Astronomy & Astrophysics manuscript no. paperdraft

May 6, 2019

The Effects of Varying Depth in Cosmic Shear Surveys

S. Heydenreich[1], Peter Schneider[1], Hendrik Hildebrandt[2,1], Catherine Heymans[3], and Marika Asgari[3]

order of authors?, KiDS-Authors?

[1] Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany

[2] Astronomisches Institut, Ruhr-Universität Bochum, Universitätsstr. 150, 44801 Bochum, Germany

[3] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

e-mail: `sven@astro.uni-bonn.de`

Received XXX; accepted XXX

ABSTRACT

Cosmic shear proves to be a powerful tool to study the properties of the local Universe. The discrepancy in the parameter S_8 between measurements in the local Universe and the Cosmic Microwave Background motivates further investigation of yet unaccounted systematic biases. Especially ground-based surveys are subject to a variation in depth. We want to understand and quantify the resulting effects. In particular, we check if they introduce a bias to the cosmological parameters and if they can be responsible for the occurence of *B*-Modes. We construct a semi-analytic model to estimate the impact on the shear correlation function and analyze the implications for cosmological parameters. Furthermore we construct COSEBIs of the correlation functions to quantify the occuring B-Modes. For the Kilo-Degree Survey this effect introduces an error in $\xi_\pm$ of the order of a few percent on small scales, which is responsible for a 0.1σ bias in Ω_m and σ_8. However, the parameter S_8 is robust against this modification. We also report the occurrence of B-Modes, although not to a significant degree. We conclude that the effects of varying optical depth for ground-based surveys on cosmological parameters are not yet significant, but should be accounted for in next-generation experiments.

Key words. gravitational lensing – weak lensing – cosmic shear

1. Introduction

The discovery of cosmic shear has provided us with a new and powerful cosmological tool to investigate the ΛCDM Model and determine its parameters. Contrary to the analysis of the CMB by Planck Collaboration et al. (2018), cosmic shear is more sensitive to the properties of the local Universe and thus provides an excellent consistency check for the standard model of cosmology. Current cosmic shear surveys are especially sensitive to the parameter $S_8 = \sigma_8 \sqrt{\Omega_{\mathrm{m}}/0.3}$, where σ_8 denotes the normalisation of the matter power spectrum and Ω_{m} is the matter density. It is interesting to note that all three current major cosmic shear results report a lower S_8 than inferred from CMB analysis: While Planck Collaboration et al. (2018) determined a value of $S_8 = 0.830 \pm 0.013$, Hikage et al. (2019) report $S_8 = 0.800^{+0.029}_{-0.028}$ from analysis of the Subaru Hyper Suprime-Cam survey, Hildebrandt et al. (2018) report $0.737^{+0.040}_{-0.036}$ from KiDS+VIKING data and the Dark Energy Survey (Troxel et al. 2018) reports $S_8 = 0.782 \pm 0.027$. Also, Heymans et al. (2013) report $S_8 = 0.759 \pm 0.020$ from analysis of CFHTLens data. This discrepancy has received a lot of attention (Verde et al. 2013). It could be interpreted as a statistical coincidence, a sign of new physics like massive neutrinos (Battye & Moss 2014), time-varying dark energy or modified gravity (Planck Collaboration et al. 2016); or as the manifestation of a systematic effect, either in the cosmic shear surveys or in the Planck mission (Addison et al. 2016), that is not yet accounted for.

As weak gravitational lensing measures a tiny signal over a large sample, it is extremely sensitive to anything that systematically biases the measurements, such that the error bars in current surveys arise to equal parts from statistical and systematic uncertainties (compare Hildebrandt et al. 2017). With next-generation surveys like the Large Synoptic Survey Telescope and Euclid right at the doorstep, systematic effects in gravitational lensing have received an unprecedented amount of attention (Asgari et al. 2018; Blake 2019; Shirasaki et al. 2019).

To check for remaining systematics, a weak lensing signal can be divided into two components, the so-called E- and B-modes (Crittenden et al. 2002; Schneider et al. 2002). To leading order, B-modes can not be created by astrophysical phenomena and are thus an excellent test for remaining systematics. As Hildebrandt et al. (2017) reported the significant detection of B-modes, it was well motivated to check for possible systematics that are not yet accounted for. Note that the non-existence of B-modes does not necessarily imply that the sample is free of remaining systematics.

One systematic effect is the variation of depth in a survey. While effects like Galactic extinction or dithering strategies do play a role in every survey, this work focuses on the effects caused by varying atmospheric conditions, that are found in ground-based surveys. To first order, this variation can be modelled by a step-like depth function which varies from pointing to pointing. In this work we assume the specifications of the Kilo-Degree Survey, namely a collection of 1 deg^2 square fields.

In Section 2 we will introduce a simple toy model to understand this effect and analyze the impact on the power spectrum. In Section 3 we will estimate the effect on the shear correlation functions $\xi_{\pm}$ using two different models. We will present our results in Section 4. In Section 5 we will discuss our results and comment on the impacts of our used simplifications. We will assume the standard weak gravitational lensing formalism, a summary of which can be found in Bartelmann & Schneider (2001).

2. Modelling the Power Spectrum

For our first analysis we will further simplify our assumptions: We imagine that all the matter between sources and observer is concentrated in a single lens plane of distance D_d from the observer. If we now distribute sources at varying distances D_s, then the convergence κ varies according to $\kappa \propto D_{\mathrm{d}s}/D_\mathrm{s}$.

