Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **479,** 454–477 (2018)
Advance Access publication 2018 May 30

doi:10.1093/mnras/sty1412

Flat-Sky Pseudo-Cls analysis for weak gravitational lensing

Marika Asgari,[1]★ Andy Taylor,[1] Benjamin Joachimi[2] and Thomas D. Kitching[3]

[1]*SUPA, Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*
[2]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK*
[3]*Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK*

Accepted 2018 May 29. Received 2018 May 29; in original form 2016 December 14

ABSTRACT

We investigate the use of estimators of weak lensing power spectra based on a flat-sky implementation of the 'Pseudo-Cl' (PCl) technique, where the masked shear field is transformed without regard for masked regions of sky. This masking mixes power and 'E'-convergence and 'B'-modes. To study the accuracy of forward-modelling and full-sky power spectrum recovery, we consider both large-area survey geometries and small-scale masking due to stars and a checkerboard model for field-of-view gaps. The power spectrum for the large-area survey geometry is sparsely sampled and highly oscillatory, which makes modelling problematic. Instead, we derive an overall calibration for large-area mask bias using simulated fields. The effects of small-area star masks can be accurately corrected for, while the checkerboard mask has oscillatory and spiky behaviour that leads to per cent biases. Apodization of the masked fields leads to increased biases and a loss of information. We find that we can construct an unbiased forward model of the raw PCls, and recover the full-sky convergence power to within a few per cent accuracy for both Gaussian and lognormal-distributed shear fields. Propagating this through to cosmological parameters using a Fisher-Matrix formalism, we find that we can make unbiased estimates of parameters for surveys up to 1200 deg^2 with 30 galaxies per arcmin^2, beyond which the per cent biases become larger than the statistical accuracy. This implies that a flat-sky PCl analysis is accurate for current surveys but a Euclid-like survey will require higher accuracy.

Key words: Gravitational lensing: weak – methods: numerical – cosmology: observations.

1 INTRODUCTION

Weak gravitational lensing, the distortion of distant galaxy images by the intervening matter, provides us with a unique probe of the mass distribution over a large range of scales in the Universe and so is sensitive to the properties of the dark matter and dark energy (for a comprehensive review of weak lensing, see Bartelmann & Schneider 2001, and for a recent review of cosmic shear, see Kilbinger 2015). In addition, it is sensitive to temporal and spatial distortions of space–time and hence can be used as a probe of gravity. One of the primary concerns of cosmology is the comparison of dark matter, dark energy, and modified gravity models and the estimation of their parameters. Cosmological model comparison and parameter estimation are usually carried out by compressing the data into a form that can be most easily compared with the models. If the data are Gaussian distributed, all of the relevant information is contained in the 2-point statistics of the data. However, non-linear evolution of the density field generates higher order correlations, which makes 2-point statistics insufficient for capturing the entire information content of the field. Nevertheless, much of the information is still retained by the 2-point statistics and a lot of effort in cosmology has gone into optimally extracting this information.

The 2-point correlations of the weak lensing shear field can be estimated directly from the shear signal on the sky or from the harmonic modes from a transformation of the data. Many cosmic shear studies have focused on the real-space shear 2-point correlation functions (2PCFs), since this statistic is not biased by the sampling of the shear field from the lensed source galaxy images, for the scales that are available in the survey. However, points in the 2PCFs are correlated in a way that does depend on the galaxy sampling including the mask and survey geometry (see Kilbinger & Schneider 2004). In addition, the 2PCFs mix linear and non-linear scales of the shear field. These non-linear scales can be more difficult to model, in particular, due to the presence of baryons that affect the evolution of structure. The cosmic shear field can also be decomposed into a convergence (even-parity or E-mode) field that is generated by the matter-density field and a divergence-free (odd-parity or B-mode) field that is mainly generated by ellipticity noise and systematics for current surveys. The 2PCFs mix these two modes; however, a full separation can be achieved using Complete

★ E-mail: ma@roe.ac.uk

Orthogonal Sets of E/B-Integrals (COSEBIs) (Schneider, Eifler & Krause 2010) that can also be used to restrict the range of scales (see Kilbinger et al. 2013; Huff et al. 2014; Asgari et al. 2017, for cosmic shear with COSEBIs).

In harmonic space, the 2-point observable is the shear power spectrum that, by definition, must be positive semi-definite for auto-spectra. The spherical harmonic modes on the celestial sphere are uncorrelated due to rotational invariance and homogeneity; moreover, the covariance of the linear shear power spectrum on the full-sky is diagonal for Gaussian perturbations before masking effects. However, for 2-point statistics, the scales with the most cosmological information are in the non-linear régime where the shear powers are correlated. In order to model accurately these non-linear regimes, we need simulations, which are accurate over a finite range in Fourier space, and so a harmonic analysis is well matched to simulated modelling. The decomposition into *E*- and *B*-modes is also straightforward in harmonic space on the full sky.

The main drawback of a harmonic analysis of cosmic shear is the effect of the source galaxy sampling. As the shear field is sampled by the background source galaxies, a shear map is defined by the position of the source galaxies. However, for data analysis it is more convenient to bin the shear data for a harmonic analysis and define a shear mask on the pixelated field where no source galaxies are detected. Stellar images may also contaminate nearby galaxy images and so these galaxies must be excluded. On the sky, the mask multiplies the shear field, and so in harmonic space it is a convolution. The convolution will correlate different scales and will bias the shear power and covariance unless it is accounted for.

This problem is well known in cosmic microwave background (CMB) analysis, where a spherical harmonic analysis is standard. One common, and fast, direct measurement of the CMB power spectrum is carried out by the Pseudo-Cl (PCl) analysis (Hivon et al. 2002). This method can rapidly analyse masked data in spherical harmonic space and has been used to analyse CMB temperature (see Planck Collaboration XVI 2014, for a full-sky analysis of the CMB temperature) and polarization data (see Brown, Castro & Taylor 2005; Brown et al. 2009, and references therein, for a full-sky analysis of simulated and QUaD (Q and U Extragalactic Sub-mm Telescope at Degree Angular Scale Interferometer) survey data). While most CMB analysis takes place on the full, curved sky where the spherical harmonic decomposition is well-defined, some studies have used a flat-sky analysis (see e.g. Memari 2009, for a flat-sky analysis of the QUaD CMB data). The main advantage of a flat-sky analysis is speed, especially if very small scales are being analysed, as Fast Fourier Transforms (FFTs) can be used, but the choice of modes is more poorly defined than a full-sky analysis, depending on the size of the patch analysed.

Hikage et al. (2011) used a curved PCl method to analyse simulated data for small masks as well as investigating flat-sky PCl estimation using similar masks. Kitching et al. (2014) used spherical-Bessel transforms of the shear field on flat sky to preform a 3D cosmic shear analysis of CFHTLenS data. Kitching et al. (2012) also used PCls as a tool to estimate the impact of shape measurement biases for the GREAT10 Challenge. Flat-sky PCl analysis has also been applied to data to estimate the cross-power spectrum of CMB and galaxy lensing maps (see Hand et al. 2015; Harnois-Déraps et al. 2016) and galaxy–galaxy lensing (see Hikage & Oguri 2016).

CMB data are also analysed using Maximum Likelihood estimators (see Planck Collaboration XVI 2014, for example), which can also be used on shear fields (see Seljak 1998; Hu & White 2001), and first applied by Brown et al. (2003) for the COMBO-17 survey and Köhlinger et al. (2016) for the CFHTLenS data. However, these methods may be too slow for current and future surveys where the number of pixels and the general resolution are high. Other methods exist that estimate the power spectrum indirectly, using 2PCFs (see e.g. Szapudi et al. 2001; Chon et al. 2004; Becker et al. 2015). Finally, an alternative approach that shares many of the advantages of the PCl technique, but can more easily treat the masked regions, is Bayesian hierarchical modelling (Alsing, Heavens & Jaffe 2016a; Alsing et al. 2016b), although this method is substantially slower than a PCl analysis.

In this paper, we study the effects of masking in the PCl approach for weak lensing, for both small and large masks, using Gaussian and lognormal simulated shear fields on a flat sky. We first go through the formalism of PCls in Section 2, where we explain how the mode mixing can be modelled via a mixing matrix. In Section 3, the resulting pseudo power spectra and the recovered power spectra are shown and compared with their expected values from theory. Finally, in Section 4 we propagate the random and mask modelling errors to the cosmological parameters using a Fisher analysis and check for significant biases.

2 FORMALISM

In this section, we review the basic formalism and go through some of the more important steps taken to calculate the mixing matrix, which models the effects of masking on the power spectra. The details of these calculations are given in Memari (2009). The following formalism is written for a flat-sky approximation. The formalism here has some differences from the one outlined in Hikage et al. (2011), and we apply an additional angular averaging to simplify the relations and speed up the calculations.

The convergence can be separated into two real parts $\kappa_{\rm E}$ and $\kappa_{\rm B}$ in real space. Weak gravitational lensing can produce only $\kappa_{\rm E}$ up to first order in the Newtonian gravitational potential. Hence, any $\kappa_{\rm B}$ would come from other effects, including systematic errors and intrinsic galaxy alignments. In Fourier space, we can write κ in terms of the Fourier transforms of $\kappa_{\rm E,B}(\boldsymbol{\vartheta})$,

$$\hat{\kappa}_{\pm}(\boldsymbol{\ell}) = \hat{\kappa}_{\rm E}(\boldsymbol{\ell}) \pm {\rm i}\hat{\kappa}_{\rm B}(\boldsymbol{\ell}), \tag{1}$$

with

$$\hat{\kappa}_{\rm E,B}(\boldsymbol{\ell}) = \int {\rm d}^2\vartheta\, \kappa_{\rm E,B}(\boldsymbol{\vartheta}) e^{-{\rm i}\boldsymbol{\ell}\cdot\boldsymbol{\vartheta}}, \tag{2}$$

where a hat refers to a Fourier-space quantity. Note that, $\hat{\kappa}_{\rm E,B}(\boldsymbol{\ell})$ are complex quantities. We can also write $\gamma_{\pm}$ as

$$\gamma_{\pm}(\boldsymbol{\vartheta}) = \gamma_1(\boldsymbol{\vartheta}) \pm {\rm i}\gamma_2(\boldsymbol{\vartheta}), \tag{3}$$

where $\gamma_{1,2}$ are the shear components in Cartesian coordinates. The Fourier transform of $\gamma_{\pm}$ is

$$\hat{\gamma}_{\pm}(\boldsymbol{\ell}) = \hat{\gamma}_1(\boldsymbol{\ell}) \pm {\rm i}\hat{\gamma}_2(\boldsymbol{\ell}), \tag{4}$$

where $\hat{\gamma}_{1,2}(\boldsymbol{\ell})$ are the Fourier transforms of $\gamma_{1,2}(\boldsymbol{\vartheta})$, respectively. To find the relation between $\hat{\kappa}_{\pm}(\boldsymbol{\ell})$ and $\hat{\gamma}_{\pm}(\boldsymbol{\ell})$, we note that they are both functions of the lensing potential, $\psi_{\pm}$, via

$$\gamma_+ = \tfrac{1}{2}\eth\eth\psi_+\,, \quad \gamma_- = \tfrac{1}{2}\eth^*\eth^*\psi_-\,, \quad \kappa_{\pm} = \tfrac{1}{2}\eth^2\psi_{\pm}, \tag{5}$$

where

$$\psi_{\pm} = \psi_{\rm E}(\boldsymbol{\vartheta}) \pm {\rm i}\psi_{\rm B}(\boldsymbol{\vartheta})\,, \qquad \psi_- = \psi_+^*, \tag{6}$$

and

$$\eth \equiv \partial_1 + {\rm i}\partial_2\,, \quad \eth^* \equiv \partial_1 - {\rm i}\partial_2 \quad \text{and} \quad \eth^2 \equiv \eth\eth^*, \tag{7}$$

where $\partial_{1,2}$ are partial derivatives with respect to θ_1, θ_2. Eliminating ψ in equation (5) results in relations between the shear and convergence,

$$\kappa_+ = \partial^*\partial^*\partial^{-2}\gamma_+ \qquad \text{and} \qquad \kappa_- = \partial\partial\partial^{-2}\gamma_-, \tag{8}$$

where the inverse Laplacian operator is

$$\partial^{-2} \equiv \int \frac{\mathrm{d}^2\boldsymbol{\vartheta}'}{2\pi} \ln|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}'|\,. \tag{9}$$

In Fourier space, the relation between κ and γ is more straightforward. Using the relation between the partial derivatives in real space with their Fourier counterparts,

$$\mathcal{F}(\partial) = \mathrm{i}\hat{\ell}\,, \qquad \mathcal{F}(\partial^*) = \mathrm{i}\hat{\ell}^*, \tag{10}$$

where $\mathcal{F}$ refers to a Fourier transform and

$$\hat{\ell} = \ell_x + \mathrm{i}\ell_y\,, \qquad \hat{\ell}^* = \ell_x - \mathrm{i}\ell_y, \tag{11}$$

we find

$$\begin{aligned}\hat{\kappa}_+(\boldsymbol{\ell}) &= \hat{\ell}^*\hat{\ell}^*|\hat{\ell}|^{-2}\hat{\gamma}_+(\boldsymbol{\ell})\,,\\ \hat{\kappa}_-(\boldsymbol{\ell}) &= \hat{\ell}\hat{\ell}|\hat{\ell}|^{-2}\hat{\gamma}_-(\boldsymbol{\ell})\,.\end{aligned} \tag{12}$$

Simplifying the above equations by substituting for $\hat{\ell}$ from

$$\hat{\ell} = \ell e^{\mathrm{i}\varphi_\ell}\,, \quad \text{with} \quad \ell = |\hat{\ell}|, \tag{13}$$

results in

$$\hat{\kappa}_\pm(\boldsymbol{\ell}) = e^{\mp 2\mathrm{i}\varphi_\ell}\hat{\gamma}_\pm(\boldsymbol{\ell}), \tag{14}$$

where φ_ℓ is the polar angle of both $\boldsymbol{\ell}$ and $\hat{\ell}$.

2.1 Masking effects on shear fields

In any realistic scenario, parts of the images are masked. Formally, we need to know only the position of source galaxies. However, analysing a gridded image is significantly faster, since FFTs can be utilized in this case. A PCl analysis relies on such gridded fields, where any region with no signal resulting from observers' choices or faulty and empty pixels produces the mask. We can choose to apodize the masked shear field with a smoothing kernel, S, to avoid sharp mask features, which make the Fourier transform of the mask challenging. The masks used in this work consist of ones and zeros exclusively. However, in practice the detector defect masks are usually smoother, due to dithering of the observed images. If the mask provided by the observer is smooth enough, then it will mimic an apodized binary mask.

A mask, W, has a multiplicative effect on the shear field,

$$\widetilde{\gamma}_\pm(\boldsymbol{\vartheta}) = W(\boldsymbol{\vartheta})\gamma_\pm(\boldsymbol{\vartheta}), \tag{15}$$

where we assume that $W(\boldsymbol{\vartheta}) = 0$ corresponds to a fully masked region. Any quantity with a tilde denotes a masked or pseudo quantity from here on. There are two ways to apodize a mask, one is to convolve the masked shear field, $\widetilde{\gamma}(\boldsymbol{\vartheta})$, with S,

$$\widetilde{\gamma}^{\mathrm{s}}_\pm(\boldsymbol{\vartheta}) = \int \mathrm{d}^2\boldsymbol{\vartheta}' S(\boldsymbol{\vartheta} - \boldsymbol{\vartheta}')W(\boldsymbol{\vartheta}')\gamma_\pm(\boldsymbol{\vartheta}')\,. \tag{16}$$

The superscript s denotes a smoothed quantity. The other method is to take the mask and smooth its edges with a kernel. Note that when this apodization method is used, the mask will maintain its original zeros while smoothly transitioning to the unmasked parts, where $W(\boldsymbol{\vartheta}) = 1$. Therefore, using this method enlarges the mask. The original mask is then replaced by the new apodized mask. We use this method for apodizing the masks in this work.

In Fourier space, the shear field is first convolved with the mask (we will drop the hat for Fourier counterparts from here on for simplicity, e.g. $\hat{\gamma}(\boldsymbol{\ell}) \to \gamma(\boldsymbol{\ell})$),

$$\widetilde{\gamma}_\pm(\boldsymbol{\ell}) = \int \frac{\mathrm{d}^2\ell'}{(2\pi)^2} W(\boldsymbol{\ell} - \boldsymbol{\ell}')\gamma_\pm(\boldsymbol{\ell}'), \tag{17}$$

and then multiplied by the smoothing kernel if the first apodization method is used,

$$\widetilde{\gamma}^{\mathrm{s}}_\pm(\boldsymbol{\ell}) = S(\boldsymbol{\ell}) \int \frac{\mathrm{d}^2\ell'}{(2\pi)^2} W(\boldsymbol{\ell} - \boldsymbol{\ell}')\gamma_\pm(\boldsymbol{\ell}')\,. \tag{18}$$

Substituting from equation (14) into the above equation, we can find a relation for the masked κ,

$$\begin{aligned}&\widetilde{\kappa}_\pm(\boldsymbol{\ell}) = \int \frac{\mathrm{d}^2\ell'}{(2\pi)^2} W(\boldsymbol{\ell} - \boldsymbol{\ell}')\kappa_\pm(\boldsymbol{\ell}')e^{\mp 2\mathrm{i}\varphi_{\ell\ell'}}\\ &\text{and} \quad \widetilde{\kappa}^{\mathrm{s}}_\pm(\boldsymbol{\ell}) = S(\boldsymbol{\ell})\widetilde{\kappa}_\pm(\boldsymbol{\ell}), \quad \text{where} \quad \varphi_{\ell\ell'} = \varphi_\ell - \varphi_{\ell'}.\end{aligned} \tag{19}$$

By adding and subtracting the equations above, we can find a relation between the masked and unmasked $\kappa_{\mathrm{E,B}}$,

$$\begin{aligned}\widetilde{\kappa}_{\mathrm{E}}(\boldsymbol{\ell}) &= \int \frac{\mathrm{d}^2\ell'}{(2\pi)^2} W(\boldsymbol{\ell} - \boldsymbol{\ell}')[\kappa_{\mathrm{E}}(\boldsymbol{\ell}')\cos 2\varphi_{\ell\ell'} + \kappa_{\mathrm{B}}(\boldsymbol{\ell}')\sin 2\varphi_{\ell\ell'}],\\ \widetilde{\kappa}^{\mathrm{s}}_{\mathrm{E}}(\boldsymbol{\ell}) &= S(\boldsymbol{\ell})\widetilde{\kappa}_{\mathrm{E}}(\boldsymbol{\ell}),\\ \widetilde{\kappa}_{\mathrm{B}}(\boldsymbol{\ell}) &= \int \frac{\mathrm{d}^2\ell'}{(2\pi)^2} W(\boldsymbol{\ell} - \boldsymbol{\ell}')[\kappa_{\mathrm{B}}(\boldsymbol{\ell}')\cos 2\varphi_{\ell\ell'} - \kappa_{\mathrm{E}}(\boldsymbol{\ell}')\sin 2\varphi_{\ell\ell'}],\\ \widetilde{\kappa}^{\mathrm{s}}_{\mathrm{B}}(\boldsymbol{\ell}) &= S(\boldsymbol{\ell})\widetilde{\kappa}_{\mathrm{B}}(\boldsymbol{\ell}),\end{aligned} \tag{20}$$

where $\widetilde{\kappa}^{\mathrm{s}}_{\mathrm{E,B}}(\boldsymbol{\ell})$ are the smoothed and masked E/B-mode κ. The above relations show that the mask affects the convergence in Fourier space by mixing some of the E-mode components into the B-modes and vice versa. Consequently, in order to utilize Fourier space information in cosmic shear analysis, the effects of the mask must be modelled.

