We can then compute the sum $\sum_{u=1}^{\#\text{coef}(\mathcal{S}_{j,i})} \mathcal{S}_{j,i}[u]$. This can be generalised to positive and negative bins using:

$$l_1^{j,i} = \sum_{u=1}^{\#\text{coef}(\mathcal{S}_{j,i})} | \mathcal{S}_{j,i}[u] | = \|\mathcal{S}_{j,i}\|_1, \tag{1}$$

where $\|.\|_1$ is the standard ℓ_1-norm (i.e. $\|x\|_1 = \sum_k |x_k|$) and the index u runs from 1 to the number of pixels in a given bin i at scale j (i.e. $\#\text{coef}(\mathcal{S}_{j,i})$). The quantity $l_1^{j,i}$ defined in Eq. (1) is nothing more than the ℓ_1-norm of the binned pixel values of the starlet coefficients of the original image κ map. In the following, we will name S_{ℓ_1}, the starlet ℓ_1-norm, as the set S_{ℓ_1} of all $l_1^{j,i}$ numbers obtained from the different scales j and bins i.

This approach enables us to extract the information encoded in the absolute value of all pixels in the map instead of characterising it only by selecting local minima or maxima. An interesting advantage is that it avoids the open issue of how to define a void (Colberg et al. 2008). It is interesting to notice that this S_{ℓ_1} statistic is also directly related to the density probability function of the starlet coefficients at different scales.

3. Experiment

3.1. Summary statistics

We provided constraints on the sum of neutrino masses M_ν, on the matter density parameter Ω_m, and on the power spectrum amplitude A_s by employing five different statistics applied to the MassiveNus[1] simulations (Liu et al. 2018). Three of them are statistics that are used in weak lensing studies: the power spectrum, combined mono-scale peak and void counts, and multi-scale peak count (using the starlet transform). The two others are the starlet extrema and the starlet ℓ_1-norm S_{ℓ_1} that we introduced in the previous section.

We computed the summary statistics on maps of the signal-to-noise field, where we define the signal to noise as the ratio between the noisy convergence κ convolved with the filter $\mathcal{W}(\theta_{\text{ker}})$ over the smoothed standard deviation of the noise for each realisation per redshift:

$$S/N \equiv \frac{(\mathcal{W} * \kappa)(\theta_{\text{ker}})}{\sigma_n^{\text{filt}}}, \tag{2}$$

where $\mathcal{W}(\theta_{\text{ker}})$ can be a single-Gaussian filter or a starlet filter. Concerning σ_n^{filt}, its definition depends on the employed filter. For a Gaussian kernel, it is given by the standard deviation of the smoothed noise maps, while for the starlet case we need to estimate the noise at each wavelet scale for each image per redshift. To estimate the noise level at each starlet scale, we followed Starck & Murtagh (1998) and used the fact that the standard deviation of the noise at the scale j is given by $\sigma_j = \sigma_j^e \sigma_I$,

---

[1] Further details on the simulations are illustrated in Appendix A.

where σ_I is the standard deviation of the noise of the image and σ_j^e are the coefficients obtained by taking the standard deviation of the starlet transform of a Gaussian distribution with standard deviation one at each scale j. For the purposes of this study, we mimicked the shape noise expected for a survey such as *Euclid*[2] (Laureijs et al. 2011; Euclid Collaboration 2020) as defined in Appendix A. We performed a tomographic analysis using four source redshifts, $z_s = \{0.5, 1.0, 1.5, 2.0\}$, with corresponding values for the galaxy number density n_{gal} per source redshift bin $n_{gal} = \{11.02, 11.90, 5.45, 1.45\}$.

A: We computed the power spectrum as defined in Eq. (A.8) on noisy convergence maps filtered with a Gaussian kernel with smoothing size $\theta_{ker} = 1$ arcmin. We considered angular scales with 24 logarithmically spaced bins in the range $\ell = [300, 5000]$.

B: Regarding mono-scale peaks and voids, we computed peaks (as pixels with values larger than their eight neighbours) and voids (as pixels with values smaller than their eight neighbours) on noisy convergence maps. The maps were filtered with a single-Gaussian kernel with smoothing size $\theta_{ker} = 2$ arcmin, with 29 linearly spaced bins for peaks between the minimum and maximum of the map in S/N and 29 linearly spaced bins for voids between the negative maximum and positive minimum of the maps in S/N.

C: Starlet peak counts were computed as pixels with values larger than their eight neighbours. They were computed on noisy convergence maps filtered with a starlet kernel with four scales corresponding to [1.6′, 3.2′, 6.4′, coarse], with 29 linearly spaced bins for each scale between the minimum and maximum values of each S/N map.

D: Starlet extrema were obtained by combining the peaks computed on maps with only $S/N > 0$ contribution with starlet voids computed on maps with only $S/N < 0$ contribution. We used a starlet as we did with (C), with 58 linearly spaced bins (29 for peaks and 29 for voids).