2.1. Effects on the Power Spectrum

Assuming that the depth, and thus the source redshift populations, vary between pointings, an observer will measure a shear-signal that is modified by a locally constant depth-function $\gamma^\mathrm{obs}(\boldsymbol{\theta}) = W(\boldsymbol{\theta})\gamma(\boldsymbol{\theta})$ with $W(\boldsymbol{\theta}) = 1 + w(\boldsymbol{\theta})$, where $\langle w(\boldsymbol{\theta})\rangle = 0$ holds. In accordance to the definition of the shear power spectrum

$$(2\pi)^2\delta(\boldsymbol{\ell} - \boldsymbol{\ell}')P(|\boldsymbol{\ell}|) = \langle\hat{\gamma}(\boldsymbol{\ell})\hat{\gamma}(\boldsymbol{\ell}')\rangle \ , \tag{1}$$

we define the observed power spectrum via

$$P^\mathrm{obs}(\boldsymbol{\ell}) \equiv \frac{1}{(2\pi)^2}\int \mathrm{d}^2\ell' \left\langle \hat{\gamma}^\mathrm{obs}(\boldsymbol{\ell})\hat{\gamma}^\mathrm{obs*}(\boldsymbol{\ell}')\right\rangle \ . \tag{2}$$

Note that due to the depth-function both the assumptions of homogeneity and isotropy break down, which means that we can neither assume isotropy in the power spectrum, nor can we assume that $\left\langle \hat{\gamma}^\mathrm{obs}(\boldsymbol{\ell})\hat{\gamma}^\mathrm{obs*}(\boldsymbol{\ell}')\right\rangle$ vanishes for $\boldsymbol{\ell} \neq \boldsymbol{\ell}'$. To model a constant depth on each individual pointing $\boldsymbol{\alpha}$, we can choose random variables $w_{\boldsymbol{\alpha}}$, that only need to satisfy $\langle w_{\boldsymbol{\alpha}}\rangle = 0$, and parametrize $w(\boldsymbol{\theta})$ as

$$w(\boldsymbol{\theta}) = \sum_{\boldsymbol{\alpha}\in\mathbb{Z}^2} w_{\boldsymbol{\alpha}}\Xi(\boldsymbol{\theta} - L\boldsymbol{\alpha}) \ \text{, with the Box-Function } \Xi(\boldsymbol{\theta}) = \begin{cases} 1 & \boldsymbol{\theta} \in \left[-\frac{L}{2}, \frac{L}{2}\right]^2 \\ 0 & \text{else} \end{cases} \ , \tag{3}$$

where L is the length of one pointing. Following the calculations in Appendix A.1, we derive

$$P^\mathrm{obs}(\boldsymbol{\ell}) = P(\boldsymbol{\ell}) + \left\langle w^2\right\rangle \int \frac{\mathrm{d}^2\boldsymbol{k}}{(2\pi)^2}\,\hat{\Xi}(\boldsymbol{\ell} - \boldsymbol{k})P(\boldsymbol{k}) \,, \tag{4}$$

where $\left\langle w^2\right\rangle \equiv \left\langle w_{\boldsymbol{\alpha}}^2\right\rangle$ is the dispersion of the depth-function. The observed power spectrum P^obs is thus composed of the original power spectrum P, plus a convolution of the power spectrum with the fourier transfrom of a box-function, scaling with the variance of the geometric lensing efficiency $\left\langle\frac{D_{\mathrm{ds}}}{D_s}\right\rangle$. In particular, the power spectrum is not isotropic anymore. Following Schneider et al. (2002), it would be interesting to extract E- and B-mode information out of this power spectrum, however Schneider et al. (2010) present a more sophisticated decomposition of E- and B-modes, that can also be applied to real data using the shear correlation function, so instead we want to focus our efforts on these parts.

3. Modelling the shear correlation functions

A convenient way to infer cosmological information from observational data are the shear correlation functions $\xi_\pm$, which are defined as

$$\xi_\pm = \langle\gamma_t\gamma_t\rangle \pm \langle\gamma_\times\gamma_\times\rangle \; . \tag{5}$$

They are the prime estimators to quantify a cosmic-shear signal since it is simple to include a weighting of the shear-measurements into the correlation functions and, contrary to the power spectrum, one does not have to worry about the shape of the survey footprint, or masked regions. Cosmologically, given two comoving distance probability distributions of sources $p_i(\chi), p_j(\chi)$, one can compute the shear correlation function from the underlying matter power spectrum P_δ via

Citation!

$$\xi_\pm[\theta, p_i, p_j] = \int_0^\infty \frac{\mathrm{d}l\, l}{2\pi} J_{0,4}(l\theta) P(l, p_i, p_j) \, , \tag{6}$$

$$P(l, p_i, p_j) = \frac{9H_0^4\Omega_\mathrm{m}^2}{4c^4} \int_0^{\chi_h} \mathrm{d}\chi \frac{g(\chi, p_i) g(\chi, p_j)}{a^2(\chi)} P_\delta\left(\frac{l}{f_K(\chi)}, \chi\right) , \tag{7}$$

$$g(\chi, p_i) = \int_\chi^{\chi_H} \mathrm{d}\chi' \, p_i(\chi') \frac{f_K(\chi' - \chi)}{f_K(\chi')} \, . \tag{8}$$

Here, $J_{0,4}$ denote the 0-th and 4-th order Bessel Functions.

3.1. Using an analytic Model

For a first simple analysis we will assume that a deeper redshift distribution just yields a stronger shear signal, in the sense that the shear field for a deeper redshift distribution gets multiplied by a weight W. While this is not true for a 3-dimensional matter distribution, it should be valid for small variations in redshift. Additionally, we assume that a higher depth does not only lead to a stronger average shear, but also to a higher galaxy number density, implying a correlation between those two quantities.

Let $N^i(\boldsymbol{\theta}), N^j(\boldsymbol{\theta})$ be the average weighted number of galaxies per pointing in redshift bins i and j and let $W^i(\boldsymbol{\theta}), W^j(\boldsymbol{\theta})$ be the weighting of average shear. The observed correlation function $\xi_\pm^{ij,\mathrm{obs}}(\theta)$ now changes from one of constant depth $\xi_\pm^{ij,\mathrm{const}}(\theta)$ via (compare Eq.)