The masks considered in this work are categorized into two groups: small- and large-scale masks. We also combine these masks to make the composite mask. Fig. 1 shows the star and checkerboard masks used throughout this work. These are plausible mask models that resemble masks used for real data (see Erben et al. 2013, for example). The left-hand panel shows the star mask that contains randomly positioned circles with random areas picked from three ranges; 2 per cent of the field is covered with stars from [0.1, 0.5], 5 per cent from [1, 25], and 3 per cent from [15, 100], square arcminutes. The checkerboard mask, which represents a Charge-Coupled Device (CCD) chip pattern or any other regular large-scale pattern, contains three dark pixels to simulate chip boundaries. These masks contain only ones and zeros. The masked regions are shown in black. The masks are zero padded to twice their size in each direction to minimize artefacts from the assumed periodic boundary conditions. As can be seen in Fig. 1, the masks have sharp features that motivate smoothing.

2.2 The mixing matrix

For a Gaussian isotropic random field, all the information is contained in the $C(\ell)$, which depends only on the absolute value of the wave numbers. However, the mask in general is not isotropic, which means that its power spectrum depends on the wave number angle, ϕ_ℓ, as well. In order to apply the mask to a theory power spectrum, averages over its angular dependences are taken and a mixing matrix is calculated. In practice, aside from the mixing matrix, the effects of ℓ-mode binning have to be included for a more accurate and comprehensive analysis. The exact steps that need to be taken

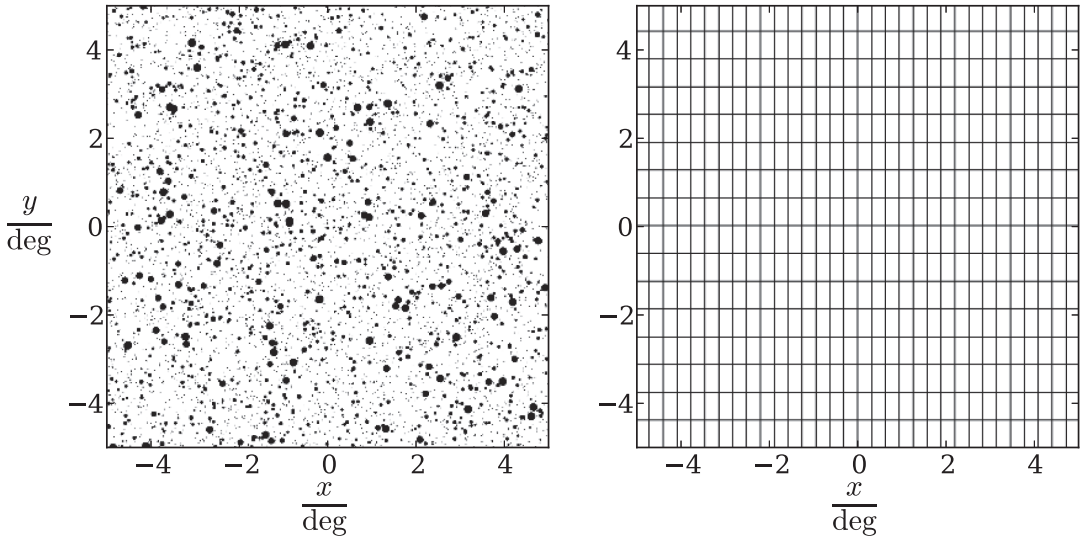

Figure 1. Star and checkerboard masks. The star mask contains randomly positioned circles with random areas picked from three ranges, 2 per cent from [0.1,0.5], 5 per cent from [1,25], and 3 per cent from [15,100] square arcminutes. The checkerboard mask mimics a CCD gap pattern. Three pixels are masked in the gaps. These two masks are also combined to simulate a more realistic scenario.

in such an analysis are explained in Section 2.4. For simplicity of the formalism, here we ignore the binning effects.

The power spectrum is defined as

$$\left\langle |\kappa_{\rm X}(\boldsymbol{\ell})\kappa^*_{\rm Y}(\boldsymbol{\ell}')| \right\rangle = (2\pi)^2\delta_{\rm D}(\boldsymbol{\ell}-\boldsymbol{\ell}')C(\ell)^{\rm XY}, \tag{21}$$

where X and Y represent E or B, and the angle brackets denote an ensemble average. Therefore, the expected average $C(\ell)$ is

$$\left\langle C(\ell)^{\rm XY} \right\rangle = \frac{1}{A}\int_0^{2\pi}\frac{{\rm d}\varphi_\ell}{2\pi}\left\langle |\kappa_{\rm X}(\boldsymbol{\ell})\kappa^*_{\rm Y}(\boldsymbol{\ell})| \right\rangle, \tag{22}$$

where the integral is a simple angle averaging and over a finite area the Dirac delta in equation (21) is replaced by

$$\delta_{\rm D}(0) \approx \frac{A}{(2\pi)^2}, \tag{23}$$

where A is the area of the field. The full $\delta_{\rm D}$ function is recovered when $A \to \infty$. In practice, because of the existing masks on the images, we can measure only a pseudo power spectrum, $\widetilde{C}(\ell)$. In the absence of noise, $\widetilde{C}(\ell)$ is defined in the same way as $C(\ell)$ in equation (22) by replacing $\kappa(\boldsymbol{\ell})$ with the masked convergence, $\widetilde{\kappa}(\boldsymbol{\ell})$. However, the cosmological models provide us with the power spectrum, $C(\ell)$. We can find a relation between $C(\ell)$ and $\widetilde{C}(\ell)$ by inserting for $\widetilde{\kappa}(\boldsymbol{\ell})$ from equation (20) into the masked version of equation (22). For example, the E-mode PCl, $\widetilde{C}^{\rm EE}(\ell)$, can be written as,

$$\begin{aligned}\left\langle \widetilde{C}(\ell)^{\rm EE} \right\rangle = \frac{1}{A}\int\frac{{\rm d}\varphi_\ell}{2\pi}|S(\boldsymbol{\ell})|^2 \\ \times \int\frac{{\rm d}^2\ell'}{(2\pi)^2}W(\boldsymbol{\ell}-\boldsymbol{\ell}')\int\frac{{\rm d}^2\ell''}{(2\pi)^2}W^*(\boldsymbol{\ell}-\boldsymbol{\ell}'') \\ \times \Big\langle \big|[\kappa_{\rm E}(\boldsymbol{\ell}')\cos 2\varphi_{\ell\ell'} + \kappa_{\rm B}(\boldsymbol{\ell}')\sin 2\varphi_{\ell\ell'}] \\ \times [\kappa^*_{\rm E}(\boldsymbol{\ell}'')\cos 2\varphi_{\ell\ell''} + \kappa^*_{\rm B}(\boldsymbol{\ell}'')\sin 2\varphi_{\ell\ell''}]\big| \Big\rangle,\end{aligned} \tag{24}$$

where $S(\boldsymbol{\ell})$ can be ignored if the mask is not smoothed or the second apodization scheme is used. The mask is not a variable between the realizations (that is, assuming that there is no correlation between the mask and the underlying shear field); therefore, we can take W out of the ensemble averages. Moreover, choosing a symmetric smoothing kernel allows us to take $|S(\boldsymbol{\ell})|^2$ out of the integral over φ_ℓ. Using equation (21), we link the $\widetilde{C}(\ell)$ to the $C(\ell)$,

$$\begin{aligned}\left\langle \widetilde{C}^{\rm EE}(\ell) \right\rangle = \frac{|S(\ell)|^2}{A}\int\frac{{\rm d}\varphi_\ell}{2\pi}\int\frac{{\rm d}^2\ell'}{(2\pi)^2}|W(\boldsymbol{\ell}-\boldsymbol{\ell}')|^2 \\ \times \left\{C^{\rm EE}(\ell')\cos^2 2\varphi_{\ell\ell'}\right. \\ + [C^{\rm EB}(\ell') + C^{\rm BE}(\ell')]\sin 2\varphi_{\ell\ell'}\cos 2\varphi_{\ell\ell'} \\ \left. + C^{\rm BB}(\ell')\sin^2 2\varphi_{\ell\ell'}\right\}.\end{aligned} \tag{25}$$

The above equation is written for the E-mode power spectrum although it can be extended to the other cases, shown in Appendix A. While $C(\ell')^{\rm XY}$ depend only on $|\boldsymbol{\ell}'|$, $W(\boldsymbol{\ell}-\boldsymbol{\ell}')$ and the trigonometric functions in equation (25) depend on the polar angles $\varphi_{\ell'}$ and φ_ℓ. Therefore, the angle averaging part of the integrals in equation (25) can be taken independent of the cosmological model (see Memari 2009). The details of the calculations are given in Appendix A. The masking effect is hence modelled in the form of a mode mixing matrix, $\mathscr{M}$,

$$\mathscr{M}(\ell,\ell') \equiv \frac{|S(\ell)|^2}{(2\pi)^2 A}\int_0^\pi {\rm d}\eta\, W_{\gamma\gamma}\big(L(\ell,\ell',\eta)\big)M_\eta(\eta), \tag{26}$$

where η is the angle between $\boldsymbol{\ell}$ and $\boldsymbol{\ell}'$,

$$L(\ell,\ell',\eta) \equiv |\boldsymbol{\ell}-\boldsymbol{\ell}'| = \sqrt{\ell^2+\ell'^2-2\ell\ell'\cos(\eta)}, \tag{27}$$

and $W_{\gamma\gamma}(L)$ is the power spectrum of the mask,

$$W_{\gamma\gamma}(L) \equiv \int_0^{2\pi}\frac{{\rm d}\varphi_L}{2\pi}|W(\boldsymbol{L})|^2, \tag{28}$$

and

$$M_\eta(\eta) \equiv \begin{pmatrix} 1+\cos 4\eta & 1-\cos 4\eta & 0 \\ 1-\cos 4\eta & 1+\cos 4\eta & 0 \\ 0 & 0 & 2\cos 4\eta \end{pmatrix}. \tag{29}$$

As a result, we can write

$$\widetilde{\boldsymbol{C}}(\ell) = \int_0^\infty {\rm d}\ell'\ell'\mathscr{M}(\ell,\ell')\boldsymbol{C}(\ell'), \tag{30}$$

where

$$\boldsymbol{C} \equiv \left(C^{\rm EE}(\ell), C^{\rm BB}(\ell), C^{\rm EB}(\ell)\right)^t. \tag{31}$$

In practice, we need to change all the integrals in the above equations into discrete finite sums. Therefore, we write equation (30) as follows for discrete values of ℓ,

$$\widetilde{\boldsymbol{C}}(\ell_i) = \sum_j \Delta\ell_j \ell_j \mathscr{M}(\ell_i, \ell_j)\boldsymbol{C}(\ell_j)\,. \tag{32}$$

The ℓ_i values depend on the size of the real-space size of the field and the binning used. We explain the details of the binning used in this work in Section 2.4. For the above equation, however, we use the smallest ℓ-bin that is allowed to keep the discrete sum as close as possible to its continuous form in equation (30). Note that the integrals in equations (26) and (28) also need to be changed into discrete sums. These approximations are the main reason for the biases that we will see in the Results sections. Full-sky PCls also suffer from biases, which are discussed in Elsner, Leistedt & Peiris (2016), where they offer a proposal to resolve them. A full-sky analysis, however, does not suffer from some of the limitations faced by flat-sky analysis since the limits on the Fourier modes and the binning are well defined.

We absorb $\Delta\ell_i\ell_i$ in equation (32) in $\mathscr{M}(\ell_i, \ell_j)$ and define a new mixing matrix,

$$M_{\ell_i\ell_j} = \Delta\ell_j \ell_j \mathscr{M}(\ell_i, \ell_j), \tag{33}$$

which satisfies this matrix relation,

$$\widetilde{C}_\ell = M_{\ell\ell'} C_{\ell'}, \tag{34}$$

where, for simplicity, we have dropped the i and j subscripts and replaced them with ℓ and ℓ' instead and Einstein summation rules apply. The $\widetilde{C}_\ell$ are the elements of a vector of $\widetilde{C}(\ell)$ for discrete values of ℓ. Consequently, we can write the inverse of equation (34) and recover the power spectrum,

$$C_\ell^{\rm rec} = (\mathbf{M}^{-1})_{\ell\ell'}\widetilde{C}_{\ell'}, \tag{35}$$

provided we can invert $\mathbf{M}$.

A prominent source of noise in weak lensing analysis is the galaxy shape noise that we model as a Gaussian random noise with zero mean and σ_ϵ dispersion corresponding to the dispersion of the complex galaxy ellipticity (see Hu 1999). The mask affects the noise in the same way as the shear field. We can write the noise as a separate source of power with no ℓ dependence,

$$N_\epsilon = \frac{\sigma_\epsilon^2}{2n_{\rm gal}}, \tag{36}$$

where $n_{\rm gal} = 30$ per square arcminutes is the mean number density of galaxies and $\sigma_\epsilon = 0.3$ is the intrinsic dispersion of galaxy ellipticities, similar to values expected from a Euclid-like future survey (see Laureijs et al. 2011).[1] As a result, the measured PCl is

$$\widetilde{\boldsymbol{C}} = \mathbf{M}[\boldsymbol{C} + N_\epsilon], \tag{37}$$

and the recovered Cl is

$$\boldsymbol{C}^{\rm rec} = (\mathbf{M}^{-1})\widetilde{\boldsymbol{C}} - N_\epsilon. \tag{38}$$

2.3 Mask smoothing: apodization

We use three Gaussian smoothing kernels to apodize the masks using the second method (smoothing the edges of $W(\boldsymbol{\vartheta})$ before applying it to the shear fields). Hence, in the equations where S appears, it should be ignored. The advantage of this method is that it allows slower variations for the integrands in equations (26) and (28), which could make their discrete approximation more accurate. As can be seen in equation (26) in the case of the first apodization method, the smoothing kernel comes into play only after the angular averages in equations (26) and (28) have been taken.

Note that using this method increases the effective masked area, since the fully masked regions will remain the same, while their edges will have a smooth transition from zero to one, which is determined by the size of the kernel. In general, any smoothing function can be chosen as the kernel. However, here we use Gaussian kernels. They are identified using the number of pixels that determines their size in real-space pixels, N, which is an odd number. We set all the values outside a box centred at the maximum of the Gaussian with N pixels on each side to zero and set the dispersion of the Gaussian kernel to $(N-1)/1.5$. The apodization is done around the edges of the masks, such that the apodized mask starts from 0 on the edge of the original mask and transitions to 1 over roughly N-1 pixels. An odd N is chosen so that the kernel is symmetric around its origin.[2] Here we use three sizes for the Gaussian kernels that are listed in Table 1. The main results are shown and compared for the original masks and these three apodization schemes.

Table 1. Apodization case name and the number of pixels, N, used in defining the kernel. The kernels used in this work are Gaussian functions with $\sigma = (N-1)/1.5$ with a range of support equal to N pixels. The first row shows the name given to each case.

	Ap1	Ap2	Ap3
N	5 pixels	11 pixels	23 pixels

2.4 Binning and pixelization effects: theory CL and PCL

In practice, the 2D Fourier fields (shear and mask) are pixelated; hence, their angle averaged values are not exact and depend on the method used. We will ignore the window function of the map pixel shape, as it is important only for very high Fourier modes that are not used in this work. To take the angular average over such a field, we choose an annulus around the centre of the field, identify all the pixels with centres lying inside the annulus, and take their average value. The Fourier mode, ℓ, which corresponds to this estimated value, is also calculated by averaging over the value of the $|\boldsymbol{\ell}|$ modes that lie in this annulus. Fig. 2 demonstrates the angle averaging scheme. The edges of the annuli are shown as concentric circles, and the pixels that correspond to an annulus are shown in different colours. The second annulus, for example, has 8 pixels with their centres lying inside it, which are painted orange. As we go to larger annuli, the number of pixels increases, which in turn increases the accuracy of the angle averaging.

To estimate $W_{\gamma\gamma}$ from equation (28), we use annuli of width equal to the smallest Fourier mode, $\Delta\ell_{\rm min}$, available in the field

$$\Delta\ell_{\rm min} = \frac{2\pi}{D}, \tag{39}$$

where D is the side length of the square field after zero padding. Then the estimated $W_{\gamma\gamma}$ values are fed into equation (26). Since $W_{\gamma\gamma}$ is estimated only for discrete values, the integral in equation (26) needs to be transformed into a sum over these values of L. This integral

[1] www.Euclid-ec.org

[2] The apodization is done by first zero-padding each mask with $(N-1)/2$ pixels and then convolving it with a 2D convolution method (filter2 in MATLAB).

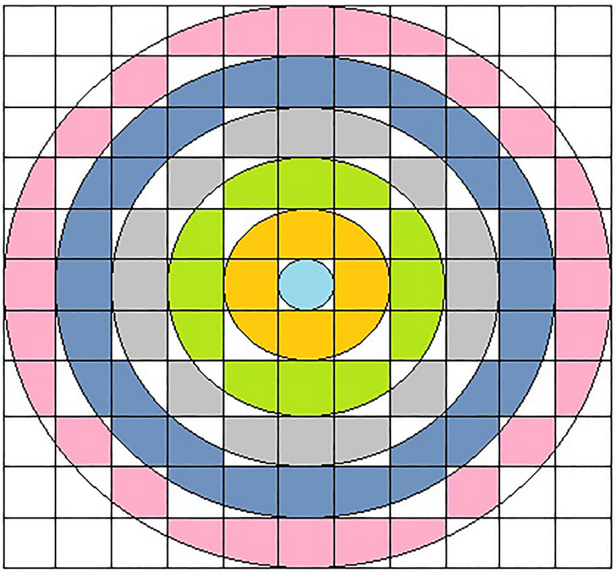

Figure 2. A visual presentation of the angle averaging scheme. For the smallest binning scheme, each annulus is defined between two adjacent circles. An average is taken over the pixels with centres lying in the annuli, painted in different colours. The ℓ value that corresponds to the angle averaged quantity is also the average of the $|\boldsymbol{\ell}|$ of the relevant pixels. The binning is done by merging two or more of the smallest annuli with each other.

is taken over η, the angle between $\boldsymbol{\ell}$ and $\boldsymbol{\ell}'$, which is calculated for each L, ℓ and $\ell^{'}$ from

$$\cos\eta = \frac{\ell^2 + \ell'^2 - L^2}{2\ell\ell'}. \tag{40}$$

The values of ℓ and $\ell^{'}$ depend on the binning scheme used. Note that in equation (26), the available η do not form a regular grid, therefore, dη is not constant.