E: The starlet ℓ_1-norm S_{ℓ_1} was computed following Eq. (1) on noisy convergence maps filtered with a starlet kernel with four scales. We followed the same process as we did for (C) and (D).

Statistics (D) and (E) are our new proposals. In all statistics where we employed the starlet decomposition, the finest frequency we considered was $\theta_{ker} = 1.6$ arcmin, corresponding to the maximum angular scale $\ell_{max} = 2149$.

3.2. Covariance matrices

The `MassiveNus` suite of simulations (described in more detail in Appendix A) also includes a cosmology with massless neutrinos $\{M_\nu, \Omega_m, 10^9 A_s\} = \{0.0, 0.3, 2.1\}$ obtained from initial conditions different from those of the massive neutrino simulations. We used this additional simulation set to compute the covariance matrix of the observable. The covariance matrix elements are computed as

$$C_{ij} = \sum_{r=1}^{N} \frac{(x_i^r - \mu_i)(x_j^r - \mu_j)}{N - 1}, \quad (3)$$

where N is the number of observations (in this case, the 10 000 realisations), x_i^r is the value of the considered observable in the ith bin for a given realisation r, and

$$\mu_i = \frac{1}{N} \sum_r x_i^r \quad (4)$$

is the mean of the observable in a given bin over all the realisations. Furthermore, we took into account the loss of information due to the finite number of bins and realisations by adopting the estimator introduced by Hartlap et al. (2007) for the inverse of the covariance matrix:

$$C^{-1} = \frac{N - n_{bins} - 2}{N - 1} C_*^{-1}, \quad (5)$$

where N is the number of realisations, n_{bins} the number of bins, and C_* the covariance matrix computed for the power spectrum and peak counts, whose elements are given by Eq. (3). For illustration, we scaled the covariance for an *Euclid*[3] sky coverage by the factor f_{map}/f_{survey}, where $f_{map} = 12.25 \deg^2$ is the size of the convergence maps and $f_{Euclid} = 15\,000 \deg^2$.

3.3. Likelihood

To perform Bayesian inference and get the probability distributions of the cosmological parameters, we used a Gaussian likelihood for a cosmology-independent covariance:

$$\log \mathcal{L}(\boldsymbol{\theta}) = \frac{1}{2}(d - \mu(\boldsymbol{\theta}))^T C^{-1}(d - \mu(\boldsymbol{\theta})), \quad (6)$$

where d is the data array, C is the covariance matrix of the observable, and μ is the expected theoretical prediction as a function of the cosmological parameters θ. In our case, the data array is the mean over the (simulated) realisations of all the separate observables, or, of their different combinations, for our fiducial model. Cosmological parameters are the ones for which simulations are available, namely $\{M_\nu, \Omega_m, 10^9 A_s\}$, but the same statistics can of course be applied to other parameters if simulations allow.

In order to predict the theoretical values of all observables given a new set of values for the cosmological parameters $\{M_\nu, \Omega_m, 10^9 A_s\}$, we employed an interpolation with Gaussian process regression (Rasmussen & Williams 2005) using the `scikit-learn` python package. The emulator used in this Letter is the same as that employed in Li et al. (2019).

3.4. Result estimators

To quantify the constraining power of the different summary statistics, we employed the following estimators, which are based on Euclid Collaboration (2020). To have an approximate quantification of the size of the parameter contours, we considered the following figure of merit (FoM):

$$\text{FoM} = \left(\det \left(\tilde{F} \right) \right)^{1/n}, \quad (7)$$

where $\tilde{F}$ is the marginalised Fisher sub-matrix that we estimated as the inverse of the covariance matrix among the set of cosmological parameters under investigation and n is equal to the parameter space dimensionality. We show the values of the FoM for our observables in Table 1. To estimate the marginalised 1σ error on a single parameter θ_α (which means having included all the degeneracies with respect to other parameters), we used the quantity:

$$\sigma_{\alpha\alpha} = \sqrt{C_{\alpha\alpha}}, \quad (8)$$

where $C_{\alpha\alpha}$ are the diagonal elements of the parameter covariance matrix. We show the values of the $\sigma_{\alpha\alpha}$ for our observables in Table 2.

Table 1. Values of the figure of merit (FoM) as defined in Eq. (7) for different parameter pairs for each observable employed in the likelihood analysis.

FoM	(M_ν, Ω_m)	(M_ν, A_s)	(Ω_m, A_s)	(M_ν, Ω_m, A_s)
Power spectrum	1585	77	1079	2063
Starlet peaks	5904	331	4856	10 147
ℓ_1-norm	11 408	619	9126	16 688
Starlet extrema	6967	442	5956	6740
Peaks + voids monoscale	5321	307	4286	6114

Notes. In the last column, we provide the 3D FoM given as the inverse of the volume in (M_ν, A_s, Ω_m) space.

Table 2. Values of 1-σ marginalised errors as defined in Eq. (8) for each cosmological parameter for the different observables.