Take eq. from Hildebrandt et al. (2017)?

$$\begin{aligned}
\xi_\pm^{ij,\mathrm{obs}}(\theta) &= \frac{\left\langle N^i(\boldsymbol{\theta}')N^j(\boldsymbol{\theta}'+\boldsymbol{\theta})\gamma_\mathrm{t}^{i,\mathrm{obs}}(\boldsymbol{\theta}')\gamma_\mathrm{t}^{j,\mathrm{obs}}(\boldsymbol{\theta}'+\boldsymbol{\theta})\right\rangle}{\langle N^i(\boldsymbol{\theta}')N^j(\boldsymbol{\theta}'+\boldsymbol{\theta})\rangle} \pm \frac{\left\langle N^i(\boldsymbol{\theta}')N^j(\boldsymbol{\theta}'+\boldsymbol{\theta})\gamma_\times^{i,\mathrm{obs}}(\boldsymbol{\theta}')\gamma_\times^{j,\mathrm{obs}}(\boldsymbol{\theta}'+\boldsymbol{\theta})\right\rangle}{\langle N^i(\boldsymbol{\theta}')N^j(\boldsymbol{\theta}'+\boldsymbol{\theta})\rangle} \\
&= \frac{\left\langle N^i(\boldsymbol{\theta}')N^j(\boldsymbol{\theta}'+\boldsymbol{\theta})W^i(\boldsymbol{\theta}')W^j(\boldsymbol{\theta}'+\boldsymbol{\theta})\right\rangle}{\langle N^i(\boldsymbol{\theta}')N^j(\boldsymbol{\theta}'+\boldsymbol{\theta})\rangle} \xi_\pm^{ij,\mathrm{const}}(\theta) \, ,
\end{aligned} \tag{9}$$

where the average $\langle \ldots \rangle$ represents both an ensemble average as well as an average over the position $\boldsymbol{\theta}'$. Assuming that depth and galaxy number density of neighbouring pointings are uncorrelated, the only important property of a galaxy pair is whether or not they lie in the same pointing, which is described by the function $E(\theta)$.

To compute the modified shear correlation functions, we parametrize the number densities $N^i(\boldsymbol{\theta}) = \left\langle N^i \right\rangle [1 + n^i(\boldsymbol{\theta})]$ and the weight $W^i(\boldsymbol{\theta}) = 1 + w^i(\boldsymbol{\theta})$ and, as in (3), interpret $n^i(\boldsymbol{\theta})$ as a function with average $\left\langle n^i \right\rangle = 0$, that is constant on each pointing. Without loss of generality we set $\left\langle N^i \right\rangle = 1$. We can see that $\left\langle n^i(\boldsymbol{\theta}')n^j(\boldsymbol{\theta}' + \boldsymbol{\theta}) \right\rangle = E(\theta) \left\langle n^i(\boldsymbol{\theta}')n^j(\boldsymbol{\theta}') \right\rangle \equiv E(\theta) \left\langle n^i n^j \right\rangle$ holds and compute:

$$\begin{aligned}
&\left\langle N^i(\boldsymbol{\theta}')N^j(\boldsymbol{\theta}' + \boldsymbol{\theta})W^i(\boldsymbol{\theta}')W^j(\boldsymbol{\theta}' + \boldsymbol{\theta}) \right\rangle \\
&\quad = 1 + \left\langle n^i w^i \right\rangle + \left\langle n^j w^j \right\rangle + E(\theta) \Big[\left\langle n^i n^j \right\rangle + \left\langle n^i w^j \right\rangle + \left\langle n^j w^i \right\rangle + \left\langle w^i w^j \right\rangle + \left\langle n^i n^j w^i \right\rangle + \left\langle n^i n^j w^j \right\rangle + \left\langle n^i w^i w^j \right\rangle \\
&\qquad + \left\langle n^j w^i w^j \right\rangle + \left\langle n^i n^j w^i w^j \right\rangle \Big]
\end{aligned} \tag{10}$$

Ignoring correlations higher than second order[1], and performing the same calculation for the denominator of Eq. (9), we get

$$\xi_\pm^{ij,\mathrm{obs}}(\theta) = \frac{1 + \left\langle n^i w^i \right\rangle + \left\langle n^j w^j \right\rangle + E(\theta) \left[\left\langle n^i n^j \right\rangle + \left\langle n^i w^j \right\rangle + \left\langle n^j w^i \right\rangle + \left\langle w^i w^j \right\rangle \right]}{1 + E(\theta) \left\langle n^i n^j \right\rangle} \xi_\pm^{ij,\mathrm{const}}(\theta) \,. \tag{11}$$

For the calculation of the reference correlation functions $\xi_\pm^{ij}(\theta)$ we distribute the *same* galaxies into our survey, only this time we will not order their weightings W or their number densities N by pointing. We can imagine this by cutting the footprint into infinitesimal elements $\mathrm{d}^2\boldsymbol{\theta}$, and redistributing those at random. When we calculate the correlation function of this survey, we note that $\left\langle n^i(\boldsymbol{\theta}') \right\rangle \left\langle n^j(\boldsymbol{\theta}' + \boldsymbol{\theta}) \right\rangle = 0$ holds for $\boldsymbol{\theta} \neq \mathbf{0}$, as the two corresponding infinitesimal elements have uncorrelated weighting and number density. Performing the same calculations as above, this yields a relation between the correlation function of constant optical depth $\xi_\pm^{ij,\mathrm{const}}$ and the modelled one $\xi_\pm^{ij}$:

$$\xi_\pm^{ij}(\theta) = \left(1 + \left\langle n^i w^i \right\rangle + \left\langle n^j w^j \right\rangle\right) \xi_\pm^{ij,\mathrm{const}}(\theta) \,. \tag{12}$$

The ratio of modelled and observed correlation function thus becomes:

$$\frac{\xi_\pm^{ij}(\theta)}{\xi_\pm^{ij,\mathrm{obs}}(\theta)} \approx \frac{1 + \left\langle n^i w^i \right\rangle + \left\langle n^j w^j \right\rangle + E(\theta) \left\langle n^i n^j \right\rangle}{1 + \langle n^i w^i \rangle + \langle n^j w^j \rangle + E(\theta) \left[\langle n^i n^j \rangle + \langle n^i w^j \rangle + \langle n^j w^i \rangle + \langle w^i w^j \rangle \right]} \,. \tag{13}$$

It is interesting to note that $\xi_\pm^{ij} = \xi_\pm^{ij,\mathrm{obs}}$ holds wherever $E(\theta) = 0$, meaning that the correlation function is not affected for large angular scales. Given a set of average redshifts, following Van Waerbeke et al. (2006), we can estimate

$$\langle |\gamma| \rangle \propto \langle z \rangle^{0.85} \,. \tag{14}$$

[1] It is not inherently obvious that this is a valid assumption. However, after performing both calculations we noticed no difference between the outcomes of both equations

We shall later see that this approximation is valid for higher tomographic redshift bins $z \gtrsim 0.5$, but starts to break down at lower redshifts. We thus want to construct a model that is valid for an arbitrary distribution of redshifts and does not rely on the assumption of a single lens plane.