To find the PCls from theory, first $\mathbf{M}_{\ell\ell'}$ is estimated for the smallest binning that corresponds to the field after zero-padding, similar to $W_{\gamma\gamma}$. Secondly, the input theory power spectrum, $C^{\rm input}(\ell)$, is laid on a grid with the same pixel size and area. Then it is angle averaged to find $C_\ell^{\rm ave}$. $C^{\rm input}(\ell)$ and $C_\ell^{\rm ave}$ show the largest differences for the smallest ℓ modes as expected, since the number of pixels that lie in the first few annuli for the angle averaging is not representative, which biases the results (see Fig. 2). Finally, the theory value of PCls is estimated by applying $\mathbf{M}_{\ell\ell'}$ to the sum of $C_\ell^{\rm ave}$ and N_ϵ. The $\widetilde{C}(\ell)$ can then be re-binned into broader bins as desired. These steps can be summarized in

$$\widetilde{\boldsymbol{C}}_{\rm b}^{\rm th} = \mathscr{B}\{\mathbf{M}(\boldsymbol{C}^{\rm ave} + N_\epsilon)\}, \tag{41}$$

where $\widetilde{\boldsymbol{C}}_{\rm b}^{\rm th}$ is the binned theory PCl and $\boldsymbol{C}^{\rm ave}$ is $C_\ell^{\rm ave}$ in vector format, $\mathbf{M}$ is the mixing matrix, and $\mathscr{B}$ is the binning matrix defined as follows:

We define n_ℓ as the number of smallest bins of size $\Delta\ell_{\rm min}$ that are combined to make the wider bins. If the total number of initial ℓ-bins is $n_{\rm tot}$, then there are $n_{\rm bin} = \lfloor n_{\rm tot}/n_\ell \rfloor$ wider bins, where $\lfloor x \rfloor$ denotes the largest integer that is smaller than x. $\mathscr{B}$ is then an $n_{\rm bin} \times n_{\rm tot}$ matrix of this form

$$\mathscr{B}_{b\ell} = \begin{cases} n_{\rm p}(\ell)\left(\sum\limits_{(b-1)n_\ell+1}^{bn_\ell} n_{\rm p}(\ell)\right)^{-1} & (b-1)n_\ell < \ell \le bn_\ell \\ 0 & \ell > bn_\ell \quad \text{or} \quad \ell \le (b-1)n_\ell, \end{cases} \tag{42}$$

where $n_{\rm p}(\ell)$ is the number of pixels in each initial ℓ-bin (see Fig. 2).

The ellipticity noise contribution, $\mathscr{B}(\mathbf{M}N_\epsilon)$, can be subtracted from the theory and measured values subsequently. We forward model the $\widetilde{C}_\ell^{\rm th}$, so it is the closest to the estimated PCls.

We can use two methods to recover a C_ℓ, which result in very different values. The first method is to apply the inverse of the mixing matrix on the $\widetilde{C}_\ell^{\rm est}$ measured from the fields and then bin the result into wider ℓ-bins,

$$\text{I}: \boldsymbol{C}_{\rm b}^{\rm rec} = \mathscr{B}(\mathbf{M}^{-1}\widetilde{\boldsymbol{C}}^{\rm est}) - N_\epsilon, \tag{43}$$

where $\boldsymbol{C}_{\rm b}^{\rm rec}$ is the binned recovered $C(\ell)$ in vector format. This recovered power spectrum can be compared with a binned $C_\ell^{\rm ave}$. The advantage of this method is that it is less computationally intensive as the mixing matrix is applied only once on the $\widetilde{C}_\ell^{\rm est}$ and not on the different theory values.

The second method is to recover $C(\ell)$ by applying the inverse of a binned mixing matrix to a binned estimated $\widetilde{C}(\ell)$

$$\text{II}: \boldsymbol{C}_{\rm b}^{\rm rec} = \mathbf{M}_{\rm b}^{-1}\widetilde{\boldsymbol{C}}_{\rm b}^{\rm est}. \tag{44}$$

We can write the predicted theory value for this recovered $C_\ell^{\rm rec}$ as

$$\boldsymbol{C}_{\rm b}^{\rm th} = \mathbf{M}_{\rm b}^{-1}\widetilde{\boldsymbol{C}}_{\rm b}^{\rm th} = \mathbf{M}_{\rm b}^{-1}\mathscr{B}\{\mathbf{M}(\boldsymbol{C}^{\rm ave} + N_\epsilon)\}, \tag{45}$$

where the noise contribution $\mathbf{M}_{\rm b}^{-1}\mathscr{B}\mathbf{M}N_\epsilon$ can be subtracted from both recovered and theory values in equations (44) and (45). To use the second method, we would need to apply the mixing matrix on the $C_\ell^{\rm ave}$ value to find $\widetilde{C}_\ell^{\rm th}$. Therefore, this method is at least as computationally demanding as the forward modelling where the theory PCls are compared to their measured values. Furthermore, as the binned mixing matrix is more diagonal, the shape of the recovered C_ℓ from this method is similar to $\widetilde{\boldsymbol{C}}_{\rm b}^{\rm th}$, instead of the underlying power spectrum.

2.5 Mask power spectra and mixing matrices

In the past sections, we explained the methods we use to find the power spectrum and mixing matrix of a mask, as well as our apodization scheme. Here we show the mask power spectra, $W_{\gamma\gamma}(\ell)$ for all the mask and apodization combinations and show examples of mixing matrices.

Fig. 3 shows the mask power spectra defined in equation (28). Each row belongs to a mask configuration, while each column shows the results for different apodization cases, which are indicated at the right-hand side of the rows and the top of the columns. The fields are zero-padded before the FFT is applied to them; therefore, even the 'No Mask' case has a large-scale square-shaped mask, which one could call the survey footprint. The zero-padding allows for a higher resolution estimate of $W_{\gamma\gamma}$, effectively interpolating between the natural ℓ values of the original pixelated field. We zero-pad the masks to double their size on each side.

$W_{\gamma\gamma}$ is presented for the smallest binning available, since it is used in this format to find the mixing matrix (see equation 26). We use the 'No Mask' option as the control case. By comparing the first row in the figure with the following rows, we see that the largest scale feature, at small ℓ values, is due to the zero-padding. The Fourier transform of a perfect square mask is a double sinc function. Therefore, the power spectrum of such a mask oscillates heavily, which is what we see for the 'No Mask' version. As we add more structures to the mask, this oscillatory behaviour is suppressed, which is what we see in Fig. 3. One of the main differences between the current work and Hikage et al. (2011) is that in contrast to this work, they assumed periodic boundary conditions and only small-scale masks to study PCls. Unlike the star mask that has a

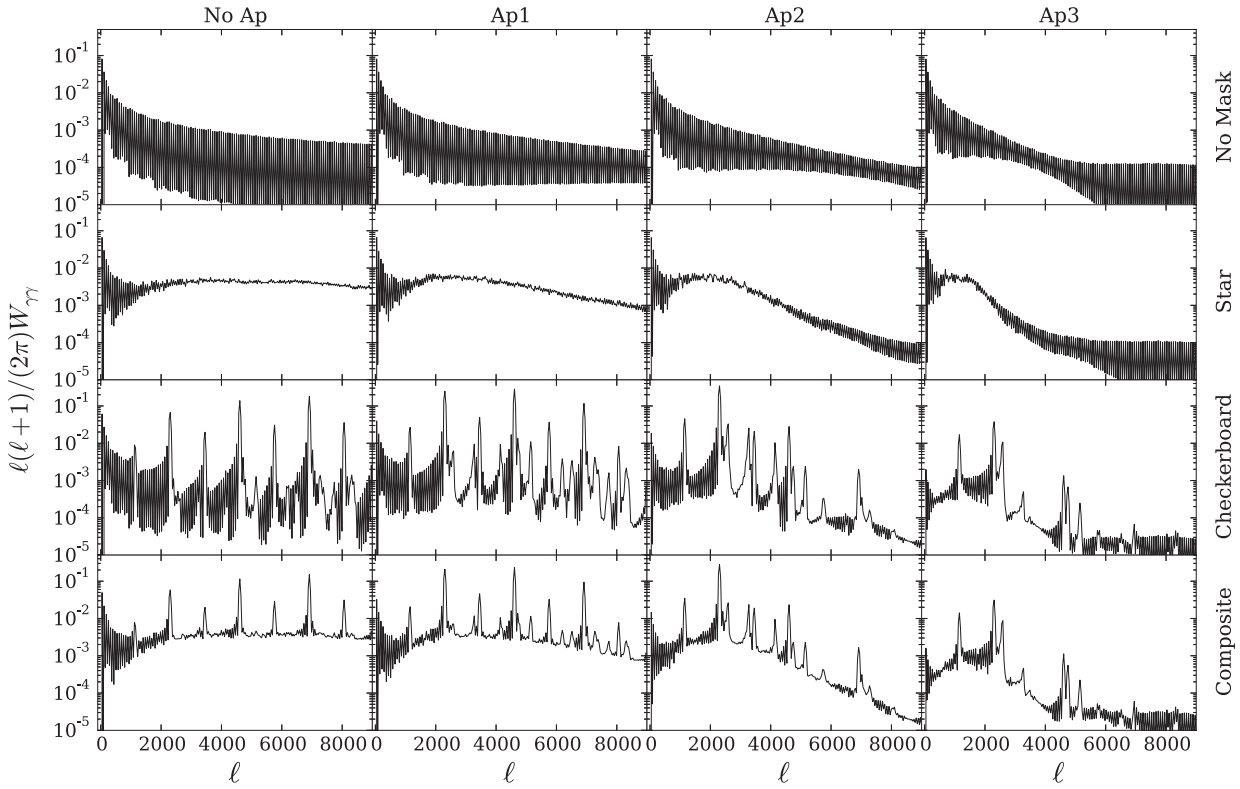

Figure 3. Normalized power spectra of all the mask and apodization combinations for the smallest binning ($\Delta\ell_{\rm min} \approx 18$/rad). The fields are all zero-padded; hence, even the 'No Mask' case has a square-shaped mask. The composite mask is one that combines the star and checkerboard patterns. 'No Ap' means that no apodization has been applied to the original mask, while 'Ap1', 'Ap2', and 'Ap3' indicate masks that are apodized using Gaussian kernel with increasing support, respectively (see Section 2.3). Note that the *x*-axis is linear while the *y*-axis is in logarithmic scale.

relatively featureless power spectrum, the checkerboard shows very prominent peaks, due to its regular pattern. The star mask shows relatively constant power over a large range of $\ell \gtrsim 2000$, since it consists of randomly positioned stars of different sizes. The star mask is basically self-similar for this range of scales. The features in the power spectrum of the checkerboard mask form a comb corresponding to the harmonics of the regular pattern. Specifically, ℓ 1150 corresponds to the first harmonic of the checkerboard pattern.

By comparing different columns in Fig. 3, we see that the apodization dampens the tail of the power spectra, the scale and strength of which depend on the size of the kernel. As a result, we see that 'Ap3', which is the largest kernel we use, has a more dramatic effect on $W_{\gamma\gamma}$ compared to the smaller kernels. Note that the apodization here only smooths the edges of the masked regions while keeping the zeros in the mask intact; hence, more apodization results in a larger effective masked area. The smoothing is not as effective on the checkerboard power spectra, which will have important consequences for the estimation of the convergence power spectra, the effects of which will become apparent in Section 4.3.

Fig. 4 shows the mixing matrix for the composite mask (star and checkerboard). The left-hand panel shows the matrix for the original ones and zeros mask, while the right-hand panel shows the same for the mask apodized with Ap1. Since the EB–EB part of the mixing matrix is independent of the EE and BB parts, it is not shown here and will not be used in any of the analysis. The mixing matrices are binned with $n_\ell = 20$ to produce 25 approximately linearly spaced bins with $\Delta\ell \approx n_\ell \times \Delta\ell_{\rm min} = 360$/rad. The matrices are plotted in terms of the logarithm of the absolute value of their elements. As can be seen in this figure, a smooth mask has a more diagonal mixing matrix and a smoother off-diagonal behaviour. The importance of this property of the mixing matrix will become clear in the next sections.

3 MEASURED POWER SPECTRA

To test the mask modelling, we use two sets of simulations: random realizations of Gaussian and lognormal shear fields. The input power spectrum is identical for both cases and is based on a cold dark matter Universe with a dominant dark energy component, with cosmological parameters given in Table 2.

The linear power spectrum is determined assuming a primordial power-law power spectrum with Eisenstein & Hu (1998) transfer function. Additionally, the halo fit formula of Smith et al. (2003) is used for calculating the non-linear scales. We use a single redshift distribution of Van Waerbeke et al. (2006) type with $0.2 < z < 1.3$, a median redshift of 0.7, $\alpha = 2$, and $\beta = 1.5$.

All the simulations are originally made for a larger field ($20^\circ \times 20^\circ$, 2048×2048 pixels) and then a $10^\circ \times 10^\circ$ field (1024×1024 pixels) is cut out of the middle to simulate the non-periodic nature of the Universe. 100 random realizations are generated for each case in the analysis. After adding a Gaussian random shape noise with $\sigma_\epsilon = 0.3$ to the shear fields, they are masked and then zero-padded before the Fourier transform. The zero-padding scheme used here changes the size of the fields to their original size, which means, doubling the size of the field on each side by adding zeros. The zero-padding ensures that a periodic boundary condition is not assumed for the field when the FFT is applied to it. Zero-padding the field more than this results in a computationally more expensive analysis

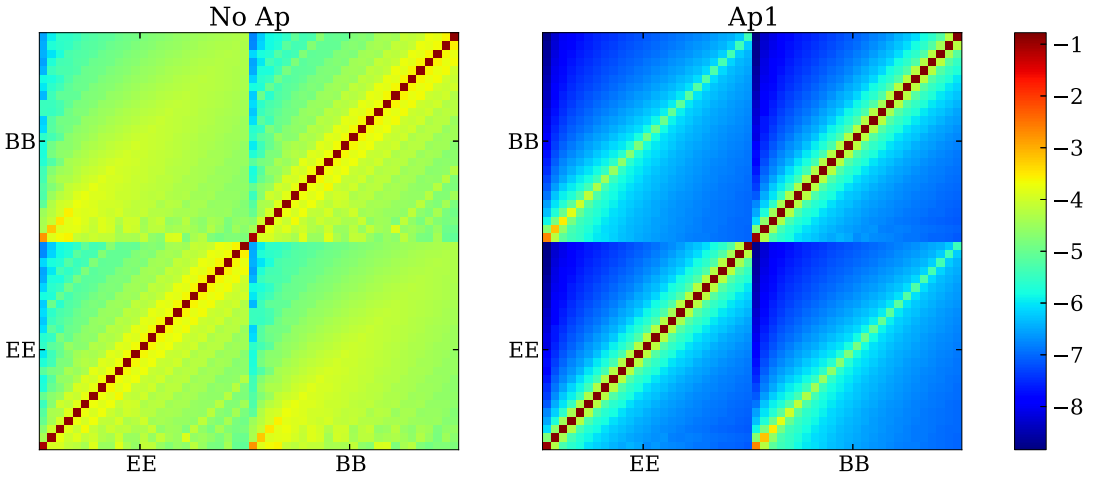

Figure 4. The logarithm of the absolute value of the mixing matrix for the composite mask (star and checkerboard). The left-hand panel shows the mixing matrix for the original mask with no apodization, whereas the right-hand panel shows the same for an apodized mask with Ap1. 25 linear ℓ-bins in [245,8830] are considered here, which correspond to $n_\ell = 20$.

Table 2. The fiducial cosmological parameters consistent with Planck 2013 results (Planck Collaboration XVI 2014). The normalization of the power spectrum, σ_8, is the standard deviation of perturbations in a sphere of radius $8h^{-1}$Mpc today. Ω_m, Ω_Λ, and Ω_b are the matter, dark energy, and baryonic matter density parameters, respectively. w_0 is the dark energy equation-of-state parameter, which is equal to the ratio of dark energy pressure to its density. The spectral index, n_s, is the power of the initial power spectrum. The dimensionless Hubble parameter, h, characterizes the rate of expansion today.

σ_8	Ω_m	Ω_Λ	w_0	n_s	h	Ω_b
0.8	0.27	0.73	−1.0	0.96	0.72	0.045

while the result remains similar. The FFT of a zero-padded field has a higher resolution; hence, zero-padding is also a non-unique form of interpolation between the Fourier modes. As a result, the resolution of the Fourier-transformed fields is 18/rad.

As we have seen in Section 2.2, the mixing matrices calculated from pixelated masks are not accurate. The inaccuracy in mask modelling is more severe for small ℓ that propagates to all scales (see equation 26). As a result, to first order a constant multiplicative bias needs to be corrected for. We measured this bias for different masks by taking the average ratio of the measured $\widetilde{C}(\ell)$ to their theory value. This bias mainly depends on the mask and is shown in Fig. 5 as

$$B_M = \left\langle \frac{\widetilde{C}_\ell^{\rm est}}{\widetilde{C}_\ell^{\rm th}} \right\rangle, \tag{46}$$

where the angled brackets mean that the average of the ratio is taken over the 100 simulations and ℓ modes. We show B_M for star, checkerboard, and three types of composite masks: composite without apodization, with Ap1, and a composite with a larger checkerboard pattern apodized with Ap1. In Fig. 1, we showed the checkerboard mask we use in the main analysis. The larger checkerboard pattern is coarser with one large rectangle instead of every four smaller rectangles shown in Fig. 1. The masks are varied so that the masked area changes. For the star mask, this is done by keeping the size range of the stars as the original but changing the percentage of star masks in each range. The checkerboard mask is varied by changing the number of dark pixels between the chips.

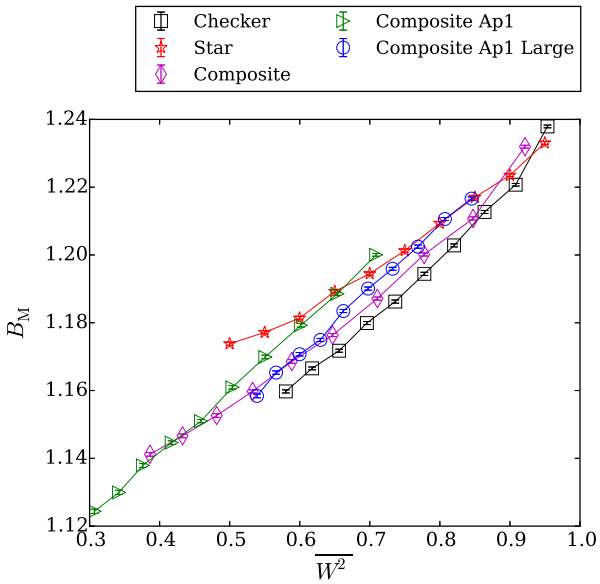

Figure 5. Constant multiplicative bias caused by inaccuracies in mask modelling on large scales. The x-axis shows the mean value of the square of the mask, which is a measure of the masked area and the scaling effect of the mask. B_M is shown for five types of masks. Star, checkerboard, and composite masks are explained in Section 2.2 and the apodization in Section 2.3. The blue circles show a modified composite mask that consists of a coarser checkerboard pattern (one rectangle for every four in the original checkerboard mask). The solid lines simply connect the symbols to guide the eye. The error bars show the variance of the mean between the 100 simulated field and are smaller than the symbols for all cases.

For the composite masks, both masks are varied simultaneously. The change in the masked area is captured by $\overline{W^2}$, corresponding to the mean value of the square of the mask before zero-padding. For a binary mask, $\overline{W^2}$ is equal to the fraction of the sky which is unmasked. Assuming no mode-mixing and zero-padding, applying a mask on the field results in the scaling of the power spectrum by $\overline{W^2}$. As can be seen in Fig. 5, the relation between B_M and $\overline{W^2}$ is roughly linear; however, it is not universal and depends on the type of mask used. Generally, for heavier masking B_M decreases, this

can be explained by looking at Fig. 3. The first peak in the plots corresponds to the effect of zero-padding. As we add more structures to the mask, the relative significance of this peak decreases, since more masking produces power on other scales. Fortunately, $B_{\rm M}$ is not very sensitive to the underlying power spectrum or the binning scheme. For example, in Fig. 5 we see that the scatter between the different realizations is very small, which shows that the value of $B_{\rm M}$ is mostly sensitive to the mask and not the exact value of the underlying power spectrum. $B_{\rm M}$ here is shown for the forward modelling case. A similar constant bias needs to be corrected for when $C_\ell^{\rm rec}$ is estimated, which has a similar behaviour and value. This constant bias is present only when large-scale masks are considered; therefore, Hikage et al. (2011), who considered only star masks with no apodization, did not report it.