$\sigma_{\alpha\alpha}$	M_ν	Ω_m	A_s
Power spectrum	0.147	0.005	0.204
Starlet peaks	0.057	0.003	0.064
ℓ_1-norm	0.041	0.002	0.051
Starlet extrema	0.040	0.004	0.062
Peaks + voids monoscale	0.054	0.004	0.067

3.5. MCMC simulations and posterior distributions

We explored and constrained the parameter space with the emcee package, which is a python implementation of the affine invariant ensemble sampler for Markov chain Monte Carlo (MCMC) introduced by Foreman-Mackey et al. (2013). We assumed a flat prior, a Gaussian likelihood function as defined in Sect. 3.3, and a model-independent covariance matrix as discussed in Sect. 3.2. The walkers were initialised in a tiny Gaussian ball of radius 10^{-3} around the fiducial cosmology $[M_\nu, \Omega_m, 10^9 A_s] = [0.1, 0.3, 2.1]$, and we estimated the posterior using 120 walkers. Our chains are stable against the length of the chain, and we verified their convergence by employing Gelman Rubin diagnostics (Gelman & Rubin 1992). To plot the contours, we used the ChainConsumer python package (Hinton 2016).

4. Results

We now illustrate forecast results on the sum of neutrino masses M_ν, on the matter density parameter Ω_m, and on the power spectrum amplitude A_s for a survey with *Euclid*-like noise in a tomographic setting with four source redshifts, $z_s = [0.5, 1.0, 1.5, 2.0]$. We compare results for the different observables defined in Sect. 3.1 and investigate the impact of the choice of the filter on the covariance matrix.

In Fig. 1, we show the comparison between the constraints obtained using different summary statistics, as described in Sect. 3.1. As expected, we see that all higher-order statistics are more constraining than the power spectrum. The new result of this Letter is represented by the starlet ℓ_1-norm: the inclusion of all pixels enables us to retrieve tighter constraints than the combination of local minima and maxima (voids + peaks). Specifically, for all parameter space planes, the ℓ_1-norm FoM values, illustrated in Table 1, are about twice as large compared to those for the combined local minima and maxima (and more than seven times larger than the power spectrum FoM values).

In this work, we have also introduced starlet extrema as a new summary statistic to constrain the parameters. Similarly to the ℓ_1-norm, starlet extrema are computed between the minimum

and maximum S/N value of the map, but they are defined as the combination of local maxima computed on $S/N > 0$ and local minima computed on $S/N < 0$ (namely, they do not encode the information present in all pixels). We see that starlet extremum FoM values are larger than starlet peaks and peaks + voids monoscales, suggesting that starlet extrema can be a good multi-scale higher-order statistic candidate. However, the ℓ_1-norm remains the statistic that performs the best in terms of constraining power with respect to all the summary statistics we have considered. In Fig. 2, we show the marginalised constraints on each cosmological parameter corresponding to the different observables. To compare the improvement obtained by employing the different statistics, we computed the 1σ marginalised error for each parameter, as summarised in Table 2. We find that the starlet ℓ_1-norm outperforms the power spectrum by 72% on M_ν, 60% on Ω_m, and 75% on A_s, and the state-of-the-art peaks + voids for a single smoothing scale by 24% on M_ν, 50% on Ω_m, and 24% on A_s. Starlet extrema outperform the power spectrum by 72% on M_ν, 20% on Ω_m, and 70% on A_s. We also quantify the improvement provided by the ℓ_1-norm with respect to our previous study (Ajani et al. 2020), finding that the starlet ℓ_1-norm outperforms starlet peaks by 28% on M_ν, 33% on Ω_m, and 20% on A_s. In Appendix B, we also compare the covariance matrices obtained when using starlet extrema and the ℓ_1-norm, and we find that starlet extrema present a more diagonal covariance for the observable.