3.2. Using a semi-analytic Model

The analysis of data from the Kilo-Degree Survey showed that the redshift-distribution of sources was highly correlated with the depth in the r-band. We thus chose to separate the survey into 10 percentiles, sorted by r-band depth, meaning that if a pointing had a worse depth than 90% of the other pointings, it would belong to the first percentile, and so on. Now for each percentile m and each tomographic redshift bin i we can extract a weighted number of galaxies N^i_m and, in case the pointing overlaps with a spectroscopic survey, a source redshift distribution $p^i_m(z)$. Using (6), we can compute the shear correlation functions $\xi^{ij}_{\pm,mn}(\theta)$ for each set of percentiles m, n and redshift bins i, j[2]. When we compute the measured shear correlation functions of a survey, we take the weighted average of tangential and cross shears of all pairs of galaxies (Hildebrandt et al. (2017) give a good overview for the process). If, for a single pair of galaxies, one galaxy lies in the m-th percentile of redshift bin i and the second one lies in the n-th percentile of redshift bin j, then their contribution to the observed correlation functions is, on average, $\xi^{ij}_{\pm,mn}(\theta)$. This means that if we know each of those single correlation functions, we can reconstruct the total correlation functions via a weighted average of the single functions. Formally, we define

Cite the same eq. as above from Hildebrandt et al. (2017)

$$\xi^{ij,\mathrm{obs}}_{\pm}(\theta) = \frac{\sum_{m,n} P^{ij}_{mn}(\theta)\xi^{ij}_{\pm,mn}(\theta)}{\sum_{m,n} P^{ij}_{mn}(\theta)}\,, \tag{15}$$

where P^{ij}_{mn} is the new weighting of the correlation functions. This weighting has to be proportional to the probability that a galaxy pair of distance θ is of percentiles m and n, as well as to the original weighting of these galaxies.

We could completely scratch this derivation and just point to the Appendix; I personally feel that this approach is more intuitive than the more mathematical one of the appendix, but if the paper gets too long we definitely do not need two methods to derive the same equation.

In this analysis, we will assume an infinitely large survey footprint with an uncorrelated distribution of depth. We will later discuss the validity of these assumptions as well as possible mitigation strategies. For $m \neq n$ we know that the pair of galaxies has to lie in different pointings, which is accounted for by including the factor $[1 - E(\theta)]$. Furthermore, the first galaxy has to lie in percentile m, the probability of which is $1/10$. When the first pointing is of percentile m, the probability that a galaxy in a different pointing is of percentile n is also equal to $1/10$. The impact of such a galaxy pair on the correlation functions scales with the product of the weighted number of galaxies N^i_m, N^j_n. We get for $n \neq m$:

$$P^{ij}_{mn}(\theta) = [1 - E(\theta)]\frac{1}{100}N^i_m N^j_n\,. \tag{16}$$

[2] For the calculation of the shear correlation functions we use the Nicaea-program. Among other things, it calculates the shear correlation functions for a given cosmology and source redshift distribution. To estimate the power spectrum on nonlinear scales we use the methods developed by Takahashi et al. (2012).

For the calculation of $P^{ij}_{mm}(\theta)$ we have to account for a different possibility: In case that the galaxy lies in the same pointing [accounted for by the factor $E(\theta)$], it automatically is of the same percentile. We therefore set

$$P^{ij}_{mm}(\theta) = E(\theta)\frac{1}{10}N^i_m N^j_m + [1 - E(\theta)]\frac{1}{100}N^i_m N^j_m\,. \tag{17}$$

We can then write $P^{ij}_{mn}(\theta)$ as:

$$P^{ij}_{mn}(\theta) = E(\theta)\frac{1}{10}N^i_m N^j_n\,\delta_{mn} + [1 - E(\theta)]\frac{1}{100}N^i_m N^j_m\,, \tag{18}$$

where δ_{mn} denotes the Kronecker delta. Inserting this into Eq. (15), we compute

$$\xi^{ij,\mathrm{obs}}_{\pm,mn}(\theta) = \frac{1}{C}\sum_{m=1}^{10} N^i_m \left\{E(\theta)N^j_m\xi^{ij}_{\pm,mm}(\theta) + \frac{[1 - E(\theta)]}{10}\sum_{n=1}^{10} N^j_n\xi^{ij}_{\pm,mn}(\theta)\right\}\,, \tag{19}$$

with the normalization

$$C = \sum_{m=1}^{10} N^i_m\left[E(\theta)N^j_m + \frac{[1 - E(\theta)]}{10}\sum_{n=1}^{10} N^j_n\right]. \tag{20}$$

A more mathematically rigorous derivation of this function can be found in Appendix A.3.