We tested several different methods to estimate the mixing matrix and concluded that the only way to systematically tackle this challenge is to find and correct for $B_{\rm M}$ using simulations. Note that while changing the integration scheme or the number of zero-padded pixels, can reduce or eliminate $B_{\rm M}$ for certain masks, it will not be applicable to other masks. For example, for a spherically symmetric mask, the best method to estimate the angular averages is to take averages over all the pixels (in Fourier space) with the exact same distance to the middle of the field in Fourier space. Using this method results in a smaller number of points for each ℓ mode, which will in turn result in a very inaccurate mask modelling, for asymmetric masks. In conclusion, the method used in this paper is the most robust approach to mask modelling for a flat-sky analysis. The results presented in this work have all been corrected with $B_{\rm M}$. Nevertheless, we acknowledge that the need for this calibration makes this flat-sky FFT method less desirable than an all-sky PCl method, which does not appear to produce this effect.

Flat-sky PCl methods have been applied to data for CMB temperature and polarization as well as cosmic infrared background anisotropy (Planck Collaboration XVIII 2011). In those previous analyses, a correction factor has been used to account for all the remaining data effects (see e.g. Ponthieu, Grain & Lagache 2011, for POKER algorithm). This factor is estimated using simulations and thus is in principle cosmology dependent. Applying this correction factor to the analysis can hide any residual masking effects. In the current work, we treat this correction separately to understand its effects on flat-sky PCl estimators.

Fig. 6 shows the average estimated and theory PCls and Cls of the lognormal fields for the composite mask, with $n_\ell = 20$ (see Section 2.4 for the details of the ℓ-binning and the definition of n_ℓ). The noise contribution is subtracted from the estimated (pseudo-)power spectra. $C(\ell)^{\rm input}$ and $C(\ell)^{\rm ave}$ are shown in green dashes and magenta solid curve, respectively. As discussed earlier, due to the pixelized nature of the fields, the input power spectrum and its angle averaged version are not identical and show differences mostly at small Fourier modes, which is apparent in the figure as the small ℓ difference between the green dashes and magenta curves. The recovered C_ℓ are shown as magenta squares. We use method I (see equation 43), which results in the closest recovery to the true power spectrum. Since the second recovery method (equation 44) does not provide any advantages (as discussed in Section 2.4), we will not use it in any of the following analyses. The PCls shown in black are rescaled by a factor of $1/\overline{W^2}$, which enables us to see the mode-mixing effects of the mask. Furthermore, the noise contribution is subtracted from the PCls. The E-mode PCls are shown as a solid curve and filled circles while the B-mode PCLs are shown as a dotted curve and empty circles for theory and estimated values, respectively. This figure shows that the composite mask has a large

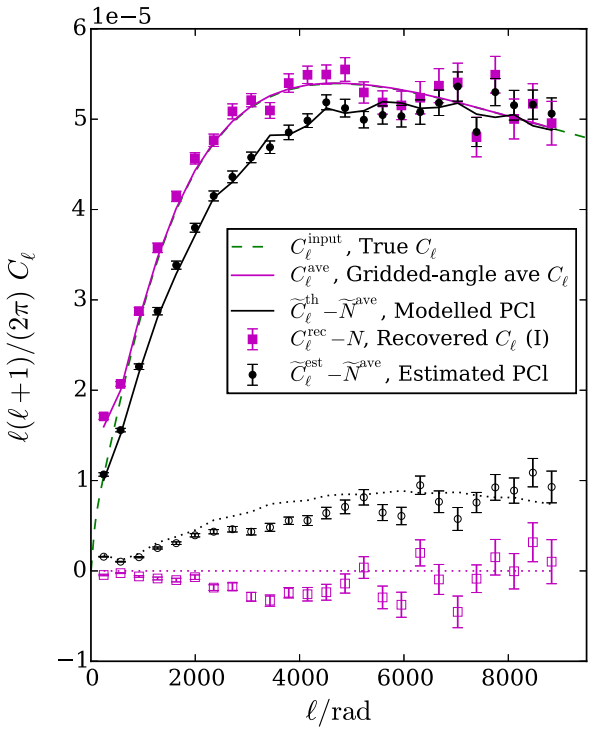

Figure 6. The estimated power spectra from the lognormal simulations and their expected values from theory for the composite mask with no apodization. 20 of the smallest ℓ-bins are merged to make these power spectra ($n_\ell = 20$, $\Delta\ell \approx 360$/rad). The curves show the expected theory values and the symbols show the estimated values from the simulations. The dashed green curve shows the input power spectrum from which the simulations are constructed. The magenta solid curve shows the input power spectrum after angle averaging. The magenta squares belong to the first recovery method defined in equation (43). The solid black curve and filled symbols show the theory and estimated E-mode PCls, while the dotted black curve and empty symbols show the theory and estimated B-mode PCls. The ellipticity noise contribution is subtracted from the results. The error bars correspond to the field-to-field variations between the realizations of the shear fields. Similar plots for other masks are shown in Appendix B.

effect on the PCls over a large range of Fourier modes. Additionally, some of the power is moved to the B-modes. The theory and the estimated values of PCls are fairly consistent for the E-modes, which is not the case for the B-modes, since the B-modes have a smaller signal they are more sensitive to inaccuracies in mask modelling and hence will not be used for the parameter estimation in the next section. Similar plots can be seen in Appendix B for the other masks.

Since the variance on the mean of 100 fields is very small, it is difficult to compare the theory to estimated $\widetilde{C}(\ell)$ or $C(\ell)$ values in Fig. 6. Therefore, the relative power spectra are plotted in Fig. 7 for the composite mask. The grey areas show the cosmic variance for the simulated fields. The cosmic variance for $C(\ell)$ assuming a Gaussian distribution is estimated using

$$\sigma^2_{\rm cosm,\,Cl} = \frac{2}{2\ell+1}\frac{1}{f_{\rm sky}\Delta\ell}(C(\ell)+N_\epsilon)^2, \tag{47}$$

where $f_{\rm sky}$ is the fraction of the sky that is not covered by masks. $\Delta\ell$ is the ℓ-bin width and N_ϵ is the noise power. For the 100 simulated fields of 100 square degrees each,

$$f_{\rm sky} \simeq 0.24\, f_{\rm image}, \tag{48}$$

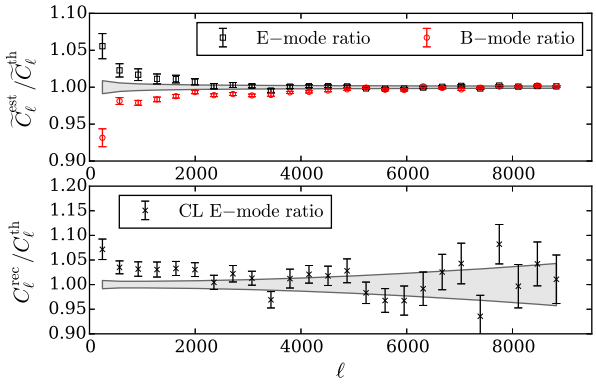

Figure 7. The ratio of estimated to theory power spectra for the composite mask with $n_\ell = 20$, $\Delta_\ell \approx 360$/rad. The top plot shows this ratio for the $\widetilde{C}(\ell)$ and the bottom for $C(\ell)$. The noise contribution is not subtracted from $\widetilde{C}(\ell)$, while it is subtracted from $C_\ell^{\rm rec}$, to recover the input power spectra. The black points show the ratio for the E-mode power spectra, whereas the red ones correspond to the B-modes. The recovered $C(\ell)$ is estimated from the first binning method, for which the theory value of B-modes is zero; hence, this ratio is not shown here. The shaded area shows the expected cosmic variance.

where $f_{\rm image}$ is the effective fraction of each image not covered by the mask. To find the cosmic variance for $\widetilde{C}(\ell)$, we need to use the mixing matrix on the $\langle \Delta C(\ell)^2 \rangle$. Doing so results in some off-diagonal terms that we are not interested in, for the purposes of this section, since we show only the diagonal terms and their associated error bars. The off-diagonal terms are incorporated in the analysis in Section 4.3. The diagonal terms are

$$\sigma^2_{\rm cosm,\,PCl} = \frac{2}{2\ell+1}\frac{1}{f_{\rm sky}\Delta\ell}(\widetilde{C}(\ell)+\widetilde{N}_\epsilon)^2, \tag{49}$$

which is the cosmic variance term for the PCls, where $\widetilde{C}(\ell)$ is their expected value from theory.

The top plot in Fig. 7 shows the ratio of the estimated PCl to its theory value for E-modes and B-modes. Unlike in Fig. 6, the noise contribution is not subtracted from the PCls, since for the forward modelling Fisher analysis in the next section, the PCls used contain noise. As can be seen in the plots, the low-ℓ ratios diverge from unity, which shows an imperfect mask modelling at these scales. The E/B-modes show an almost anti-correlated behaviour on these scales, which suggests that theory PCls do not account for all the mode-mixing and E/B leakage. The bottom plot shows the ratio of $C(\ell)^{\rm rec}$ defined in equation (43) to its theory value, which is equal to the angle averaged input power spectrum, $C(\ell)^{\rm ave}$. Again here we see that the agreement between the theory and estimated values is better at scales above $\ell \sim 2000$. Similar figures for other masks are provided in Appendix B

4 ERROR PROPAGATION

The ultimate goal of a cosmic shear analysis is to constrain cosmological models and their parameters in a typical scenario. Here we use a Fisher analysis to put upper limits on the constraining power of PCls. We use both $\widetilde{C}(\ell)$ (forward modelling) and $C(\ell)$ (backward modelling) to compare the constraints and biases on model parameters.

4.1 Fisher analysis: formalism

The main purpose of this study is to determine the accuracy of a PCl analysis and its limitations. The ratio of the bias on a deduced model parameter to the errors associated with it can give us an indication of the accuracy of such analysis. Ideally, the estimators are unbiased; however, as we have seen throughout this paper, there are sources of bias in a PCls analysis, originating from inaccuracies in mask modelling and binning effects.

Formally, the Fisher matrix is defined as the ensemble average of second derivatives of the negative log-likelihood function at the maximum likelihood point,

$$F_{ij} \equiv \left\langle \frac{\partial^2 \mathfrak{L}}{\partial\phi_i\,\partial\phi_j} \right\rangle, \tag{50}$$

where ϕ_i are the parameters to be inferred.

Assuming a Gaussian likelihood distribution, to calculate the Fisher matrix we can use the following relation,

$$F_{ij} = \langle \mathfrak{L}_{,ij} \rangle = \frac{1}{2}{\rm Tr}[\mathbb{C}^{-1}\,\mathbb{C}_{,i}\,\mathbb{C}^{-1}\,\mathbb{C}_{,j} + \mathbb{C}^{-1}\,M_{ij}], \tag{51}$$

where $\mathbb{C}$ is the data covariance, $\mathbb{C}_{,i}$ is the derivative of $\mathbb{C}$ with respect to ϕ_i, and M_{ij} is a matrix composed of the derivatives of $\boldsymbol{\mu}$, the expected value of the data vector,

$$M_{ij} = \boldsymbol{\mu}_{,i}\,\boldsymbol{\mu}^{\rm t}_{,j} + \boldsymbol{\mu}_{,j}\,\boldsymbol{\mu}^{\rm t}_{,i}, \tag{52}$$

where $\boldsymbol{\mu}^{\rm t}$ is the transpose of $\boldsymbol{\mu}$ (for the details of the derivation of equation 51, see Tegmark, Taylor & Heavens 1997, for example). The second term in equation (51) is dominant for a survey with relatively large area as the covariance matrix is scaled inversely by the area (see Eifler, Schneider & Hartlap 2009; Asgari, Schneider & Simon 2012, where the exclusion of the first term is shown to have a negligible effect). Hence, we will use only the second term to calculate our Fisher matrices.

The Fisher matrix can also be used to propagate the bias in the measured observables to the estimated parameters. Taylor et al. (2007) showed that for a Gaussian-distributed likelihood, the linear bias for a parameter, ϕ_μ, given the bias in the observables, $\mathbf{x}$, is

$$B_\mu = \left(F^{-1}\right)_{\mu\nu}\mu_{i,\nu}\left(\mathbb{C}^{-1}\right)_{ij}\left(\mu_j - x_j\right), \tag{53}$$

where μ_j is the expected value of x_j and Einstein summation rules apply (also see Knox, Scoccimarro & Dodelson 1998; Kim et al. 2004).

4.2 Weak lensing covariance

To find the value for the Fisher matrix and the bias from equations (51) and (53), we need to find the covariance of $C(\ell)$ and $\widetilde{C}(\ell)$. In this work, we use lognormal shear fields as default and compare the final results with Gaussian shear fields. The covariance of a Gaussian field has been calculated in the literature (see Kaiser 1998; Joachimi, Schneider & Eifler 2008, for example). In Appendix C, we show a general calculation for finding the moments of a lognormal field, which is then used to find the covariance of the shear power spectra.

A lognormal distribution provides a more realistic characterization of the convergence field, $\kappa(\theta)$ (see Hilbert, Hartlap & Schneider 2011, and references therein). Therefore, we can use the formalism in Appendix C to estimate the covariance of such a field. The equations in Appendix C are derived for the moments of the density contrast, which has a minimum of -1, unlike $\kappa(\theta)$. Hence, we need to incorporate this difference to find the covariance of a lognormal

$\kappa(\theta)$ field. We can write the lognormal convergence field in terms of a Gaussian field,

$$\kappa(\theta) = e^{n(\theta)} - \kappa_0, \tag{54}$$

where $n(\theta)$ is a Gaussian random field and κ_0 is the absolute value of the minimum convergence (see Hilbert et al. 2011). For this work, the value assumed for κ_0 is 0.012, which corresponds to the value found by Hilbert et al. (2011) for the Millennium simulation with source galaxy redshift of 0.76. Using this definition instead of the one for $\delta^{\rm ln}$ in equation (C1) introduces extra constant coefficients in equation (C11). The final result after applying these changes is shown here.

The covariance matrix of the power spectrum of a lognormal convergence field can be written in terms of the sum of the covariance of a Gaussian field and a purely lognormal term,

$$\begin{aligned}\mathbb{C}^{\rm tot}(\ell,\ell') &\equiv \mathbb{C}^{\rm ln}(\ell,\ell') + \mathbb{C}^{\rm G}(\ell,\ell') \\ &= \langle \hat{C}(\ell)\hat{C}(\ell')\rangle - C(\ell)C(\ell'),\end{aligned} \tag{55}$$

where $C(\ell)$ is the expected value of $\hat{C}(\ell)$. $\mathbb{C}^{\rm G}(\ell,\ell')$ is given in Kaiser (1998) as

$$\mathbb{C}^{\rm G}(\ell,\ell') = \frac{4\pi}{A\ell\Delta\ell}(C(\ell)+N_\epsilon)^2\delta_{\ell\ell'}, \tag{56}$$

where $\delta_{\ell\ell'}$ is the Kronecker delta, which makes this covariance diagonal. N_ϵ is the noise power spectrum given in equation (36). $\Delta\ell$ is the width of the ℓ-bin and A is the area of the field. Note that this a model covariance matrix that assumes a simple survey geometry.

The purely lognormal term, $\mathbb{C}^{\rm ln}(\ell,\ell')$, can be calculated using the purely lognormal terms of the fourth-order lognormal moments. We put $k_1=\ell$, $k_2=-\ell$, $k_3=\ell'$, and $k_4=-\ell'$ in equation (C18); following Hilbert et al. (2011), we ignore all terms but *IV*, which we simplify to find the desired relation,

$$\begin{aligned}\mathbb{C}^{\rm ln}(\ell,\ell') &\simeq \frac{1}{A^2\kappa_0^2}\int\frac{{\rm d}\varphi_\ell}{2\pi}\int\frac{{\rm d}\varphi'_\ell}{2\pi}\langle\hat{\kappa}(\ell)\hat{\kappa}(-\ell)\hat{\kappa}(\ell')\hat{\kappa}(-\ell')\rangle_{IV} \\ &= \frac{1}{A\kappa_0^2}\big\{2[C^2(\ell)C(\ell')+C^2(\ell')C(\ell)] \\ &\quad +[C(\ell)+C(\ell')]^2 \\ &\quad \times\int\frac{{\rm d}\varphi_{\ell\ell'}}{2\pi}[C(|\boldsymbol{\ell}-\boldsymbol{\ell}'|)+C(|\boldsymbol{\ell}+\boldsymbol{\ell}'|)]\big\},\end{aligned} \tag{57}$$

where $\varphi_{\ell\ell'}=\varphi_\ell-\varphi_{\ell'}$ is the angle between $\boldsymbol{\ell}$ and $\boldsymbol{\ell}'$.

To find the binned covariance matrix, we first calculate the covariance matrix for the smallest binning and then apply the binning matrix to it,

$$\mathbb{C}^{\rm tot}_{\rm b} = \mathscr{B}\mathbb{C}^{\rm tot}\mathscr{B}^{\rm t}, \tag{58}$$

where $\mathscr{B}$ is the binning matrix defined in Section 2.4. The covariance matrix for a binned $\widetilde{C}(\ell)$ is then,

$$\widetilde{\mathbb{C}}^{\rm tot}_{\rm b} = \mathscr{B}\mathbf{M}\mathbb{C}^{\rm tot}\mathbf{M}^{\rm t}\mathscr{B}^{\rm t}. \tag{59}$$

The cosmic variance term in equation (47) is basically the same as the diagonal terms in equation (56) with $2\ell+1\approx 2\ell$. The area in equation (56) is the area of the field before zero-padding; hence, here we ignored the masking effect. But in equation (47), we use the effective area of the field. The reason behind this difference is that we use the unmasked covariance on the right-hand side of equation (59) to find the masked one, which accounts for the loss of area. Note that the covariance in equation (56) is an analytic estimate for a Gaussian field, which is simply connected and when all relevant angles are smaller than the extent of the field (see Joachimi et al. 2008). In the next section, we use the area before masking to measure the covariance of the recovered Cl. This is not an accurate representation of this covariance as we are assuming that all the lost information due to masking is recovered. Scaling this covariance with the effective area is not a fair representation either. However, in Table 3, we provide effective area scaling factors that can be used to rescale the covariance values in the following section.

Table 3. The area scaling due to mask. The values in the table correspond to the ratio of the field area after to its area before masking ($f_{\rm sky}/A$).

No Ap	Ap1	Ap2	Ap3	
1.00	1.00	1.00	0.99	No Mask
0.90	0.90	0.90	0.89	Star
0.86	0.70	0.49	0.17	Checkerboard
0.78	0.63	0.43	0.15	Composite

Here we limit our theory and simulation comparison to power spectra and leave a covariance matrix comparison to future work. For our finest binning, we have 500 ℓ-modes. To get an unbiased estimate of the inverse covariance, at least 625 simulations are needed according to Anderson (2003) or ∼4000 for a more accurate estimate (better than 5 per cent, see Hartlap, Simon & Schneider 2007).

4.3 Fisher analysis: results

In this section, we investigate the significance of the possible biases in the estimated cosmological parameter due to a PCl analysis and compare that to the errors on the estimations. We calculate the bias on the parameter estimation from equation (53) for each simulated field by comparing the expected value of the power spectra to their observed value. The mean and standard deviation of the bias are then measured from 100 realizations of the shear field for each set of the simulations. The error on the other hand is calculated analytically using a Fisher analysis (see equation 51) and a theoretical covariance matrix for the power spectra (see Section 4.2) for a single 100 deg^2 field. The error can be scaled by 1/10 to obtain the error for a 10 000-deg^2 field.