5. Conclusions

In this Letter, we have proposed using starlet ℓ_1-norm statistics on weak lensing converge maps to constrain cosmological parameters. The measure of multi-scale peak amplitudes can be seen as a measure of the ℓ_1-norm of a subset of positive wavelet coefficients. Similarly, the measure of void amplitudes can be seen as a measure of the ℓ_1-norm of a subset of negative wavelet coefficients. Wavelets therefore provide a great framework for a joint peak and void analysis, in which information from all wavelet coefficients is included. We therefore propose using a very simple ℓ_1-norm statistic, defined as the sum of the ℓ_1-norm of all coefficients in a given S/N (pixel) bin for each wavelet scale, as defined in Eq. (1). We investigate the impact of employing the starlet ℓ_1-norm as summary statistics computed on weak lensing convergence maps to estimate cosmological parameters, and we find that the ℓ_1-norm outperforms the two state-of-the-art summary statistics, the power spectrum and the combination of mono-scale peaks and voids, by, respectively, 72% and 24% on M_ν, 60% and 50% on Ω_m, and 75% and 24% on A_s. We have furthered proposed starlet extrema and compared them to the ℓ_1-norm: in this case as well, the latter performs better in terms of constraining power, within the current ideal setting, while the former present the advantage of a more diagonal covariance
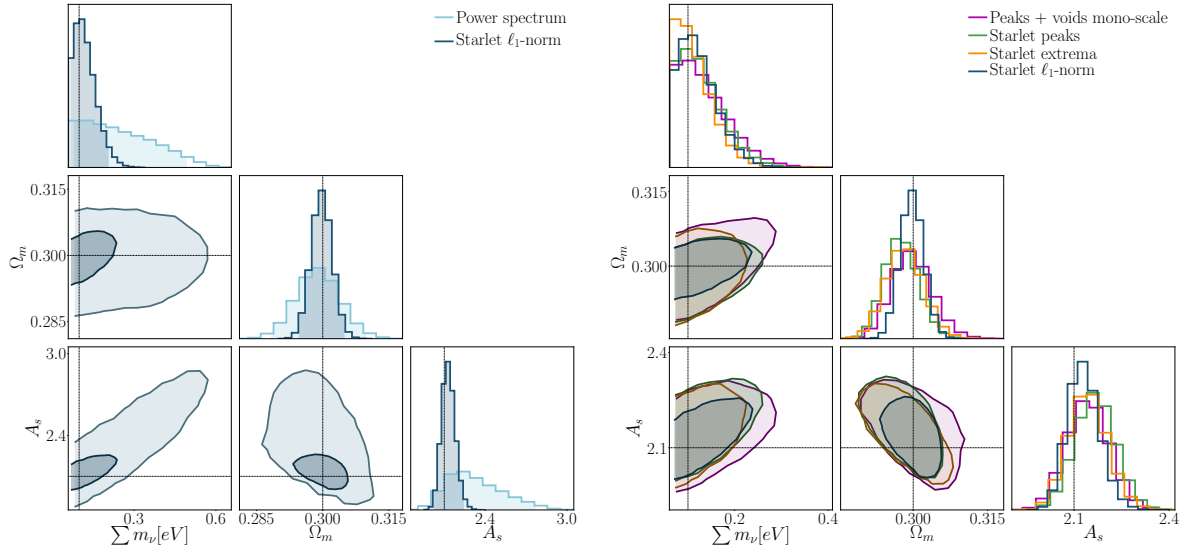
Fig. 1. 95% confidence contour tomography with source redshifts z_s = [0.5, 1.0, 1.5, 2.0] and corresponding galaxy number densities n_gal = [11.02, 11.90, 5.45, 1.45]. The dotted black line is the fiducial model: [$\sum m_\nu$, Ω_m, $10^9 A_\mathrm{s}$] = [0.1, 0.3, 2.1]. *Left panel*: constraints from the power spectrum (light blue contours) computed on noisy maps smoothed with a Gaussian filter with θ_ker = 1 arcmin, compared to constraints from the starlet ℓ_1-norm (dark blue contours) computed on noisy maps filtered with a four-scale starlet kernel. *Right panel*: constraints from the combination of peaks and voids (magenta contours) computed on noisy maps smoothed with a Gaussian filter with θ_ker = 2 arcmin compared to constraints from starlet peak counts (green contours), starlet extrema (orange contours), and the ℓ_1-norm (dark blue contours), computed on noisy maps filtered with a four-scale starlet kernel.
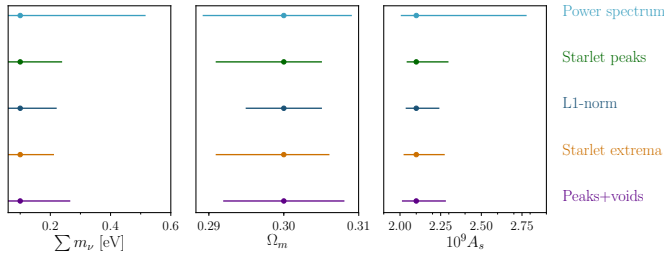

Fig. 2. Marginalised errors for the observables described in Sect. 3.1 on each parameter, showing the 2.5 and 97.5 percentiles with respect to the fiducial model. They refer to a tomographic setting with z = [0.5, 1.0, 1.5, 2.0], with the fiducial model set to [M_ν, Ω_m, $10^9 A_\mathrm{s}$] = [0.1, 0.3, 2.1]. The last observable refers to mono-scale peaks and voids, as described in the text.

matrix. We are aware that the statistical power alone is not sufficient to serve as a robust probe for precision cosmology; for their usage, it will be important to test how these statistics react in a non-ideal setting and how their performance is impacted by systematics in the signal. We will dedicate a future study to this aspect.

We conclude that the new statistic proposed here presents several advantages in the context of cosmological parameter inference: it provides a fast calculation of the full void and peak distribution; it does not rely on a specific definition of peaks and voids or some arbitrary threshold; it instead encodes information of the entire pixel distribution without excluding pixels that are not local minima or maxima; and it leads to tighter constraints, at least within an ideal setting. The starlet decomposition therefore provides a very powerful framework for a joint multi-scale peak and void analysis.