If we want to compute this for all 5 redshift bins of the KV450-survey, this forces us to calculate 1275 correlation functions and add them, thus yielding potential numerical errors (apart from being computationally expensive). However, if we examine Eq. (8), we see that the comoving distance distribution of sources factors in linearly. This in turn implies that in Equations (7) and (6) both source distance distributions factor in linearly. This basically means that, instead of adding correlation functions, we can add their respective redshift distributions and compute the correlation functions of that. In particular, we can define the *combined number of galaxies* N^i and *average redshift distribution* $p^i(z)$ of tomographic bin i as

$$N^i \equiv \sum_m N^i_m\,, \qquad p^i(z) = \frac{\sum_m N^i_m p^i_m(z)}{\sum_m N^i_m}\,. \tag{21}$$

If we define $\xi^{ij}_\pm$ as the correlation functions between the average redshift distributions $p^i(z)$ and $p^j(z)$, then we observe:

$$\sum_{m,n} N^i_m N^j_n \xi^{ij}_{\pm,mn} = N^i N^j \xi^{ij}_\pm\,. \tag{22}$$

Consequently, we can apply this to (19), yielding

$$\xi^{ij,\mathrm{obs}}_\pm(\theta) = \frac{1}{C}\left\{E(\theta)\left[\sum_{m=1}^{10} N^i_m N^j_m \xi^{ij}_{\pm,mm}(\theta)\right] + \frac{[1 - E(\theta)]}{10}\xi^{ij}_\pm(\theta)N^i N^j\right\}\,. \tag{23}$$

For each set of redshift bins we thus only have to compute eleven correlation functions, which reduces the number of functions to compute from 1275 to 165. We can see that for large distances

θ, such that $E(\theta) = 0$ holds, the observed correlation functions agree with the ones that we would usually calculate.

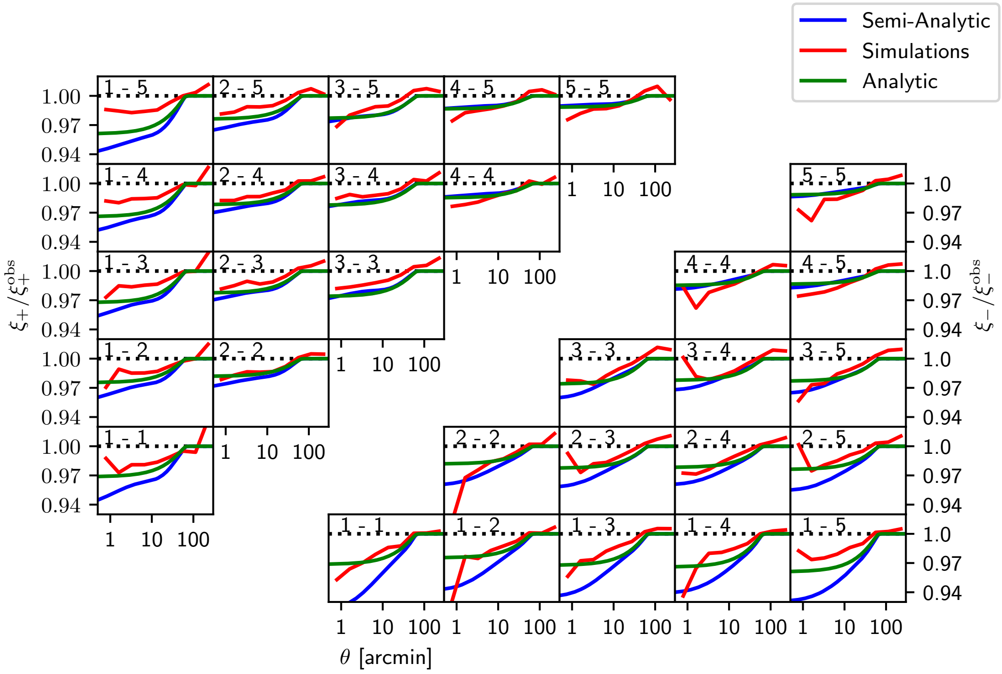


Fig. 1: The ratio of observed to modeled correlation functions for the analytic method (green), the semi-analytic method (blue) and the numerical simulations (red) for a cross-correlation of all redshift bins.

4. Results

We applied both our methods to data from the KV450 survey and computed the ratio of observed and modeled correlation functions. Furthermore, we conducted numerical simulations investigating the same issue: A 100 deg^2 field in the Scinet Light Cone Simulations (SLICS) (Harnois-Déraps et al. 2018) was randomly separated into 10 percentiles. For each tomographic redshift bin of each percentile, galaxies were placed to trace the respective redshift distribution. Afterwards, their expected shear was determined (shape-noise in the form of intrinsic ellipticities of galaxies was not included). This was compared to a set of simulations where the galaxies were simply distributed according to the combined redshift distribution of each respective tomographic bin. As in Hildebrandt et al. (2018), we have separated the data in 5 tomographic redshift bins and performed our analysis for a cross-correlation of all bins. The result can be seen in Figure 1. We can see that for high redshift bins, the analytic and the semi-analytic methods are consistent, whereas for low redshift bins they significantly diverge. We explain this due to the facts that the analyic method uses simplifications that are redshift-dependent and only hold for small variations in redshift, which is not fulfilled in the low redshift bins.[3]

Catherine, is this what you did? If you are unhappy with any of these explanations, or feel like I missed something, please feel free to add it :)

Should I include the plot number density vs average redshift from the talk? Maybe in the appendix to avoid too many Figures?

[3] We also observe that in the semi-analytic model, ξ_- seems to be much stronger affected by this effect: Following Equation (6), ξ_+ is computed by filtering the power spectrum with the 0-th order Bessel function. This function peaks at $\ell\theta = 0$, meaning that for all values of θ, the correlation function ξ_+ is sensitive to small values of ℓ, corresponding to large separations θ. However, ξ_- is obtained by filtering with the 4-th order Bessel function, which peaks at approximately $\ell\theta \approx 5$, so for different θ this function is sensitive to varying parts of the convergence power spectrum. A more detailed analysis of this can be found in the Appendix of Köhlinger et al. (2017).

The simulations seem to be in rough agreement with the models, but there are some features that can not be explained. Especially we observe that the ratios of the correlation functions for large values of θ seem to consistently surpass unity, which can not be explained by our models. The simulations seem to be in rough agreement with the models, but there are some significant differences. After a thorough analysis we explain these discrepancies the following way: The simulations were performed on a $100\,\mathrm{deg}^2$ field, which means that shot-noise of the fields plays a significant role. After performing the same simulations for a different distribution of depth between the pointings and obtaining completely different results, we are quite certain that this is the dominating effect. The implications of this and possible mitigation strategies will be discussed in Section 5.