We limit our study to two free parameters, σ_8 and $\Omega_{\rm m}$, to demonstrate the validity of the PCl method for weak gravitational lensing. The number of free parameters does not change the main results.

We use a fixed ℓ-mode range [∼(18–9000)rad^{-1}] but in reality a redshift-dependent maximum ℓ-mode should be used in order to create a consistent k-mode selection. This relation is approximately $\ell_{\rm max}=k_{\rm max}r(z)$, where $r(z)$ is the co-moving distance. We choose a maximum ℓ-mode of ∼9000 that corresponds to the largest angular mode that would be probed at the largest co-moving distance for a $k_{\rm max}\simeq 1$. Some surveys might truncate at smaller ℓ, especially due to uncertainty in non-linear modelling. In addition, the S/N of a power spectrum estimator is relatively flat between $\ell=1000$ and $\ell=10\,000$ (see e.g. Sato & Nishimichi 2013). Nevertheless, here we want to demonstrate the constraints from the full ℓ-range, which can be relevant for baryonic physics.

Recall that in forward modelling, we apply the mixing matrix on the theory power spectra, while in the backward modelling, we instead correct the masking effects by applying the inverse mixing matrix to the observed PCls. We use only the E-mode power spectra for this analysis.

Figs 8 –10 show the absolute value of the mean bias, $|\bar{B}|$, the 1 σ error, and their ratio for σ_8, while $\Omega_{\rm m}$ is marginalized over,

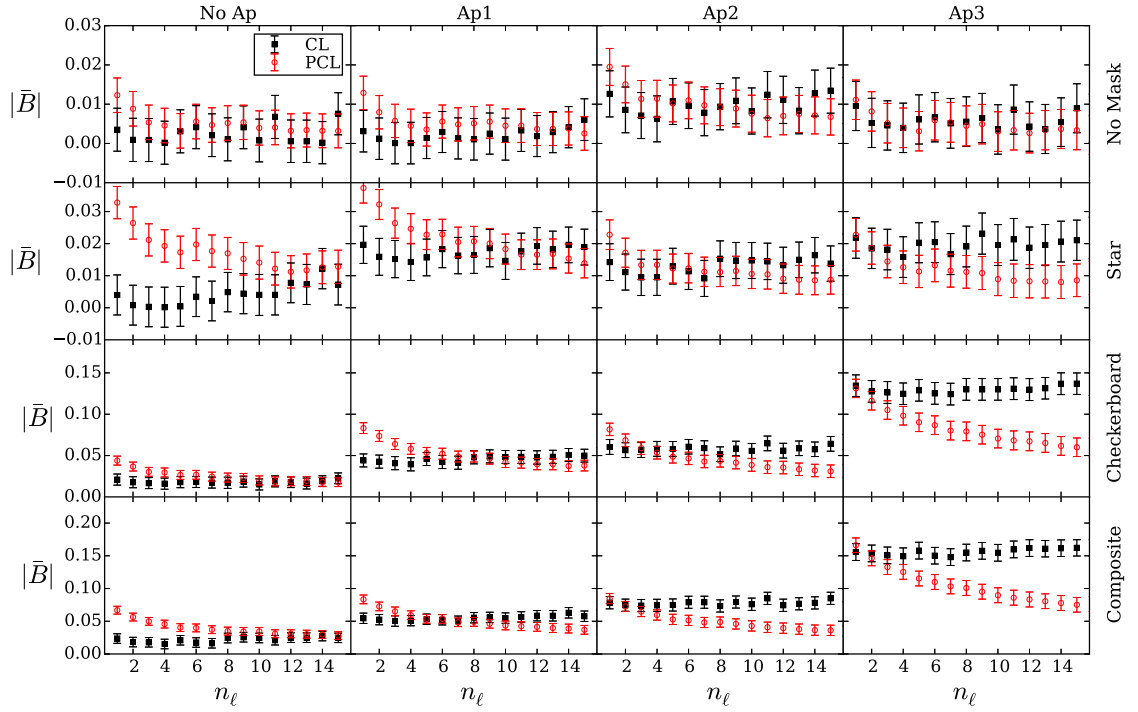

Figure 8. The average bias on the estimated parameter, σ_8, with respect to the bin size. n_ℓ is the number of original ℓ bins that are combined to make the wider bins. All the other cosmological parameters are fixed to their fiducial values in Table 2, aside from Ω_m which is marginalized over. The red empty circles denote the forward modelling scheme, where the PCls are the observables and the theory mixing matrix is applied to the theory power spectra, while the black full squares show the values for the backward modelling, where the mask correction is applied to the data instead of the theory power spectra. Each row belongs to a different mask and each column to a different apodization scheme (see Section 2.3). A Fisher analysis is used for estimating the bias on the estimated parameter. The error bars show the error on the mean estimated from the field-to-field variance between the 100 simulated fields. The lognormal simulations are used here.

respectively. The results for Ω_m are not shown here as they closely follow the ones for σ_8. Each row belongs to a different mask which is named at the right-hand side of the row and each column to a different apodization scheme named at the top of the column. All three figures show the results for all mask and apodization combinations. For the description and definition of the masks and apodization cases, see Sections 2.2 and 2.3. The x-axis in these plots, n_ℓ, is the number of the smallest bins that are combined to make a wider bin; therefore, the width of the corresponding bin is n_ℓ times larger than that of the smallest bin. The smallest bin width is 18/rad. The red empty circles and dashed lines belong to forward modelling, while the black full squares and solid lines show the backward modelling values. In Fig. 10, blue stars show the case where no mask correction, other than a constant area factor, has been applied. This is shown for a comparison with the corrected versions. The σ for the blue stars is the same as the backward modelling case. Note that the σ in Fig. 9 is an analytic calculation which is noise free; however, the bias calculation changes for each field and hence there is a scatter between them which is captured by the error bars in Figs 8 and 10.

By studying the three Figs 8 –10, we can see that for the cases with no apodization, the best method to use is the backward modelling, as the bias is consistently the lowest as well as the bias to error ratio. Note that in Fig. 9, the black solid line remains constant over all the panels as the backward modelling covariance has no information about the mask (see equation 58 and the discussion that follows). On the contrary, the covariance for forward modelling depends on the mask.

The forward modelling bias decreases as n_ℓ and hence the bin width increases in agreement with Asgari & Schneider (2015), who showed that narrower band power spectra are generally more biased. We apply a binning matrix (see equation 42) to our modelling to minimize this effect. The binning works better for the backward modelling as can be seen from the approximately flat behaviour of the bias with respect to n_ℓ in Fig. 8.

A comparison of different masks in Fig. 8 shows that the large-scale mask is generally more difficult to model and results in a larger bias. This effect is more pronounced for the composite mask where all the scales are affected. Note that in all cases in this work, zero padding is present, which affects the small ℓ-modes and propagates through to all modes in the mixing matrix estimation, and this was corrected for earlier by a multiplicative factor which also depends on the other properties of the mask (see Fig. 5). Apodizing the mask increases the area covered by the mask, which in turn results in a more biased estimate for σ_8. In Fig. 8, we see that the apodization affects the backward modelling more than the forward case, especially when the checkerboard mask is present. Ap3, which has the largest kernel out of the three smoothing schemes, has a drastic effect on both modelling schemes for the smallest ℓ-bins, but the PCl values recover after binning. Furthermore, in Fig.10 we see that the ratio of bias to error is hardly affected for forward modelling and seems to improve in contrast to backward modelling. This can be explained by looking at Fig. 9 where we see that the 1 σ error on σ_8 is adjusted in the forward modelling case by the mixing matrix that allows for a lower bias to error ratio.

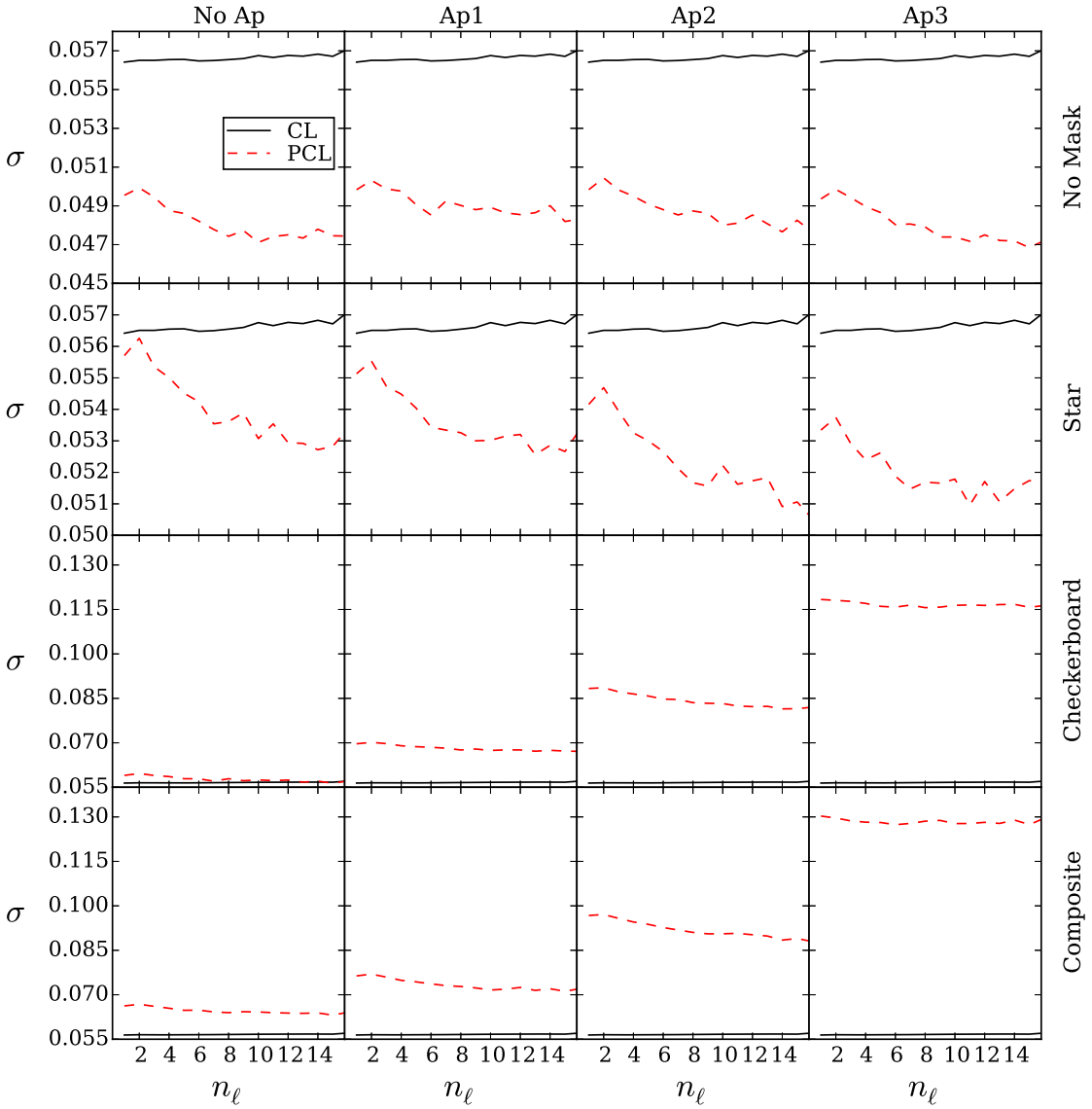

Figure 9. The 1σ error on the estimated parameter, σ_8, with respect to the bin size. n_ℓ is the number of original ℓ-bins that are combined to make the wider bins. All the other cosmological parameters are fixed to their fiducial values in Table 2, aside from Ω_m which is marginalized over. The red dashed line belongs the forward modelling scheme, where the PCls are the observables and the theory mixing matrix is applied to the theory power spectra, while the black solid line shows the values for the backward modelling, where the mask correction is applied to the data instead of the theory power spectra. Each row belongs to a different mask and each column to a different apodization scheme (see Section 2.3). Note that the black solid line remains constant between the different panels, since it is unaffected by the mask. A Fisher analysis is used for estimating the error on the estimated parameter.

The general conclusion from inspecting Figs 8 –10 is that if the data are masked in a binary manner (ones and zeros mask), which is the 'no Ap' case, a backward modelling where the recovered $C(\ell)$ are measured provides a better method, while for a non-binary mask or other effects that can mimic such masks, the forward modelling provides a better choice. For example, the inverse variance weight on shape measurements that has a multiplicative effect on the measured ellipticities can be interpreted as an apodized mask (see Miller et al. 2013, for the definition of the inverse variance weights). If the combination of these weights and the already present masks form a uniform structure in the images, they will resemble the apodized checkerboard mask and hence we expect them to behave similarly.

Fig. 11 summarizes the main conclusions in this section. It shows the Fisher constrains and the linear bias in the Ω_m–σ_8 plane. The ellipses show the 95 per cent confidence regions and they are shifted from the fiducial position according to their bias value. The fiducial position of the parameters is shown as a red x. The results are shown for the composite mask with no apodization and 'Ap2' for backward (recovering C_ℓ) and forward modelling (applying the mixing matrix to the theory). The non-apodized cases are shaded, the dashed black one shows the backward modelling case, while the solid green one belongs to forward modelling. The ellipse sizes do not change for backward modelling as we saw in Fig. 9. The empty ellipses belong to the apodized case with 'Ap2' (see Table 1), where the dotted blue refers to forward and the dashed red to backward modelling. The results are shown for the binning case with $n_\ell = 14$, $\Delta\ell \approx 252$. Aside from the backward modelling for the apodized mask, all the biases are within the 95 per cent contours. The apodization changes the size of the forward modelling ellipse, since apodizing the composite mask results in significant loss of area that is captured by the mixing

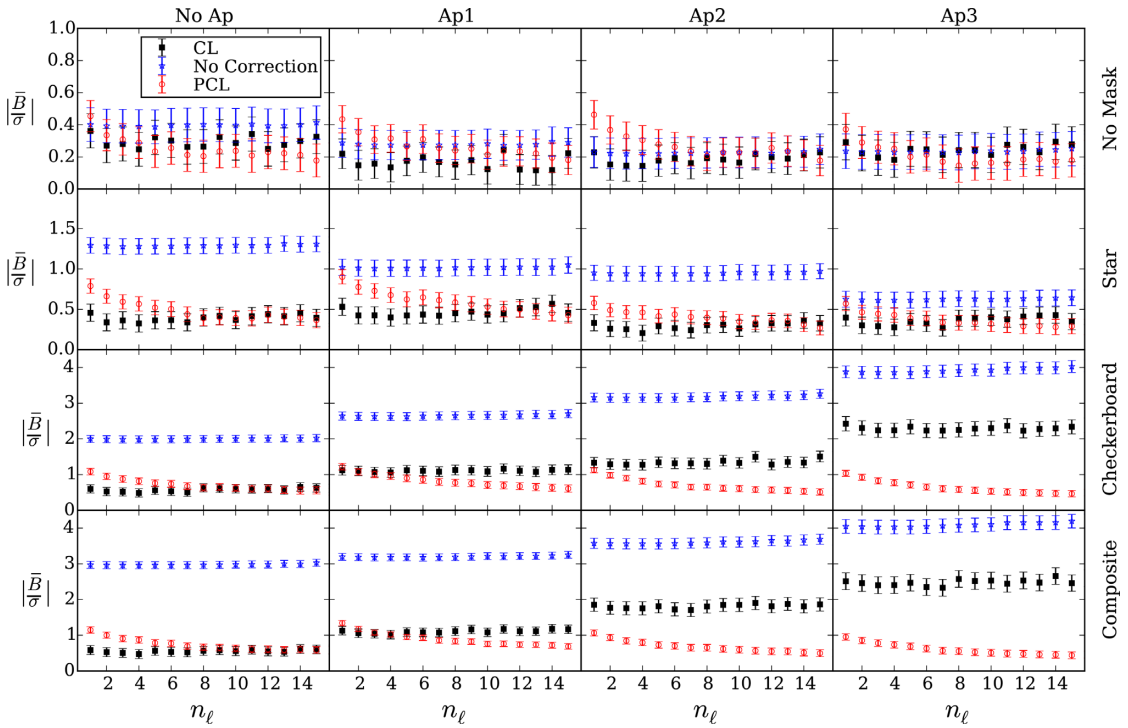

Figure 10. The ratio of the average bias to the 1σ error on the estimated parameter, σ_8, with respect to the bin size. n_ℓ is the number of original ℓ-bins that are combined to make the wider bins. All the other cosmological parameters are fixed to their fiducial values in Table 2, aside from Ω_m which is marginalized over. The red empty circles belong the forward modelling scheme, where the PCls are the observables and the theory mixing matrix is applied to the theory power spectra, while the black solid line shows the values for the backward modelling, where the mask correction is applied to the data instead of the theory power spectra. The blue stars show the ratio for the case where no mask correction is applied to either the theory or the observed power spectra aside from a multiplicative area correction factor. The blue symbols show the level of importance of the mask correction for different masks. Each row belongs to a different mask and each column to a different apodization scheme (see Section 2.3). A Fisher analysis is used for estimating the bias and error on the estimated parameter. The error bars show the error on the mean estimated from the field-to-field variance between the 100 simulated fields. The Gaussian simulations are used here.

matrix (see equation 59). However, the analytical covariance for the backward modelling in equation (58) assumes a simplistic survey geometry and hence is not exact, especially for a heavily masked region. Consequently, the backward modelling shows large bias to error ratios for an apodized mask with regular feature, such as the checkerboard case (see Fig. 1).

5 CONCLUSION

PCl analysis is a method that models the effects of masks on the Fourier transform of a field. It provides a crude but fast, FFT-based direct measurement of the power spectrum in the presence of masks but may require calibration to simulated data. In this paper, we have applied the flat-sky PCl approximation on simulated shear fields to investigate the accuracy and potential biases in this method for weak gravitational lensing analysis. This is particularly interesting for both current surveys, where the flat-sky approximation is used, and for future large-scale surveys such as Euclid[3] and LSST[4] where rapid methods may be useful. However, we note that, given the substantial need for calibration, running a flat-sky analysis on data that covers large parts of the sky is not advantageous as the gain in computational speed would at best be minimal. It remains potentially competitive on small survey patches, e.g. early Euclid data.

Although flat-sky PCl has been used for CMB analysis, it has never been tested to the extent that is done here. Here we show, for the first time, the effects of incomplete mask modelling in a flat-sky implementation of the PCl method on the estimated cosmological parameters.

Masking introduces mixing of Fourier modes in the shear field, and when the masked field is decomposed into E-convergence and B-curl modes, there is mixing between the E/B-modes. In order to forward-model this effect, or recover the all-sky power from a masked field by deconvolution, the mask mixing matrix has to be modelled to a high accuracy. We have shown the need to carefully model the mask mixing matrix, taking into consideration the numerical estimation of integrals on a pixelized field. To investigate the effects of masking, we applied large-area masks, corresponding to the limits of a survey, and small-area sub-masks that would model the presence of star masks in a field, and a checkerboard pattern to model the effects of field-of-view boundaries for a mosaicked survey observing strategy. In a previous study, Hikage et al. (2011) performed an analysis of simulated shear fields with full and flat-sky PCls. However, they considered only small-scale star masks, with periodic boundary conditions, and did not propagate the errors to the measured parameters. Hence, the analysis provided here complements and goes beyond Hikage et al. (2011). To maintain realism and avoid periodic boundary conditions, all the fields used in this work were cut out of larger fields for both sets of Gaussian and lognormal simulated fields. As a result, when analysing the fields we first zero-padded them to their original size.