References

Abbott, T. M. C., Abdalla, F. B., Avila, S., et al. 2019, Phys. Rev. D, 99, 123505
Ajani, V., Peel, A., Pettorino, V., et al. 2020, Phys. Rev. D, 102, 103531
Allys, E., Marchand, T., Cardoso, J.-F., et al. 2020, Phys. Rev. D, 102, 103506
Chan, K. C., & Blot, L. 2017, Phys. Rev. D, 96, 023528
Chang, C., Pujol, A., Mawdsley, B., et al. 2018, MNRAS, 475, 3165
Cheng, S., Ting, Y.-S., Ménard, B., & Bruna, J. 2020, MNRAS, 499, 5902
Colberg, J. M., Pearce, F., Foster, C., et al. 2008, MNRAS, 387, 933
Coulton, W. R., Liu, J., Madhavacheril, M. S., et al. 2019, JCAP, 2019, 043
Coulton, W. R., Liu, J., McCarthy, I. G., et al. 2020, MNRAS, 495, 2531
Davies, C. T., Cautun, M., Giblin, B., et al. 2020, MNRAS, submitted [arXiv:2010.11954]
Dietrich, J. P., & Hartlap, J. 2010, MNRAS, 402, 1049
Euclid Collaboration (Blanchard, A., et al.) 2020, A&A, 642, A191
Fluri, J., Kacprzak, T., Refregier, A., et al. 2018a, Phys. Rev. D, 98, 123518
Fluri, J., Kacprzak, T., Sgier, R., et al. 2018b, JCAP, 2018, 051
Foreman-Mackey, D., Hogg, D. W., Lang, D., et al. 2013, PASP, 125, 306
Gatti, M., Chang, C., Friedrich, O., et al. 2020, MNRAS, 498, 4060
Gelman, A., & Rubin, D. B. 1992, Stat. Sci., 7, 457
Giocoli, C., Moscardini, L., Baldi, M., et al. 2018, MNRAS, 478, 5436
Gupta, A., Matilla, J. M. Z., Hsu, D., et al. 2018, Phys. Rev. D, 97, 103515
Hamana, T., Oguri, M., Shirasaki, M., & Sato, M. 2012, MNRAS, 425, 2287
Hartlap, J., Simon, P., & Schneider, P. 2007, A&A, 464, 399
Heymans, C., Van Waerbeke, L., Miller, L., et al. 2012, MNRAS, 427, 146
Heymans, C., Tröster, T., Asgari, M., et al. 2021, A&A, in press, https://doi.org/10.1051/0004-6361/202039063
Hilbert, S., Marian, L., Smith, R. E., & Desjacques, V. 2012, MNRAS, 426, 2870
Hinton, S. R. 2016, J. Open Sour. Softw., 1, 00045
Kaiser, N., & Squires, G. 1993, ApJ, 404, 441
Kilbinger, M. 2015, Rep. Progr. Phys., 78, 086901
Kratochvil, J. M., Haiman, Z., & May, M. 2010, Phys. Rev. D, 81, 043519
Kratochvil, J. M., Lim, E. A., Wang, S., et al. 2012, Phys. Rev. D, 85, 103513
Kruse, G., & Schneider, P. 1999, MNRAS, 302, 821

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv e-prints [arXiv:1110.3193]

Leonard, A., Pires, S., & Starck, J.-L. 2012, MNRAS, 423, 3405

Lesgourgues, J., & Pastor, S. 2012, Adv. High Energy Phys., 2012, 608515

Li, Z., Liu, J., Zorrilla Matilla, J. M., et al. 2019, Phys. Rev. D, 99, 063527

Lin, C.-A., & Kilbinger, M. 2015, A&A, 583, A70

Lin, C.-A., Kilbinger, M., & Pires, S. 2016, A&A, 593, A88

Liu, J., Petri, A., Haiman, Z., et al. 2015, Phys. Rev. D, 91, 063507

Liu, J., Bird, S., Zorrilla Matilla, J. M., et al. 2018, JCAP, 2018, 049

LSST Science Collaboration (Abell, P. A., et al.) 2009, ArXiv e-prints [arXiv:0912.0201]

Mandelbaum, R., & Hyper Suprime-Cam (HSC) Collaboration 2017, Am. Astron. Soci. Meet. Abstr., 229, 226.02

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

Marques, G. A., Liu, J., Zorrilla Matilla, J. M., et al. 2019, JCAP, 2019, 019

Martinet, N., Bartlett, J. G., Kiessling, A., & Sartoris, B. 2015, A&A, 581, A101

Martinet, N., Schneider, P., Hildebrandt, H., et al. 2018, MNRAS, 474, 712

Martinet, N., Harnois-Déraps, J., Jullo, E., & Schneider, P. 2021, A&A, in press, https://doi.org/10.1051/0004-6361/202039679