As the next step we computed a reference correlation function given a fiducial cosmology for each combination of redshift bins, and modified said correlation function according to our semi-analytic model. Then we ran a Markov-Chain Monte Carlo simulation[4] to check for a potential bias in the cosmological parameters, using the covariance-matrix computed in Hildebrandt et al. (2017). As our main interest lies in the Ω_m - σ_8 combination, we restricted our analysis to those two parameters. As can be seen in Figure **??**, the impact of varying depth is noticeable, but insignificant compared to the uncertainties. However, to get a rough estimate for the impact on a Euclid-like survey, we divided the used covariance-matrix by 30, to account for the increased survey area. As can be seen in Figure **??**, here the impact on both Ω_m and σ_8 is significant, however it seems that the parameter S_8 is extremely robust against this effect.

To estimate the B-Modes created by this effect, we have extracted the *Complete Orthogonal Sets of E- and B-Mode Integrals* (COSEBIS, compare Schneider et al. (2010)), once of a reference set of correlation functions, to estimate numerical inaccuracies,[5] and then for the correlation functions that have been modified to account for a varying depth. We report a consistent B-Mode pattern across all redshift-bins, which can be seen in Figure 2. Although we did not determine the error bars, that a KiDS-like survey would imply on those functions, Asgari & Heymans (2018) calculated the COSEBIs of the KV450 survey for the same range in θ. The measured B-Modes were about one order of magnitude higher than the ones created by the varying depth, and still found to be not significant. Given these facts, we conclude that the creation of B-Modes due to varying depth in the KV450 survey is not significant. However, the pattern seems to be very characteristic, so when one encounters B-Modes in next generation surveys, which show a similar pattern, this would suggest that they are created by a similar effect (although we can not exclude other effects that just create the same B-Mode pattern). We also note that the difference in E-Modes is as large as the B-Modes, which suggests that any significant change in the cosmological parameters due to a varying depth will also yield the significant detection of B-Modes.

[4] The code for this was developed by Jan-Luca van den Busch and used in a modified version.
[5] For a reference correlation function the B-Modes should be consistent with zero.

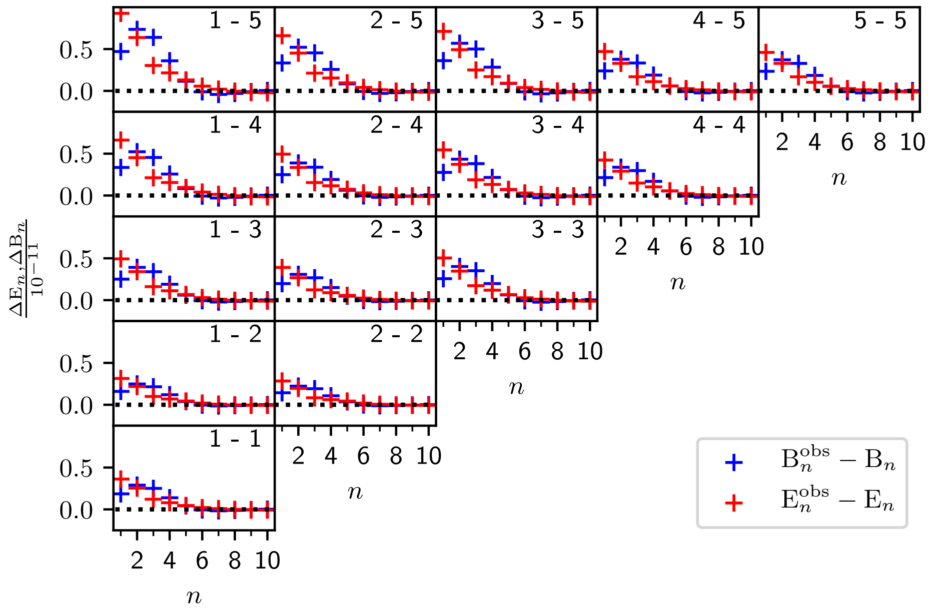


Fig. 2: Difference in the logarithmic E- and B-modes between the reference and the observed correlation functions for an angular range of $\theta_{min} = 0.'5$, $\theta_{max} = 100'$.

5. Discussion

With our semi-analytic model we try to describe the impacts of varying depth in ground-based surveys. During our analysis we have assumed a few simplifications, which we will discuss.

1. In the most general terms, we are analysing the effects of a position-based selection function on cosmic shear surveys. In our analysis, this selection function was governed by the r-band depth of a pointing. This neglects a number of other effects: The depth in different bands and the seeing of a pointing will also modify the number densities and redshift distributions on the scale of a pointing, whereas dithering strategies as well as imperfections in the telescope and CCD cause modifications on sub-pointing scales. However, we believe that these effects are subdominant compared to the variations caused by the r-band depth.
2. We have assumed an infinitely large survey area with an uncorrelated distribution of the depth-function. While the boundary effects arising from a finite survey footprint would have a small impact on the shape of the function $E(\theta)$[6], the governing factor is the shot noise of the depth-distribution. We have assumed that the probability that a neighbouring pointing is of percentile n is exactly the expectation value, namely $1/10$. While this is true for an infinitely large survey with an uncorrelated distribution of the depth-function, Figure 1 clearly shows that it is not valid for a $100\,\mathrm{deg}^2$ field. Whether our assumption holds true for the $450\,\mathrm{deg}^2$ footprint of the KV450 survey, or even the $1350\,\mathrm{deg}^2$ footprint of the final survey is an important question. Also, we have assumed an uncorrelated distribution of the depth-function, both in the simulations as well as in the models. While this is likely true as a rough approximation (very few pointings of the survey were taken on the same night and thus under the same weather conditions), effects like airmass, lunar phase, Galactic extinction and seasonal weather are likely to influence the depth on scales larger than one pointing.