[3] http://sci.esa.int/euclid/
[4] http://www.lsst.org/lsst/

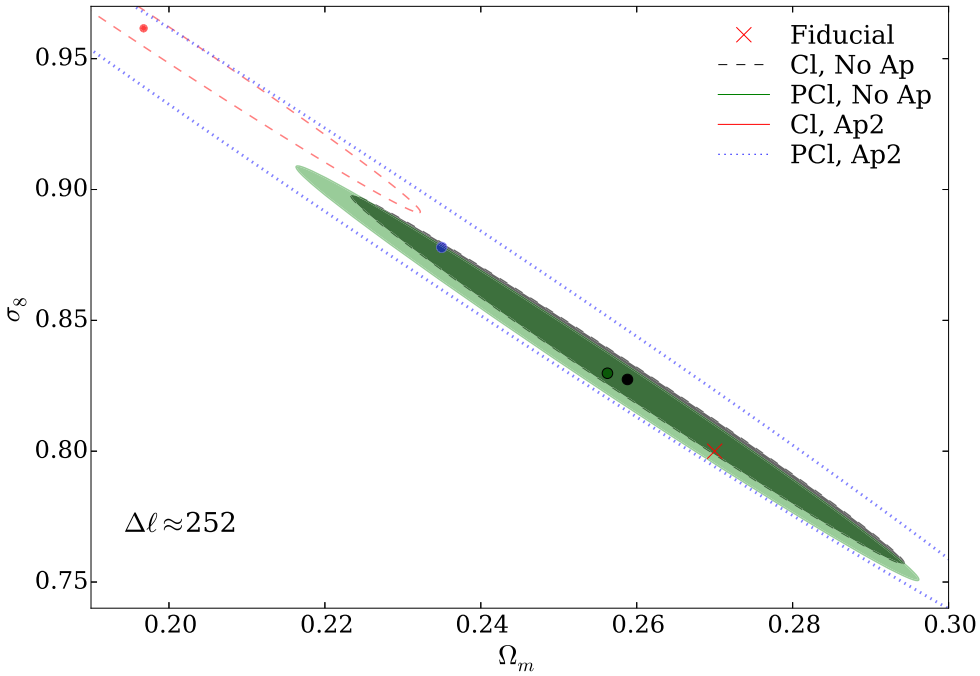

Figure 11. The 95 per cent Fisher contours for four different cases. The fiducial values of σ_8 and $\Omega_{\rm m}$ are marked by a red cross. The ellipses are shifted according to the mean bias over 100 lognormal shear field realizations. The composite mask is considered here, with and without apodization. The empty contours refer to the two apodized cases using Ap2 (see Table 1), while the filled ones belong to the non-apodized masks. The filled dashed (grey) and the empty dashed (red) contours belong to backward modelling, where the power spectrum is recovered, whereas the green filled solid and the blue dashed contours show the results for forward modelling. The power spectra are binned with $n_\ell = 14$, which corresponds to an approximately linear binning with $\Delta\ell \approx 252$. The size of these contours corresponds to a survey of 100 deg^2 and $n_{\rm gal} = 30$/arcmin^2.

Consequently, a large-scale mask was present for all masks in this analysis.

We find that for a flat-sky implementation of the PCl weak lensing power spectrum analysis an overall, constant calibration correction for large-scale masks is required due to the sparsely sampled low-l modes, and rapid oscillations of all modes of the mask power spectrum (see Fig. 3). We have shown that this calibration bias is insensitive to the input power spectrum. For small-area star masks, the forward modelling and all-sky recovery both work well, resulting in slight biases in the lensing power spectra at a few per cent for $l < 2000$. There is also a slight improvement with modest apodization of the mask. For the checkerboard mask, again there is good modelling and all-sky recovery on all scales above $l = 2000$, but the large-angle bias is slightly worse, again a few per cent, which apodization makes worse. While apodization suppresses small-scale mask power, it does not help with the large-scale power and rapidly oscillating power which lead to biases. Indeed, apodization introduces significant biases into the full-sky retrieval, while notably increasing both bias and errors in forward modelling due to the loss of effective sky area.

While investigating the recovered all-sky power spectrum, we found that the choice of binning of the mixing matrix made a significant difference. If we bin the angular wavenumber in wider bins, the mask mixing matrix becomes more diagonal which results in an overall scaling of the masked power spectra. This arises due to the loss of the fine structure in the mask power spectrum by l-binning. However, if we evaluate the mixing matrix per-ℓ, deconvolve, and then re-bin, we preserve the fine tructure of the mask power, and the recovery of the all-sky lensing power is as a result less biased compared to the alternative case. In short, the binning of the mixing matrix should be preferably as fine as possible but the power spectrum needs to be re-binned for a better cosmological parameter estimation.

Overall, following the large-area bias correction, we find that the weak lensing convergence power spectrum can be both forward modelled and the full-sky convergence power reconstructed on scales greater than $l = 2000$ for smaller scale masking composed of both stars and a checkerboard (field-of-view) pattern. On scales less than $l = 2000$, we see a slight residual excess of a few per cent.

Propagating our lensing power spectra into the error and bias on the cosmological parameters, σ_8 and Ω_m, using the Fisher matrix formalism for both Gaussian and lognormal fields, we find an unbiased measurement compared to the expected errors for a simulated survey of 100 deg^2. To estimate the Fisher matrix, the covariance matrix of the shear power is needed, and so in Appendix C we provide a novel algorithm for calculating the moments of a lognormal field that we used to estimate the covariance of the power spectra for the simulated lognormal shear fields.

From our analysis, we typically find a bias of around 1 per cent on parameters, while the error for 100 deg^2 is 3 per cent (with galaxy mean number density of 30/arcmin^2). Our results imply that this approach will remain unbiased for surveys up to 1200 deg^2. However, further study will be needed to improve this for larger surveys where the curvature of the sky will begin to be important.

In summary, we find that we can apply a flat-sky PCl analysis for a masked finite survey with stellar and checkerboard masks and that unbiased forward modelling and all-sky recovered convergence power can be recovered on angular scales above $l = 2000$, with a few per cent residual bias at lower wavenumbers. The flat-sky PCl method requires calibration to simulations to correct for an overall constant bias due to the survey geometry but requires no other calibration. Both forward-modelled and all-sky-recovered power

propagate into a small, percentage bias in measured cosmological parameter, which remain below the statistical accuracy for surveys of less than 1200 deg^2. Given similar results for forward modelling and all-sky recovery, there may be a slight preference for an all-sky recovery, since the forward modelling requires a convolution of the theory power spectrum at each point in parameter space.

Finally, we conclude that a flay-sky PCl method is suitable for the current generation of Weak Lensing surveys but would be unsuited to surveys with high galaxy number density that are larger than 1200 deg^2, such as 15 000deg^2 Euclid. The main bias seems to come from the large-area treatment of the survey geometry, where curved-sky effects will also become important. This could be alleviated with an all-sky spherical harmonic treatment. Given this can be slower, as FFT methods can be used only in the azimuthal directions, it may be that a hybrid method using flat-sky PCls on small scales could be optimal.

ACKNOWLEDGEMENTS

We thank Alan Heavens and Catherine Heymans for many great discussions, Lee Whittaker for discussions about curved sky and flat-sky analysis, and an anonymous referee for their constructive criticism. MA acknowledges support from the European Research Council, Scottish Universities Physics Alliance and Euclid. ANT thanks the Royal Society for a Wolfson Research Merit Award. BJ acknowledges support by an STFC Ernest Rutherford Fellowship, grant reference ST/J004421/1. TDK is supported by a Royal Society University Research Fellowship.

REFERENCES

Alsing J., Heavens A., Jaffe A. H., Kiessling A., Wandelt B., Hoffmann T., 2016b, MNRAS, 455, 4452
Alsing J., Heavens A. F., Jaffe A. H., Kiessling A., Wandelt B., Hoffmann T., 2016a, MNRAS, 455, 4452
Anderson T. W., 2003, An Introduction to Multivariate Statistical Analysis. Wiley, New York, NY
Asgari M., Heymans C., Blake C., Harnois-Deraps J., Schneider P., Van Waerbeke L., 2017, MNRAS, 464, 1676
Asgari M., Schneider P., 2015, A&A, 578, A50
Asgari M., Schneider P., Simon P., 2012, A&A, 542, A122
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Becker M. R. et al., 2016, Phys. Rev. D, 94, 022002
Brown M. L., Castro P. G., Taylor A. N., 2005, MNRAS, 360, 1262
Brown M. L., Taylor A. N., Bacon D. J., Gray M. E., Dye S., Meisenheimer K., Wolf C., 2003, MNRAS, 341, 100
Brown M. L. et al., 2009, ApJ, 705, 978
Chon G., Challinor A., Prunet S., Hivon E., Szapudi I., 2004, MNRAS, 350, 914
Eifler T., Schneider P., Hartlap J., 2009, A&A, 502, 721
Eisenstein D. J., Hu W., 1998, ApJ, 496, 605
Elsner F., Leistedt B., Peiris H. V., 2017, MNRAS, 465, 1847
Erben T. et al. , 2013, MNRAS, 433, 2545
Gradshteyn I., Ryzhik I., 1994, Table of Integrals, Series, and Products . Academic Press
Hand N. et al., 2015, Phys. Rev. D, 91, 062001
Harnois-Déraps J. et al., 2016, MNRAS, 460, 434
Hartlap J., Simon P., Schneider P., 2007, A&A, 464, 399
Hikage C., Oguri M., 2016, MNRAS, 462, 1359
Hikage C., Takada M., Hamana T., Spergel D., 2011, MNRAS, 412, 65
Hilbert S., Hartlap J., Schneider P., 2011, A&A, 536, A85
Hivon E., Górski K. M., Netterfield C. B., Crill B. P., Prunet S., Hansen F., 2002, ApJ, 567, 2
Huff E. M., Eifler T., Hirata C. M., Mandelbaum R., Schlegel D., Seljak U., 2014, MNRAS, 440, 1322
Hu W., 1999, ApJ, 522, L21
Hu W., White M., 2001, ApJ, 554, 67
Joachimi B., Schneider P., Eifler T., 2008, A&A, 477, 43
Kaiser N., 1998, ApJ, 498, 26
Kilbinger M., 2015, Rep. Prog. Phys., 78, 086901
Kilbinger M., Schneider P., 2004, A&A, 413, 465
Kilbinger M. et al., 2013, MNRAS, 430, 2200
Kim A. G., Linder E. V., Miquel R., Mostek N., 2004, MNRAS, 347, 909
Kitching T. D. et al., 2012, MNRAS, 423, 3163
Kitching T. D. et al. 2014, MNRAS, 442, 1326
Knox L., Scoccimarro R., Dodelson S., 1998, Phys. Rev. Lett., 81, 2004
Köhlinger F., Viola M., Valkenburg W., Joachimi B., Hoekstra H., Kuijken K., 2016, MNRAS, 456, 1508
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Memari Y., 2009, Cosmological Parameter Estimation with QUaD CMB Polarization and Temperature Experiment. Edinburgh Research Archive, University of Edinburgh
Miller L. et al., 2013, MNRAS, 429, 2858
Planck Collaboration XVI, 2014, A&A, 571, A16
Planck Collaboration XVIII, 2011, A&A, 536, A18
Ponthieu N., Grain J., Lagache G., 2011, A&A, 535, A90
Sato M., Nishimichi T., 2013, Phys. Rev. D, 87, 123538
Schneider P., Eifler T., Krause E., 2010, A&A, 520, A116
Seljak U., 1998, ApJ, 506, 64
Smith R. E. et al., 2003, MNRAS, 341, 1311
Szapudi I., Prunet S., Pogosyan D., Szalay A. S., Bond J. R., 2001, ApJ, 548, L115
Taylor A. N., Kitching T. D., Bacon D. J., Heavens A. F., 2007, MNRAS, 374, 1377
Tegmark M., Taylor A. N., Heavens A. F., 1997, ApJ, 480, 22
Van Waerbeke L., White M., Hoekstra H., Heymans C., 2006, Astropart. Phys., 26, 91

APPENDIX A: MIXING MATRIX

In Section 2.2, we skipped some of the steps in calculating the mixing matrix. Here we show the details of the formalism (based on Memari 2009). The pseudo power spectrum can be written for EE, EB, and BB correlations of the kappa map. In Section 2.2, we started from equation (21), which shows the estimator used for the power spectrum, given the convergence on a finite patch of sky, and derived equation (24) that connects PCls to the underlying convergence maps. Then we applied the ensemble averages to $\kappa_{\rm E,\,B}$ and used equation (21) to find a relation between the PCls and Cls. Here we write equation (25) for all combinations of the convergence maps,

$$\left\langle \widetilde{C}^{\rm EE}(\ell)\right\rangle = \frac{1}{A}\int\frac{{\rm d}\varphi_\ell}{2\pi}\int\frac{{\rm d}^2\ell'}{(2\pi)^2}|W(\boldsymbol{\ell}-\boldsymbol{\ell}')|^2\left\{C^{\rm EE}(\ell')\cos^2 2\varphi_{\ell\ell'}+[C^{\rm EB}(\ell')+C^{\rm BE}(\ell')]\sin 2\varphi_{\ell\ell'}\cos 2\varphi_{\ell\ell'}+C^{\rm BB}(\ell')\sin^2 2\varphi_{\ell\ell'}\right\},$$

$$\left\langle \widetilde{C}^{\rm EB}(\ell)\right\rangle = \frac{1}{A}\int\frac{{\rm d}\varphi_\ell}{2\pi}\int\frac{{\rm d}^2\ell'}{(2\pi)^2}|W(\boldsymbol{\ell}-\boldsymbol{\ell}')|^2\left\{C^{\rm EB}(\ell')\cos^2 2\varphi_{\ell\ell'}+[C^{\rm BB}(\ell')-C^{\rm EE}(\ell')]\sin 2\varphi_{\ell\ell'}\cos 2\varphi_{\ell\ell'}-C^{\rm BE}(\ell')\sin^2 2\varphi_{\ell\ell'}\right\},$$

$$\left\langle \widetilde{C}^{\rm BB}(\ell)\right\rangle = \frac{1}{A}\int\frac{{\rm d}\varphi_\ell}{2\pi}\int\frac{{\rm d}^2\ell'}{(2\pi)^2}|W(\boldsymbol{\ell}-\boldsymbol{\ell}')|^2\left\{C^{\rm BB}(\ell')\cos^2 2\varphi_{\ell\ell'}-[C^{\rm BE}(\ell')+C^{\rm EB}(\ell')]\sin 2\varphi_{\ell\ell'}\cos 2\varphi_{\ell\ell'}+C^{\rm EE}(\ell')\sin^2 2\varphi_{\ell\ell'}\right\}, \tag{A1}$$

where $\varphi_{\ell\ell'}=\varphi_\ell-\varphi_{\ell'}$. Note that we have dropped $S(\ell)$ here as it is not used in the current work. Since $C(\ell)$ do not have angular dependencies unlike $W(\boldsymbol{\ell}-\boldsymbol{\ell}')$ and the trigonometric functions, we can take the integrals over φ and $\varphi^{'}$ separately. To do so, we first write,

$$|W(\boldsymbol{\ell}-\boldsymbol{\ell}')|^2 = 2\pi\int_0^\infty {\rm d}LL\int_0^{2\pi}\frac{{\rm d}\varphi_L}{2\pi}|W(\boldsymbol{L})|^2\delta_{\rm D}\big(\boldsymbol{L}-(\boldsymbol{\ell}-\boldsymbol{\ell}')\big)\,, \tag{A2}$$

where $\delta_{\rm D}$ is the Dirac delta function, which we exchange with its integral form,

$$\delta_{\rm D}\big(\boldsymbol{L}-(\boldsymbol{\ell}-\boldsymbol{\ell}')\big)=\int\frac{{\rm d}^2\theta}{(2\pi)^2}{\rm e}^{-{\rm i}\boldsymbol{\theta}.(\boldsymbol{L}-(\boldsymbol{\ell}-\boldsymbol{\ell}'))}\,, \tag{A3}$$

and define

$$W_{\gamma\gamma}(L)\equiv\int_0^{2\pi}\frac{{\rm d}\varphi_L}{2\pi}|W(\boldsymbol{L})|^2\,. \tag{A4}$$

Substituting for $|W(\boldsymbol{\ell}-\boldsymbol{\ell}')|^2$ in equation (A1) leads to

$$\begin{aligned}
\left\langle \widetilde{C}^{\rm EE}(\ell)\right\rangle &= \frac{1}{A}\int_0^\infty\frac{{\rm d}\ell'\ell'}{(2\pi)^2}\int_{|\ell-\ell'|}^{|\ell+\ell'|}{\rm d}LLW_{\gamma\gamma}(L)\left(\int{\rm d}^2\theta{\rm e}^{-{\rm i}\boldsymbol{\theta}.\boldsymbol{L}}\right)\\
&\quad\times\int_0^{2\pi}\frac{{\rm d}\varphi_\ell}{2\pi}\int_0^{2\pi}\frac{{\rm d}\varphi_{\ell'}}{2\pi}{\rm e}^{{\rm i}\boldsymbol{\theta}.\boldsymbol{\ell}}{\rm e}^{-{\rm i}\boldsymbol{\theta}.\boldsymbol{\ell}'}\left\{C^{\rm EE}(\ell')\cos^2 2\varphi_{\ell\ell'}+[C^{\rm EB}(\ell')+C^{\rm BE}(\ell')]\sin 2\varphi_{\ell\ell'}\cos 2\varphi_{\ell\ell'}+C^{\rm BB}(\ell')\sin^2 2\varphi_{\ell\ell'}\right\},\\
\left\langle \widetilde{C}^{\rm EB}(\ell)\right\rangle &= \frac{1}{A}\int_0^\infty\frac{{\rm d}\ell'\ell'}{(2\pi)^2}\int_{|\ell-\ell'|}^{|\ell+\ell'|}{\rm d}LLW_{\gamma\gamma}(L)\left(\int{\rm d}^2\theta{\rm e}^{-{\rm i}\boldsymbol{\theta}.\boldsymbol{L}}\right)\\
&\quad\times\int_0^{2\pi}\frac{{\rm d}\varphi_\ell}{2\pi}\int_0^{2\pi}\frac{{\rm d}\varphi_{\ell'}}{2\pi}{\rm e}^{{\rm i}\boldsymbol{\theta}.\boldsymbol{\ell}}{\rm e}^{-{\rm i}\boldsymbol{\theta}.\boldsymbol{\ell}'}\left\{C^{\rm EB}(\ell')\cos^2 2\varphi_{\ell\ell'}+[C^{\rm BB}(\ell')-C^{\rm EE}(\ell')]\sin 2\varphi_{\ell\ell'}\cos 2\varphi_{\ell\ell'}-C^{\rm BE}(\ell')\sin^2 2\varphi_{\ell\ell'}\right\},\\
\left\langle \widetilde{C}^{\rm BB}(\ell)\right\rangle &= \frac{1}{A}\int_0^\infty\frac{{\rm d}\ell'\ell'}{(2\pi)^2}\int_{|\ell-\ell'|}^{|\ell+\ell'|}{\rm d}LLW_{\gamma\gamma}(L)\left(\int{\rm d}^2\theta{\rm e}^{-{\rm i}\boldsymbol{\theta}.\boldsymbol{L}}\right)\\
&\quad\times\int_0^{2\pi}\frac{{\rm d}\varphi_\ell}{2\pi}\int_0^{2\pi}\frac{{\rm d}\varphi_{\ell'}}{2\pi}{\rm e}^{{\rm i}\boldsymbol{\theta}.\boldsymbol{\ell}}{\rm e}^{-{\rm i}\boldsymbol{\theta}.\boldsymbol{\ell}'}\left\{C^{\rm BB}(\ell')\cos^2 2\varphi_{\ell\ell'}-[C^{\rm BE}(\ell')+C^{\rm EB}(\ell')]\sin 2\varphi_{\ell\ell'}\cos 2\varphi_{\ell\ell'}+C^{\rm EE}(\ell')\sin^2 2\varphi_{\ell\ell'}\right\}.
\end{aligned} \tag{A5}$$