Maturi, M., Fedeli, C., & Moscardini, L. 2011, MNRAS, 416, 2527

Parroni, C., Cardone, V. F., Maoli, R., & Scaramella, R. 2020, A&A, 633, A71

Peel, A., Pettorino, V., Giocoli, C., et al. 2018, A&A, 619, A38

Peel, A., Lalande, F., Starck, J.-L., et al. 2019, Phys. Rev. D, 100, 023508

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565

Petri, A. 2016, Astron. Comput., 17, 73

Petri, A., Liu, J., Haiman, Z., et al. 2015, Phys. Rev. D, 91, 103511

Petri, A., May, M., & Haiman, Z. 2016, Phys. Rev. D, 94, 063534

Pires, S., Leonard, A., & Starck, J.-L. 2012, MNRAS, 423, 983

Pisani, A., Massara, E., Spergel, D. N., et al. 2019, BAAS, 51, 40

Rasmussen, C. E., & Williams, C. K. I. 2005, Gaussian Processes for Machine Learning, Adaptive Computation and Machine Learning (The MIT Press)

Ribli, D., Pataki, B. Á., Zorrilla Matilla, J. M., et al. 2019, MNRAS, 490, 1843

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009

Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses (Berlin, Heidelberg: Springer-Verlag)

Shirasaki, M., Yoshida, N., Ikeda, S., Oogi, T., & Nishimichi, T. 2019, ArXiv e-prints [arXiv:1911.12890]

Starck, J.-L., & Murtagh, F. 1998, PASP, 110, 193

Starck, J. L., Fadili, J., & Murtagh, F. 2007, IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society, 16, 297

Starck, J.-L., Murtagh, F., & Fadili, J. 2016, Sparse Image and Signal Processing: Wavelets and Related Geometric Multiscale Analysis (Cambridge: Cambridge University Press)

Takada, M., & Jain, B. 2004, MNRAS, 348, 897

To, C., Krause, E., Rozo, E., et al. 2020, PRL, submitted [arXiv:2010.01138]

Vicinanza, M., Cardone, V. F., Maoli, R., et al. 2018, Phys. Rev. D, 97, 023519

Vicinanza, M., Cardone, V. F., Maoli, R., et al. 2019, Phys. Rev. D, 99, 043534

Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. 2013, Phys. Rep., 530, 87

Zürcher, D., Fluri, J., Sgier, R., et al. 2020, ArXiv e-prints [arXiv:2006.12506]

Appendix A: Background

A.1. Weak lensing

The effect of gravitational lensing at co-moving angular distance $f_K(\chi)$ can be described by the lensing potential,

$$\psi(\boldsymbol{\theta}, \chi) \equiv \frac{2}{c^2} \int_0^{\chi} \mathrm{d}\chi' \frac{f_K(\chi - \chi')}{f_K(\chi) f_K(\chi')} \Phi(f_K(\chi')\boldsymbol{\theta}, \chi'), \tag{A.1}$$

which defines how much the gravitational potential Φ arising from a mass distribution changes the direction of a light path. In this expression, K is the spatial curvature constant of the universe, χ is the co-moving radial coordinate, $\boldsymbol{\theta}$ is the angle of observation, and c is the speed of light. As we are working under the assumption of a Lambda cold dark matter (ΛCDM) model, the two Bardeen gravitational potentials are here assumed to be equal and the metric signature is defined as $(+1, -1, -1, -1)$. In particular, under the Born approximation, the effect of the lensing potential on the shapes of background galaxies in the weak regime can be summarised by its variation with respect to $\boldsymbol{\theta}$. Formally, this effect can be described by the elements of the lensing potential Jacobi matrix,

$$A_{ij} = \delta_{ij} - \partial_i \partial_j \psi, \tag{A.2}$$

which can be parametrised as

$$A = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}, \tag{A.3}$$

where (γ_1, γ_2) are the components of a spin-2 field γ called shear and κ is a scalar quantity called convergence. They describe, respectively, the anisotropic stretching and the isotropic magnification of the galaxy shape when light passes through large-scale structure. Equations (A.2) and (A.3) define the shear and the convergence fields as second-order derivatives of the lensing potential:

$$\gamma_1 \equiv \frac{1}{2}(\partial_1\partial_1 - \partial_2\partial_2)\psi \quad \gamma_2 \equiv \partial_1\partial_2\psi, \tag{A.4}$$

$$\kappa \equiv \frac{1}{2}(\partial_1\partial_1 + \partial_2\partial_2)\psi = \frac{1}{2}\nabla^2\psi. \tag{A.5}$$