   An outlook for the mitigation of finite field effects, boundaries and a correlated distribution of depth is given in Appendix B. As a preliminary result we find that finite field effects are not significant for a $450\,\mathrm{deg}^2$ or $1000\,\mathrm{deg}^2$-field, if the distribution of depth is uncorrelated.
3. In our MCMC simulations we did not account for degeneracies with other cosmological parameters or observational effects. Especially intrinsic alignments and baryonic feedback also modify the correlation functions especially on small scales, so they are probably degenerate with the effect of varying depth. In an actual MCMC simulation that accounts for these effects, we suspect that the parameters for intrinsic alignments and baryonic feedback change to mitigate this effect, and the impact on cosmological parameters is actually smaller than in our results. Also, possible degeneracies between S_8 and other cosmological parameters might bias the resulting values.

Despite these repercussions, we are confident to say that the effects of varying depth are not significant for the KV450 survey. The cosmological parameters did not change significantly and the

[6] This would be due to the fact that a pointing next to a boundary has less neighbours, therefore making it more likely that a galaxy pair is in the same pointing.

main parameter, S_8, is especially robust against this effect. In particular this means that a varying depth can not explain the discrepancy between observations of the local Universe and results from analysis of the CMB.

We have shown that this effect can create B-modes. However, Asgari et al. (2018) measured the B-modes of the KV450 survey in the same θ-range. Those B-modes are at least one order of magnitude larger and still consistent with zero, so it is safe to say that the modes created by varying depth are negligible. An interesting observation is that the change in E-modes is as big as the created B-modes (compare Figure **??**). This means that as soon as this effect causes significant biases in the cosmological parameters, it will also create significant B-modes[7]. Additionally, the created pattern is very characteristic, which makes it easy to recognize in a B-mode analysis of an actual survey.

For next-generation surveys like Euclid, this effect will be significant. Although Euclid is a space-based telescope, the photometric redshift determination will still be done by ground-based telescopes and therefore suffer from the same effects. While we did not yet perform a quantitative study of this effect for Euclid, we are certain that a such a study should be conducted.

As an outlook it would be interesting to perform a full MCMC simulation, including both cosmological and nuisance parameters. One might also check the most popular extensions to the standard model of cosmology, and see whether varying depth could cause a bias towards one of these models. Furthermore, the method developed in Appendix B could be applied to the actual footprint of the KV450 survey to see whether finite field effects or a correlated distribution of r-band depth are significant. Additionally is interesting to note that $E(\theta)$ is the azimuthal average of the function $E(\boldsymbol{\theta})$, which is not isotropic. Therefore, it would be possible to observe a direction-dependent correlation function $\xi_\pm^{ij,\mathrm{obs}}(\boldsymbol{\theta})$ in future surveys. An anisotropy in the observed correlation function could be a sign for the influence of varying depth.

[7] While this is no big surprise, it is not trivial. It could be possible that a systematic effect only creates E-modes and no B-modes, which would be extremely unfortunate as it could bias cosmological parameters without ever being detected by a B-mode analysis.

Acknowledgements. Something should probably be put in here...

References

Addison, G. E., Huang, Y., Watts, D. J., et al. 2016, ApJ, 818, 132

Asgari, M. & Heymans, C. 2018, arXiv e-prints [arXiv:1811.10596]

Asgari, M., Heymans, C., Hildebrandt, H., et al. 2018, arXiv e-prints [arXiv:1810.02353]

Bartelmann, M. & Schneider, P. 2001, Phys. Rep., 340, 291

Battye, R. A. & Moss, A. 2014, Physical Review Letters, 112, 051303

Blake, C. 2019, arXiv e-prints [arXiv:1902.07439]

Crittenden, R. G., Natarajan, P., Pen, U.-L., & Theuns, T. 2002, ApJ, 568, 20

Harnois-Déraps, J., Amon, A., Choi, A., et al. 2018, MNRAS, 481, 1337

Heymans, C., Grocutt, E., Heavens, A., et al. 2013, MNRAS, 432, 2433

Hikage, C., Oguri, M., Hamana, T., et al. 2019, PASJ[arXiv:1809.09148]

Hildebrandt, H., Köhlinger, F., van den Busch, J. L., et al. 2018, arXiv e-prints [arXiv:1812.06076]

Hildebrandt, H., Viola, M., Heymans, C., et al. 2017, MNRAS, 465, 1454

Köhlinger, F., Viola, M., Joachimi, B., et al. 2017, MNRAS, 471, 4412

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A14

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, ArXiv e-prints [arXiv:1807.06209]

Schneider, P., Eifler, T., & Krause, E. 2010, A&A, 520, A116

Schneider, P., van Waerbeke, L., & Mellier, Y. 2002, A&A, 389, 729

Shirasaki, M., Hamana, T., Takada, M., Takahashi, R., & Miyatake, H. 2019, arXiv e-prints [arXiv:1901.09488]

Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. 2012, ApJ, 761, 152

Troxel, M. A., MacCrann, N., Zuntz, J., et al. 2018, Phys. Rev. D, 98, 043528

Van Waerbeke, L., White, M., Hoekstra, H., & Heymans, C. 2006, Astroparticle Physics, 26, 91

Verde, L., Protopapas, P., & Jimenez, R. 2013, Physics of the Dark Universe, 2, 166