The angular dependence of the integral in parentheses can be taken separately to yield

$$\int{\rm d}^2\theta{\rm e}^{-{\rm i}\boldsymbol{\theta}.\boldsymbol{L}}=2\pi\int_0^\infty{\rm d}\theta\theta J_0(L\theta)\,, \tag{A6}$$

where J_0 is the zeroth-order Bessel function of the first kind. We can now take the integrals over φ_ℓ and φ'_ℓ, using the following relations,

$$J_n(x)=\frac{1}{2\pi{\rm i}^n}\int_0^{2\pi}{\rm d}\varphi\,{\rm e}^{{\rm i}x\cos\varphi}{\rm e}^{{\rm i}n\cos\varphi}\,,\qquad J_{-n}(x)=(-1)^nJ_n(x)\,, \tag{A7}$$

and

$$\cos\varphi=\frac{{\rm e}^{{\rm i}\varphi}+{\rm e}^{-{\rm i}\varphi}}{2}\,,\qquad \sin\varphi=\frac{{\rm e}^{{\rm i}\varphi}-{\rm e}^{-{\rm i}\varphi}}{2i}\,, \tag{A8}$$

which result in these equations,

$$\int_0^{2\pi}\tfrac{{\rm d}\varphi_\ell}{2\pi}\int_0^{2\pi}\tfrac{{\rm d}\varphi_{\ell'}}{2\pi}{\rm e}^{{\rm i}\theta\ell\cos\varphi_\ell}{\rm e}^{-{\rm i}\theta\ell'\cos\varphi_{\ell'}}\cos^2 2\varphi_{\ell\ell'}=\tfrac{1}{2}\left[J_0(\ell\theta)J_0(\ell'\theta)+J_4(\ell\theta)J_4(\ell'\theta)\right]\,, \tag{A9}$$

$$\int_0^{2\pi}\tfrac{{\rm d}\varphi_\ell}{2\pi}\int_0^{2\pi}\tfrac{{\rm d}\varphi_{\ell'}}{2\pi}{\rm e}^{{\rm i}\theta\ell\cos\varphi_\ell}{\rm e}^{-{\rm i}\theta\ell'\cos\varphi_{\ell'}}\sin^2 2\varphi_{\ell\ell'}=\tfrac{1}{2}\left[J_0(\ell\theta)J_0(\ell'\theta)-J_4(\ell\theta)J_4(\ell'\theta)\right]\,, \tag{A10}$$

$$\int_0^{2\pi}\tfrac{{\rm d}\varphi_\ell}{2\pi}\int_0^{2\pi}\tfrac{{\rm d}\varphi_{\ell'}}{2\pi}{\rm e}^{{\rm i}\theta\ell\cos\varphi_\ell}{\rm e}^{-{\rm i}\theta\ell'\cos\varphi_{\ell'}}\cos 2\varphi_{\ell\ell'}\sin 2\varphi_{\ell\ell'}=0\,, \tag{A11}$$

where J_4 is the fourth-order Bessel function of the first kind. After these simplifications, we are left with combinations of three Bessel functions that are the only functions that depend on θ. We then find analytic solutions for the integrals over θ using,

$$\int_0^\infty \mathrm{d}\theta\, \theta J_0(L\theta) J_n(\ell\theta) J_n(\ell'\theta) = \frac{\cos n\eta}{\pi \ell\ell' \sin\eta}\,, \tag{A12}$$

where η is the angle between ℓ and ℓ' (see Gradshteyn & Ryzhik 1994). Note that by definition $\boldsymbol{L} = \boldsymbol{\ell} - \boldsymbol{\ell}'$ (see equation A2), which means that they form a triangle of area $\frac{1}{2}\ell\ell' \sin\eta$. Substituting for the θ, φ_ℓ, and $\varphi_{\ell'}$ integrals in equation (A5), we find

$$\begin{aligned}
\left\langle \widetilde{C}^{\mathrm{EE}}(\ell) \right\rangle &= \frac{1}{A}\int_0^\infty \frac{\mathrm{d}\ell'\ell'}{(2\pi)^2}\int_0^\pi \mathrm{d}\eta\, W_{\gamma\gamma}(L)\left\{(1+\cos 4\eta)C^{\mathrm{EE}}(\ell') + (1-\cos 4\eta)C^{\mathrm{BB}}(\ell')\right\}\,,\\
\left\langle \widetilde{C}^{\mathrm{EB}}(\ell) \right\rangle &= \frac{1}{A}\int_0^\infty \frac{\mathrm{d}\ell'\ell'}{(2\pi)^2}\int_0^\pi \mathrm{d}\eta\, W_{\gamma\gamma}(L)(2\cos 4\eta)C^{\mathrm{EB}}(\ell')\,,\\
\left\langle \widetilde{C}^{\mathrm{BB}}(\ell) \right\rangle &= \frac{1}{A}\int_0^\infty \frac{\mathrm{d}\ell'\ell'}{(2\pi)^2}\int_0^\pi \mathrm{d}\eta\, W_{\gamma\gamma}(L)\left\{(1-\cos 4\eta)C^{\mathrm{EE}}(\ell') + (1+\cos 4\eta)C^{\mathrm{BB}}(\ell')\right\}\,,
\end{aligned} \tag{A13}$$

where we used $L^2 = \ell^2 + \ell'^2 - 2\ell\ell'\cos\eta$ to replace $\mathrm{d}LL/(\ell\ell'\sin\eta)$ with $\mathrm{d}\eta$. Equation (A13) shows that the EB power spectrum does not mix with EE and BB. Consequently, we ignored this term in Section 2.2. Note that the above calculations are accurate for an idealistic case where all the angles are available. The mixing matrix is then formed directly from the above equations.

APPENDIX B: POWER SPECTRUM PLOTS

The $C(\ell)$ and $\widetilde{C}(\ell)$ plots for the composite mask were shown in Fig. 6. Similar plots for all the masks are shown in this appendix, including a control case without a mask. All the cases are zero-padded before the measurements, even the control case. In addition, we show plots of estimated to theory ratios for both the recovered $C(\ell)$ and $\widetilde{C}(\ell)$ for all the masks. These plots were shown for the composite mask in Fig. 7.

In total, we have four mask types: 'No Mask', 'Star', 'Checkerboard', and 'Composite', as well as four types of apodization: ' No Ap', 'Ap1', 'Ap2', and 'Ap3'. 'No Ap' means no apodization was applied and the rest of the apodization options are explained in Section 2.3. The masks are discussed in Section 2.2.

Fig. B1 shows the estimated and theory values of $C(\ell)$ and $\widetilde{C}(\ell)$ for all the mask configurations with a wide binning ($n_\ell = 20$, $\Delta\ell \approx 360$). In Fig. B1, the magenta curves and squares show the theory and recovered $C(\ell)$, while the black curves and circles show the theory and estimated $\widetilde{C}(\ell)$ (explained in more detail in the caption). The control case with no masks shows very little difference between the $C(\ell)$ and the $\widetilde{C}(\ell)$ for most scales. The largest difference is at very large and very small scales. At very large scales, the difference is due to the zero-padding, which effectively acts as a large-scale mask. The small-scale differences appear at the scales where noise is dominant ($\ell \gtrsim 5000$). Additionally, we see more fluctuations at these scales for $\widetilde{C}(\ell)^{\mathrm{est}}$ and $C(\ell)^{\mathrm{rec}}$ as expected. The apodization has very little effect on the no mask case.

The second row of Fig. B1 shows the results for the star mask, i.e. small circular masks. Looking at the left most plot in this row, the 'No Ap' case, we see that the overall effect of the star mask is to lower the amplitude of the small-to-mid-range E-mode $\widetilde{C}(\ell)$ by shifting the power to B-modes. Apodization moves the ℓ-modes at which the leakage takes place, which is directly related to the size of the smoothing kernel (a larger kernel stops leakage for a larger range of ℓ). As a result, the Star and Ap3 plot resembles the control plots.

The third row of Fig. B1 shows the results for the checkerboard mask, i.e. large CCD patterns. Similar to the star mask, the overall effect is a leakage of E-modes into B-modes that decreases the amplitude of the E-mode $\widetilde{C}(\ell)$. However, the regular patterns in this mask produce structures in the $\widetilde{C}(\ell)$. Apodization makes the resulting $\widetilde{C}(\ell)$ smoother at large ℓ and pushes the structures to smaller ℓ-modes. The mask modelling fails to capture the structures accurately in the presence of apodization, as can be seen in the plots.

Finally, the last row of Fig. B1 shows the results for the composite mask. As both components of this mask reduce the amplitude of the E-modes by moving the power to B-modes, the effect is more pronounced for the composite. The structures of the checkerboard mask can also be seen here. The B-mode modelling is poor at certain scales for this mask. These scales are pushed to lower ℓ-modes with apodization.

In order to objectively investigate the mask modelling and its limitations, we need to look at ratio plots rather than Fig. B1, since the error bars are very small which makes any judgement from Fig. B1 difficult. Hence, Fig. B2 shows the ratio of the estimated $\widetilde{C}(\ell)$ to its theory value for all the mask and apodization configurations. The y-range here is different from that in Fig. 7 for better inspection. The noise contribution is not subtracted from $\widetilde{C}(\ell)$. The plots are shown for the lognormal fields and the largest binning ($\ell \gtrsim 5000$). The red circles show the ratio of the B-modes and the black squares the E-modes. The grey shaded area shows the expected cosmic variance contribution for each case (see equation 49). The control case with no masking shows discrepancies at small ℓ between the estimated and the theory $\widetilde{C}(\ell)$ specially for the B-modes. However, the rest is within the cosmic variance band. The star mask cases also show a similar behaviour, whereas the checkerboard mask pushes the small ℓ discrepancies to larger values. The apodized checkerboard mask covers a much larger area of the field compared to the non-apodized version, specially for Ap3. Consequently, the cosmic variance increases rapidly for this mask with larger smoothing kernels. The discrepancies seen for the star and checkerboard masks add up for the composite mask.

Fig. B3 shows the ratio of the $C(\ell)^{\mathrm{rec}}$ in equation (43) (method I) to the input $C(\ell)^{\mathrm{ave}}$. The noise contribution has been subtracted from the recovered $C(\ell)$. The grey shaded area shows the cosmic variance centred at 1. The No Mask row used as the control case shows a good agreement between the theory and recovered values. The Star cases show a similar behaviour except for a slightly overestimated recovery of the lowest ℓ-mode. The checkerboard cases show disagreements up to $\ell \approx 2000$, which can also be seen for the composite mask.

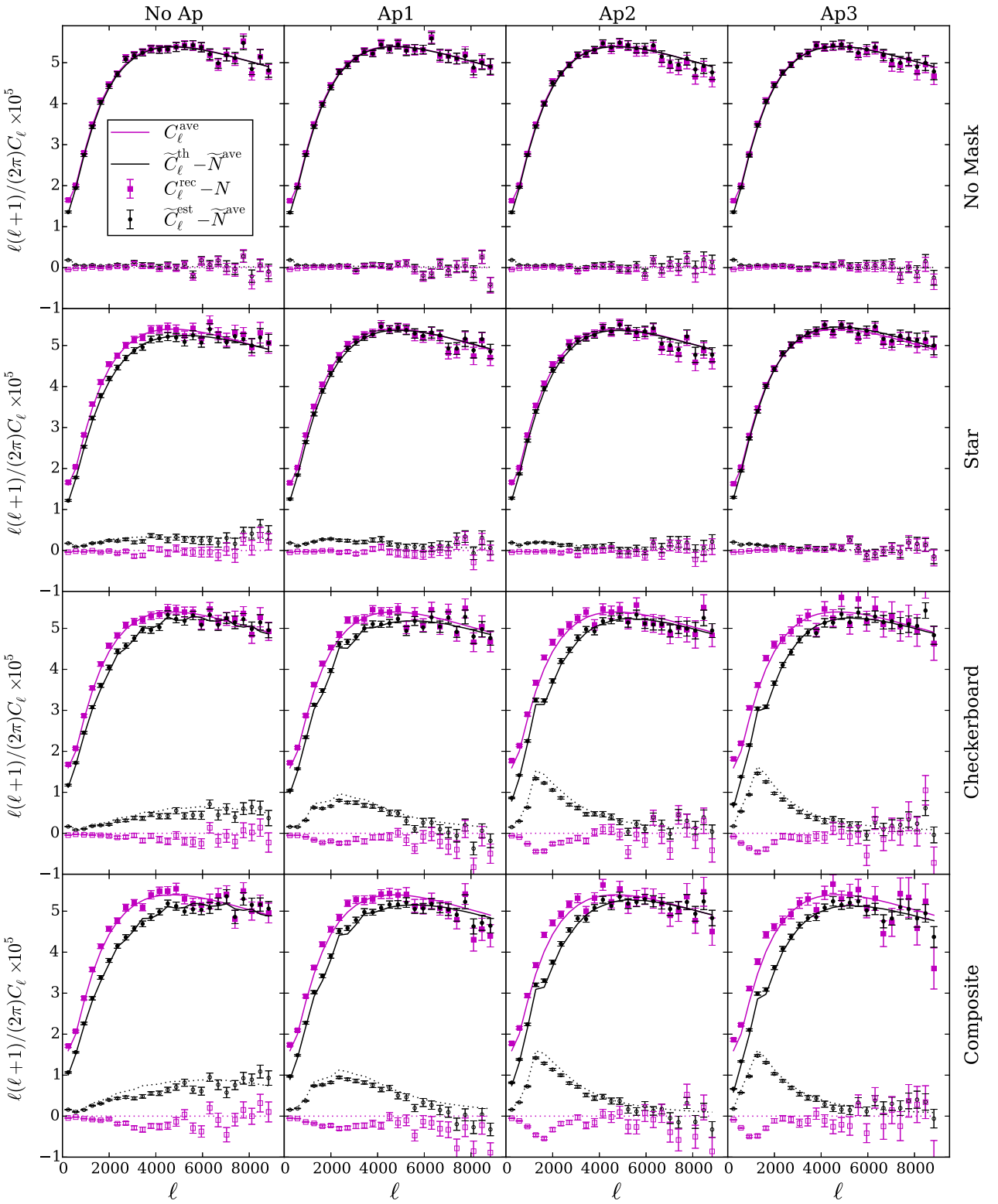

Figure B1. $C(\ell)$ and $\widetilde{C}(\ell)$ plots for all masks and apodization configurations. The magenta solid curve shows the input angle-averaged power spectrum. The magenta solid and open squares show the recovered $C(\ell)$ from method 1 (see equation 43) for E/B-modes, respectively. The solid and dotted black curves show the theory values of the $\widetilde{C}(\ell)$ for E/B-modes, while the black solid and open circles show their estimated value from the simulated fields. The noise contribution is subtracted here. The error bars show the variance of the mean of the 100 fields. The columns show the apodization used, whereas the rows show the mask type used for each plot. 'No Ap' and 'No Mask' mean no apodization and no mask was used. The first row is used as the control case.

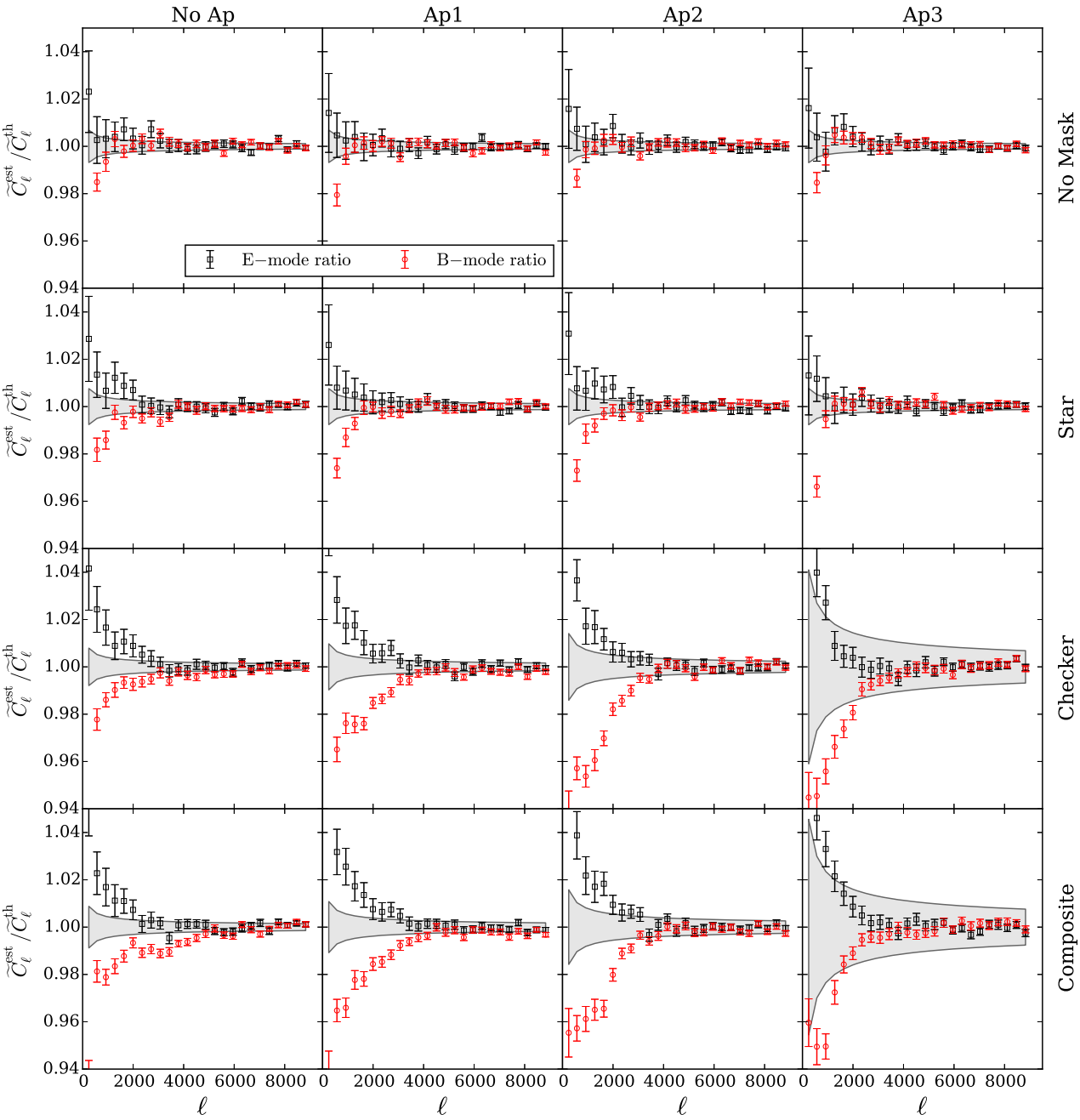

Figure B2. $\widetilde{C}(\ell)$ ratio plots for all masks and apodization configurations. The black squares show the ratio for the E-modes, while the red circles belong to B-modes. The grey shaded area shows the expected cosmic variance centred at 1. We expect to find a good agreement between the estimated and the theory values within the cosmic variance band if the mask modelling is accurate. The error bars show the variance of the mean and are estimated from the simulations.