The weak lensing field is a powerful tool for cosmological inference. The shear is more closely related to actual observables (galaxy shapes), while the convergence, as a scalar field, can be more directly understood in terms of the matter density distribution along the line of sight. This can be seen by inserting the lensing potential defined in Eq. (A.1) inside Eq. (A.5) and using the fact that the gravitational potential Φ is related to the matter density contrast $\delta = \Delta\rho/\bar{\rho}$ through the Poisson equation $\nabla^2\Phi = 4\pi G a^2 \bar{\rho}\delta$. Expressing the mean matter density in terms of the critical density $\rho_{c,0} = 3H_0^2/(8\pi G)$, the convergence field can be rewritten as

$$\kappa(\boldsymbol{\theta}) = \frac{3H_0^2\Omega_m}{2c^2} \int_0^{\chi_{\lim}} \frac{\mathrm{d}\chi}{a(\chi)} g(\chi) f_K(\chi) \delta(f_K(\chi)\boldsymbol{\theta}, \chi), \tag{A.6}$$

where H_0 is the Hubble parameter at its present value and

$$g(\chi) \equiv \int_{\chi}^{\chi_{\lim}} \mathrm{d}\chi' n(\chi') \frac{f_K(\chi' - \chi)}{f_K(\chi')} \tag{A.7}$$

is the lensing efficiency. Equation (A.6) relates the convergence κ to the 3D matter over-density field $\delta(f_K(\chi)\boldsymbol{\theta}, \chi)$, and it describes how the lensing effect on the matter density distribution is quantified by the lensing strength at a distance χ, which directly depends on the normalised source galaxy distribution $n(z)\mathrm{d}z = n(\chi)\mathrm{d}\chi$ and on the geometry of the universe through $f_K(\chi)$ along the line of sight. A complete derivation can be found in Kilbinger (2015) and Schneider et al. (1992).

Convergence power spectrum

To provide a statistical estimate of the distribution of the convergence field, the first non-zero order is given by its second moment, which is commonly described by the 'two-point correlation function' (2PCF) in real space $\langle\kappa(\theta)\kappa(\theta')\rangle$, or by its counterpart in Fourier space, the 'convergence power spectrum':

$$C_\kappa(\ell) = \frac{9\Omega_m^2 H_0^4}{4c^4} \int_0^{\chi_{\lim}} \mathrm{d}\chi \frac{g^2(\chi)}{a^2(\chi)} P_\delta\left(\frac{\ell}{f_K(\chi)}, \chi\right), \tag{A.8}$$

where P_δ represents the 3D matter power spectrum, directly related to the matter density distribution δ in Eq. (A.6), of the weak lensing convergence field. In this study, we computed the power spectrum of the noisy filtered convergence maps: for a given cosmology, we added Gaussian noise to each realisation of κ. To filter the maps, we employed a Gaussian kernel with smoothing size $\theta_{\mathrm{ker}} = 1$ arcmin and considered angular scales with logarithmically spaced bins in the range $\ell = [300, 5000]$. We computed the power spectra using LensTools for each of the $10\,000$ realisations per cosmology, and then we took the average over the realisations. We parallelised our code using joblib[4] to accelerate processing due to the large number of realisations per cosmology.

A.2. Simulations

In this Letter, we used the Cosmological Massive Neutrino Simulations (MassiveNus), a suite of publicly available N-body simulations released by the Columbia Lensing group[5]. It contains 101 different cosmological models obtained by varying the sum of neutrino masses M_ν, the total matter density parameter Ω_m, and the primordial power spectrum amplitude A_s at the pivot scale $k_0 = 0.05\,\mathrm{Mpc}^{-1}$ in the ranges $M_\nu = [0, 0.62]$ eV, $\Omega_m = [0.18, 0.42]$, and $A_s \times 10^9 = [1.29, 2.91]$. The reduced Hubble constant $h = 0.7$, the spectral index $n_s = 0.97$, the baryon density parameter $\Omega_b = 0.046$, and the dark energy equation of state parameter $w = -1$ were kept fixed under the assumption of a flat universe. The fiducial model was set at $\{M_\nu, \Omega_m, 10^9 A_s\} = \{0.1, 0.3, 2.1\}$. The complete description of the implementation and the products is illustrated in Liu et al. (2018). We used simulated convergence maps as mock data for our analysis. When dealing with real data, the actual observable was the shear field, which can be converted into the convergence field following Kaiser & Squires (1993). We bypassed this step from γ to κ and worked with the convergence maps that were directly provided as products from MassiveNus. The maps were generated using the public ray-tracing package LensTools (Petri 2016) for each of the 101 cosmological models at five source redshifts, $z_s = \{0.5, 1.0, 1.5, 2.0, 2.5\}$. Each redshift has $10\,000$ different map realisations obtained by rotating and shifting the spatial planes. Each κ map has 512^2 pixels, corresponding to a $12.25\,\mathrm{deg}^2$ total angular size area in the range $\ell \in [100, 37\,000]$ with a resolution of 0.4 arcmin. We mimicked *Euclid*-like shape

---

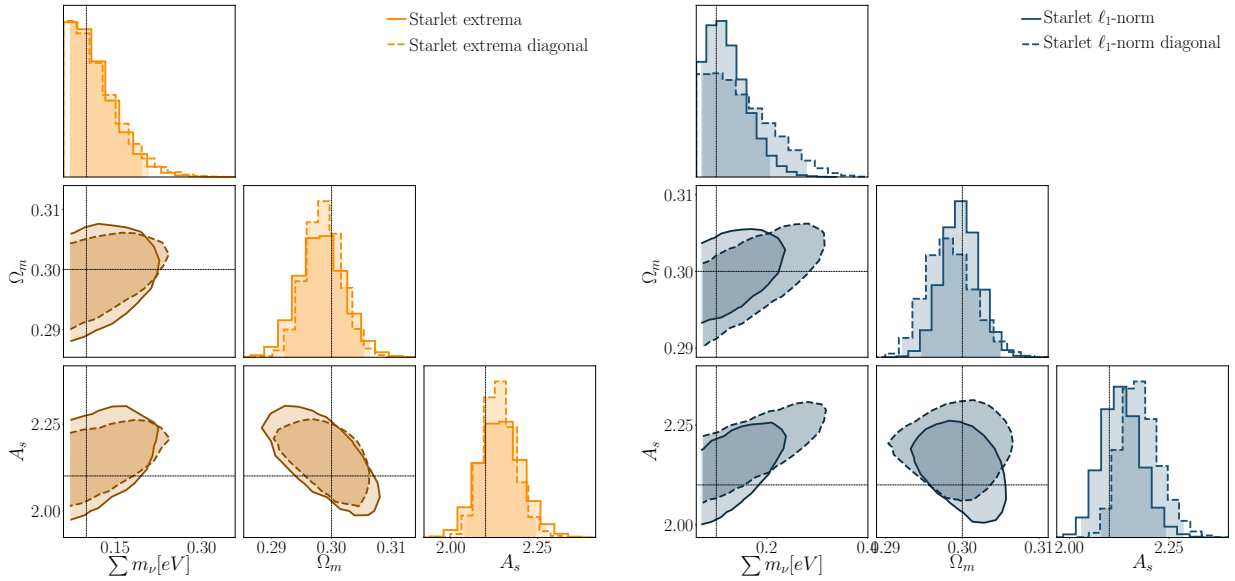4 https://joblib.readthedocs.io/
5 http://columbialensing.org

Fig. A.1. 95% confidence contour tomography with source redshifts $z_s = [0.5, 1.0, 1.5, 2.0]$ and corresponding galaxy number densities: $n_{gal} = [11.02, 11.90, 5.45, 1.45]$. The dotted black line is the fiducial model: $[\sum m_\nu, \Omega_m, 10^9 A_s] = [0.1, 0.3, 2.1]$. *Left panel*: constraints from starlet extrema with the full covariance matrix (continuous contours) computed on noisy maps filtered with a four-scale starlet kernel against constraints from starlet extrema with the only diagonal elements of the covariance matrix (dashed contours). *Right panel*: constraints from the ℓ_1-norm with the full covariance matrix (continuous contours) computed on noisy maps filtered with a four-scale starlet kernel against constraints from ℓ_1-norm with the only diagonal elements of the covariance matrix (dashed contours).

noise at each source redshift assuming Gaussian noise with mean zero and variance:

$$\sigma_n^2 = \frac{\langle \sigma_\epsilon^2 \rangle}{n_{gal} A_{pix}}, \tag{A.9}$$

where we set the dispersion of the ellipticity distribution to $\sigma_\epsilon = 0.3$ and the pixel area is given by $A_{pix} \simeq 0.16 \, \text{arcmin}^2$.

Appendix B: Diagonal covariances

As recalled in the introduction, Ajani et al. (2020) found that starlet peak counts have the interesting feature that the corresponding covariance matrix is nearly diagonal. Motivated by this, we have tested in this work whether this feature is also maintained for starlet extrema and for the ℓ_1-norm. Interestingly, we find that starlet extrema keep this characteristic, while the ℓ_1-norm show more correlations in the off-diagonal term. This can be seen by looking at Fig. A.1: in the left panel, we show the constraints for the starlet extrema when using the full covariance matrix (continuous contours) and compare them with

the starlet extrema when using only the diagonal elements of the covariance matrix (dashed contours). Analogously, we show the same comparison for the ℓ_1-norm in the right panel. When using starlet extrema as summary statistics, there is no loss of information on M_ν, and a slight loss of information of 25% on Ω_m and 22% A_s if we employ a covariance matrix with only its diagonal elements in the likelihood analysis. Concerning the ℓ_1-norm, it is sufficient to look at the contours to notice how the contours with only diagonal terms are considerably different with respect to the contours obtained with the full covariance matrix: they appear shifted and present a different degree of degeneracy for (Ω_m, A_s). Hence, considering this result along with the result from Sect. 4, we conclude that the ℓ_1-norm outperforms starlet extrema in terms of constraining power when considering the full covariance in the analysis, but presents a less diagonal matrix than starlet extrema. We conclude that, depending on the context, the ℓ_1-norm could be a convenient choice when the priority is the constraining power, while starlet extrema might be more useful when one is interested in speeding up the analysis or when the covariance matrix can be difficult to invert.