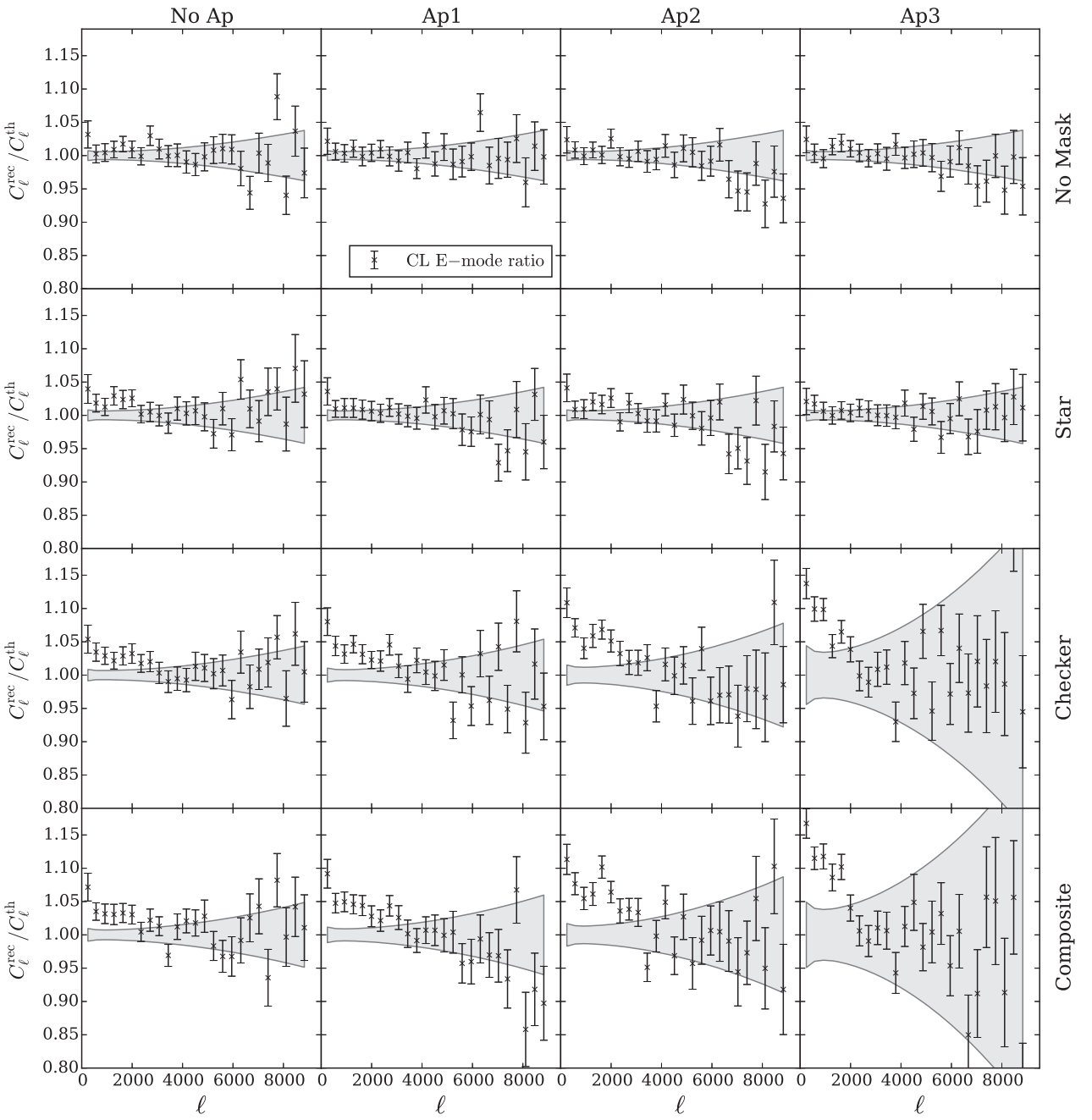

Figure B3. $C(\ell)$ ratio plots for all masks and apodization configurations. The symbols show the ratios for the recovered E-mode $C(\ell)$ to the input angle averaged $C(\ell)$. The grey area shows the cosmic variance contribution, which widens with an increased masked area. The error on the mean is estimated from the field-to-field variance of 100 lognormal simulations.

APPENDIX C: LOGNORMAL MOMENTS

We can find the moments of a lognormal field using its relation to a Gaussian field. These moments will be used to calculate the covariance matrix of the power spectrum of the lognormal realization, which will then be used to estimate the Fisher matrices. In this Appendix, we will show how all the moments of a lognormal field can be written in terms of its power spectrum by taking the following steps.

A lognormal field, $\delta^{\rm ln}$, is defined with respect to a Gaussian field, δ, with zero mean, as

$$\delta^{\rm ln}(\mathbf{x}) \equiv e^{\delta(\mathbf{x})-\sigma^2/2} - 1\,, \tag{C1}$$

where σ^2 is the variance of the Gaussian field. This implies that the variance of the lognormal field is $\exp\sigma^2$. To find the moments of this lognormal field, we take the following steps. In Fourier space, the Nth lognormal moment for a non-zero $\mathbf{k}_i$ can be written as

$$\left\langle \prod_i^N \hat{\delta}^{\rm ln}(\mathbf{k}_i) \right\rangle = \left\langle \prod_i^N [(2\pi)^n \delta_{\rm D}(\mathbf{k}_i) + \hat{\delta}^{\rm ln}(\mathbf{k}_i)] \right\rangle , \tag{C2}$$

where $\delta_{\rm D}(\mathbf{k}_i)$ is the Dirac delta function. Note that $\mathbf{k}_i$ has n dimensions and $\delta(\mathbf{k}_i)$ is a 1D quantity on a multidimensional grid. Next, we write the lognormal moments with respect to their real-space counterparts,

$$\left\langle \prod_i^N [(2\pi)^n \delta_{\rm D}(\mathbf{k}_i) + \hat{\delta}^{\rm ln}(\mathbf{k}_i)] \right\rangle = \left\langle \prod_i^N \int {\rm d}\mathbf{x}_i e^{-{\rm i}\mathbf{k}_i.\mathbf{x}_i} [1 + \delta_i^{\rm ln}] \right\rangle , \tag{C3}$$

where $\delta_i^{\rm ln} \equiv \delta^{\rm ln}(\mathbf{x}_i)$. We can take the ensemble average inside the integral and rewrite the above equation as

$$\left\langle \prod_i^N \hat{\delta}^{\rm ln}(\mathbf{k}_i) \right\rangle = \prod_i^N \left[\int {\rm d}\mathbf{x}_i e^{-{\rm i}\mathbf{k}_i.\mathbf{x}_i} \right] \left\langle \prod_i^N [1 + \delta_i^{\rm ln}] \right\rangle . \tag{C4}$$

The ensemble average, $\langle \prod_i^N [1 + \delta_i^{\rm ln}] \rangle$, can be expressed in terms of the 2PCFs by writing the lognormal fields in terms of their Gaussian generators from equation (C1),

$$\left\langle \prod_i^N (1 + \delta_i^{\rm ln}) \right\rangle = \int_{-\infty}^{+\infty} {\rm d}\boldsymbol{\delta}\ p(\boldsymbol{\delta}) \prod_i^N \left[e^{\delta_i}\ e^{-\sigma_i^2/2} \right] , \tag{C5}$$

where $\boldsymbol{\delta} \equiv (\delta_1, \delta_2, ..., \delta_N)$ and

$$p(\boldsymbol{\delta}) = \frac{e^{-\boldsymbol{\delta}\,\mathbf{C}^{-1}\boldsymbol{\delta}^t/2}}{\sqrt{(2\pi)^N \det \mathbf{C}}} \tag{C6}$$

is the multivariate Gaussian-distributed probability of $\boldsymbol{\delta}$ with $\mathbf{C}$ as the covariance.

Since $\mathbf{C}$ is a covariance, i.e. symmetric and positive definite, we can write it in terms of its eigenvalues and eigenvectors as

$$\mathbf{C} = \mathbf{O}\mathbf{D}\mathbf{O}^{\rm t} , \tag{C7}$$

where $\mathbf{O}$ is the orthogonal matrix made out of the eigenvectors of $\mathbf{C}$ and $\mathbf{D}$ is a diagonal matrix of the eigenvalues of $\mathbf{C}$. As a result,

$$\mathbf{C}^{-1} = \mathbf{O}\mathbf{D}^{-1}\mathbf{O}^{\rm t} . \tag{C8}$$

Substituting for $\mathbf{C}^{-1}$ in equation (C5) and defining $X \equiv \boldsymbol{\delta}\mathbf{O}$ yields

$$\left\langle \prod_i^N (1 + \delta_i^{\rm ln}) \right\rangle = \frac{\prod_i^N e^{-\sigma_i^2/2}}{\sqrt{(2\pi)^N \det \mathbf{C}}} \int_{-\infty}^{+\infty} {\rm d}X\ e^{-X\mathbf{D}^{-1}X^t/2} \prod_i^N \left[e^{X_j O_{ji}^{\rm t}} \right] , \tag{C9}$$

where $|\det \mathbf{O}| = 1$ was used to simplify the result. Rewriting equation (C9) in terms of its components results in

$$\left\langle \prod_i^N (1 + \delta_i^{\rm ln}) \right\rangle = \frac{\prod_i^N e^{-\sigma_i^2/2}}{\sqrt{(2\pi)^N \det \mathbf{C}}} \prod_j^N \left[\int_{-\infty}^{+\infty} {\rm d}X_j\ e^{-X_j^2 D_{jj}^{-1}/2 + X_j \sum_i O_{ij}} \right] . \tag{C10}$$

With the aid of another variable change, $Y_i = X_i \sqrt{D_i^{-1}}$, and completing the square, we solve this integral and find the desired relationship,

$$\left\langle \prod_i^N (1 + \delta_i^{\rm ln}) \right\rangle = \exp\left(\frac{\sum_{ij} C_{ij} - \sigma_i^2}{2} \right) = \exp\left(\sum_{i<j} C_{ij} \right) . \tag{C11}$$

Inserting for $\langle \prod_i^N \hat{\delta}^{\rm ln}(k_i) \rangle$ from the above equation into equation (C4) results in

$$\langle \prod_i^N \hat{\delta}^{\rm ln}(k_i) \rangle = \prod_i^N \int {\rm d}x_i e^{-{\rm i}k_i.x_i} \prod_{i<j}^N e^{C_{ij}} . \tag{C12}$$

The covariance of the Gaussian and lognormal fields is related via

$$e^{C_{ij}} = 1 + \langle \delta_i^{\rm ln} \delta_j^{\rm ln} \rangle = 1 + \xi_{ij}^{\rm ln} , \tag{C13}$$

where $\xi_{ij}^{\rm ln}$ is the correlation between $\delta_i^{\rm ln}$ and $\delta_j^{\rm ln}$. Consequently, we can write the lognormal moments in equation (C12) in terms of their 2PCFs or alternatively their power spectra,

$$\langle \prod_i^N \hat{\delta}^{\rm ln}(k_i) \rangle = \frac{1}{(2\pi)^{Mn}} \prod_i^N \int {\rm d}x_i e^{-{\rm i}k_i.x_i} \prod_{j>i;m}^{N;M} \int {\rm d}l_m e^{{\rm i}l_m.(x_i - x_j)} [(2\pi)^n \delta_{\rm D} + P](l_m) , \tag{C14}$$

where $P(l)$ is the power spectrum of the lognormal field, n is the dimension of the field (for gravitational lensing $n = 2$), the subscript m belongs to each pair of $\delta_i^{\rm ln}$ and $\delta_j^{\rm ln}$ which make $\langle\delta_i^{\rm ln}\delta_j^{\rm ln}\rangle = \xi_{ij}^{\rm ln}$, and $M = N(N-1)/2$. The above integrals can be simplified by integrating with respect to x_i, since $[(2\pi)^n\delta_{\rm D} + P](l_m)$ have no dependence on x_i. The x_i integrals will result in N delta functions of dimension n that depend on k_i and l_m. There are M, l_m integrals and 2^M, $[(2\pi)^n\delta_{\rm D}(l_m) + P(l_m)]$ combinations. Writing the Delta functions found from the x_i integrals in the following form,

$$\delta_i^{\bar{\sum} j - \bar{\sum} k} \equiv \delta_i^{+j+j'+\ldots-k-k'-\ldots} \equiv \delta_{\rm D}(k_i + l_j + l_{j'} + \ldots - l_k - l_{k'} - \ldots)\,, \tag{C15}$$

will simplify the notation. Note that $\bar{\sum}$ is not a real sum. We find that the two sums over the positive and negative l_m modes can be formulated as follows,

$$\bar{\sum} j = \bar{\sum_{r=1}^{i-1}}(r-1)N + i - r(r+1)/2 - \bar{\sum} k = -\bar{\sum_{r=i}^{N-1}}(i-1)N - i(i-1)/2 + r - i + 1\,, \tag{C16}$$

for a given N and i. We are now left with M integrals with 2^M components for each,

$$\langle\prod_i^N \hat{\delta}^{\rm ln}(k_i)\rangle = \frac{1}{(2\pi)^{n(M-N)}}\prod_{j>i;m}^{N;M}\int {\rm d}l_m[(2\pi)^n\delta_{\rm D} + P](l_m)\prod_i^N \delta_i^{+\bar{\sum}_{r=1}^{i-1}(r-1)N+i-r(r+1)/2-\bar{\sum}_{r=i}^{N-1}(i-1)N-i(i-1)/2+r-i+1}\,. \tag{C17}$$

The remaining M integrals over l_m can be simplified using the N delta functions and the delta functions in the 2^M combinations of $\delta_{\rm D}(l_m)$ and $P(l_m)$. Some of these integrals vanish after considering the delta functions. In any integral if we come about a $\delta_{\rm D}(k_i)$, then that term is equal to zero since we are not interested in $k_i = 0$ terms and for the rest of the values, the delta function vanishes.

We can immediately see that for the third moment $N = M = 3$, i.e. only one integral will remain after the simplifications and the rest of the term will either vanish or are products of power spectra and Delta functions which depend on several k_i modes. We have developed an algorithm that can simplify the moments for any given N.

The fourth moment of the lognormal fields is essential for calculating the covariance of their power spectra. Therefore, here the results for the fourth-order moment will be explicitly shown. The fourth moment has many terms. These terms can be divided into four groups, depending on the number of remaining integrals over the power spectra and an extra group that contains the Gaussian-only contribution. Hence, in the following, each group will be represented separately. The fourth lognormal moment in Fourier space can be written as

$$\langle\hat{\delta}^{\rm ln}(k_1)\hat{\delta}^{\rm ln}(k_2)\hat{\delta}^{\rm ln}(k_3)\hat{\delta}^{\rm ln}(k_4)\rangle = (2\pi)^n\delta_{\rm D}(k_1+k_2+k_3+k_4)\{I + II + III + IV\} + G\,, \tag{C18}$$

where G is the pure Gaussian term,

$$\begin{aligned} G = (2\pi)^{2n}\big[&\delta_{\rm D}(k_2+k_3)\delta_{\rm D}(k_1+k_4)P(k_1)P(k_2)\\ &+\delta_{\rm D}(k_1+k_3)\delta_{\rm D}(k_2+k_4)P(k_1)P(k_2)\\ &+\delta_{\rm D}(k_1+k_2)\delta_{\rm D}(k_3+k_4)P(k_1)P(k_3)\big]\,, \end{aligned} \tag{C19}$$

and I, II, III, and IV are the pure lognormal terms shown below. The highest number of integrals remaining after the simplifications is 3. There is only a single term of this form,

$$I = \int {\rm d}l_4{\rm d}l_5{\rm d}l_6 P(l_4)P(l_5)P(l_6)P(l_4+l_5-k_2)P(l_6-l_4-k_3)P(k_4+l_5+l_6)\,. \tag{C20}$$

There are six terms with two integrals, which can be factorized as

$$\begin{aligned} II = &\int {\rm d}l_5{\rm d}l_6 P(l_5)P(l_6)P(k_4+l_5+l_6)\big[P(l_5+l_6-k_2-k_3)P(l_5-k_2) + P(l_5+l_6-k_2-k_3)P(l_6-k_3) + P(l_5-k_2)P(l_6-k_3)\big]\\ &+\int {\rm d}l_4{\rm d}l_6 P(l_4)P(l_6)P(l_6-l_4-k_3)P(k_4+l_6)\big[P(l_4-l_6+k_1+k_3) + P(l_4-k_2)\big]\\ &+\int {\rm d}l_4{\rm d}l_5 P(l_4)P(l_5)P(l_4+l_5-k_2)P(l_4+k_3)P(k_4+l_5)\,. \end{aligned} \tag{C21}$$

The 15 terms that have one remaining integral are

$$\begin{aligned} III = &\int {\rm d}l_6 P(l_6)\big[P(k_1)P(l_6-k_3)P(l_6-k_2-k_3) + P(k_1)P(k_4+l_6)P(l_6-k_1-k_3)\\ &+P(k_1)P(k_4+l_6)P(l_6-k_3) + P(k_2)P(l_6-k_3)P(l_6-k_1-k_3)\\ &+P(k_2)P(k_4+l_6)P(l_6-k_2-k_3) + P(k_2)P(k_4+l_6)P(l_6-k_3)\\ &+P(k_4+l_6)P(l_6-k_2-k_3)P(l_6-k_3) + P(k_4+l_6)P(l_6-k_1-k_3)P(l_6-k_3)\big]\\ &+\int {\rm d}l_5 P(l_5)\big[P(k_4+l_5)P(l_5-k_2)P(l_5-k_2-k_3) + P(k_3)P(l_5-k_2)P(l_5-k_1-k_2)\\ &+P(k_3)P(k_4+l_5)P(l_5-k_2-k_3) + P(k_3)P(l_5-k_2)P(k_4+l_5)\big]\\ &+\int {\rm d}l_4 P(l_4)P(k_4)\big[P(l_4-k_2)P(l_4-k_1-k_2) + P(l_4+k_3)P(l_4+k_1+k_3) + P(l_4-k_2)P(l_4+k_3)\big]\,. \end{aligned} \tag{C22}$$

And finally, there are 16 terms that do not have any remaining integrals and depend only on the power spectra of the lognormal modes,

$$\begin{aligned} IV = {} & P(k_1)P(k_2)P(k_3) + P(k_1)P(k_2)P(k_4) + P(k_1)P(k_3)P(k_4) + P(k_2)P(k_3)P(k_4) \\ & + [P(k_1)P(k_2) + P(k_3)P(k_4)][P(k_1 + k_3) + P(k_2 + k_3)] \\ & + [P(k_1)P(k_3) + P(k_2)P(k_4)][P(k_1 + k_2) + P(k_2 + k_3)] \\ & + [P(k_1)P(k_4) + P(k_2)P(k_3)][P(k_1 + k_2) + P(k_1 + k_3)] \,. \end{aligned} \tag{C23}$$

IV has the highest contribution out of all of the pure lognormal terms as was shown by Hilbert et al. (2011) for the covariance of the 2PCFs. Ergo, to find the covariance of the power spectra for a lognormal field, we neglect *I*, *II*, and *III*.

This paper has been typeset from a TEX/LATEX file prepared by the author.