Appendix A: Detailed Calculations

Appendix A.1: Calculation of the power spectrum

In this Section we will perform the calculation for the observed power spectrum $P^{\mathrm{obs}}(\boldsymbol{\ell})$. For this, we assume an infinitely large field in order to perform our integration over $\mathbb{R}^2$. In reality, finite field effects would play a role here. We begin with the calculation of the correlation for the Fourier transformed shear:

$$
\begin{aligned}
&\left\langle \hat{\gamma}^{\mathrm{obs}}(\boldsymbol{\ell})\hat{\gamma}^{\mathrm{obs}*}(\boldsymbol{\ell}')\right\rangle \\
&\quad= \left\langle \int \mathrm{d}^2\theta \int \mathrm{d}^2\theta'\; W(\boldsymbol{\theta})W(\boldsymbol{\theta}')\gamma(\boldsymbol{\theta})\gamma^*(\boldsymbol{\theta}')\exp(\text{ß}\boldsymbol{\ell}\boldsymbol{\theta} - \text{ß}\boldsymbol{\ell}'\boldsymbol{\theta}')\right\rangle \\
&\quad= \left\langle \int \mathrm{d}^2\theta \int \mathrm{d}^2\theta'\; W(\boldsymbol{\theta})W(\boldsymbol{\theta}')\exp(\text{ß}\boldsymbol{\ell}\boldsymbol{\theta} - \text{ß}\boldsymbol{\ell}'\boldsymbol{\theta}') \int \frac{\mathrm{d}^2k}{(2\pi)^2}\int \frac{\mathrm{d}^2\ell}{(2\pi)^2}\,\tilde{\gamma}(\boldsymbol{k})\tilde{\gamma}^*(\boldsymbol{\ell})\exp(-\text{ß}\boldsymbol{k}\boldsymbol{\theta} + \text{ß}\boldsymbol{\ell}\boldsymbol{\theta}')\right\rangle \\
&\quad= \left\langle \int \mathrm{d}^2\theta \int \mathrm{d}^2\theta' \int \frac{\mathrm{d}^2k}{(2\pi)^2}\int \frac{\mathrm{d}^2\ell}{(2\pi)^2}\,P(\boldsymbol{k})(2\pi)^2\delta(\boldsymbol{k}-\boldsymbol{\ell})\exp[\text{ß}(\boldsymbol{\ell}\boldsymbol{\theta} - \boldsymbol{\ell}'\boldsymbol{\theta}' - \boldsymbol{k}\boldsymbol{\theta} + \boldsymbol{\ell}\boldsymbol{\theta}')]W(\boldsymbol{\theta})W(\boldsymbol{\theta}')\right\rangle \\
&\quad= \left\langle \int \frac{\mathrm{d}^2k}{(2\pi)^2}\,P(\boldsymbol{k}) \int \mathrm{d}^2\theta\, W(\boldsymbol{\theta})\exp[\text{ß}\boldsymbol{\theta}(\boldsymbol{\ell}-\boldsymbol{k})] \int \mathrm{d}^2\theta'\; W(\boldsymbol{\theta}')\exp[-\text{ß}\boldsymbol{\theta}(\boldsymbol{\ell}'-\boldsymbol{k})]\right\rangle \\
&\quad= \left\langle \int \frac{\mathrm{d}^2k}{(2\pi)^2}\,P(\boldsymbol{k})\widetilde{W}(\boldsymbol{\ell}-\boldsymbol{k})\widetilde{W}^*(\boldsymbol{\ell}'-\boldsymbol{k})\right\rangle
\end{aligned}
\tag{A.1}
$$

It is important to keep in mind that the ensemble averages of the weight function are independent of the ensemble averages of the shear values, meaning $\langle W(\boldsymbol{\theta})\gamma(\boldsymbol{\theta})\rangle = \langle W(\boldsymbol{\theta})\rangle\langle\gamma(\boldsymbol{\theta})\rangle$. We can define $W(\boldsymbol{\theta}) = 1 + w(\boldsymbol{\theta})$ with $\langle w(\boldsymbol{\theta})\rangle = 0$, which leads to the expession

$$
\begin{aligned}
&\left\langle \hat{\gamma}^{\mathrm{obs}}(\boldsymbol{\ell})\hat{\gamma}^{\mathrm{obs}*}(\boldsymbol{\ell}')\right\rangle \\
&\quad= \Bigg\langle \int \frac{\mathrm{d}^2k}{(2\pi)^2}\,P(\boldsymbol{k})\Big[(2\pi)^4\delta(\boldsymbol{\ell}-\boldsymbol{k})\delta(\boldsymbol{\ell}'-\boldsymbol{k}) + (2\pi)^2[\tilde{w}(\boldsymbol{\ell}-\boldsymbol{k})\delta(\boldsymbol{\ell}'-\boldsymbol{k}) + \tilde{w}^*(\boldsymbol{\ell}'-\boldsymbol{k})\delta(\boldsymbol{\ell}-\boldsymbol{k})] \\
&\qquad + \tilde{w}(\boldsymbol{\ell}-\boldsymbol{k})\tilde{w}(\boldsymbol{\ell}'-\boldsymbol{k})\Big]\Bigg\rangle \\
&\quad= (2\pi)^2\delta(\boldsymbol{\ell}-\boldsymbol{\ell}')P(\boldsymbol{\ell}) + [\langle\tilde{w}(\boldsymbol{\ell}-\boldsymbol{\ell}')\rangle P(\boldsymbol{\ell}') + \langle\tilde{w}^*(\boldsymbol{\ell}'-\boldsymbol{\ell})\rangle P(\boldsymbol{\ell})] + \left\langle \int \frac{\mathrm{d}^2k}{(2\pi)^2}\,\tilde{w}(\boldsymbol{\ell}-\boldsymbol{k})\tilde{w}^*(\boldsymbol{\ell}'-\boldsymbol{k})P(\boldsymbol{k})\right\rangle \\
&\quad\overset{(*)}{=} (2\pi)^2\delta(\boldsymbol{\ell}-\boldsymbol{\ell}')P(\boldsymbol{\ell}) + \left\langle \int \frac{\mathrm{d}^2k}{(2\pi)^2}\,\tilde{w}(\boldsymbol{\ell}-\boldsymbol{k})\tilde{w}^*(\boldsymbol{\ell}'-\boldsymbol{k})P(\boldsymbol{k})\right\rangle,
\end{aligned}
\tag{A.2}
$$

where in $(*)$ we have used that the average $\langle\tilde{w}(\boldsymbol{\ell})\rangle$ vanishes. Up until now, we have not specified our weight-function w. We parametrize it as

$$
w(\boldsymbol{\theta}) = \sum_{\boldsymbol{\alpha}\in\mathbb{Z}^2} w_{\boldsymbol{\alpha}}\Xi(\boldsymbol{\theta} - L\boldsymbol{\alpha})\,,\ \text{with the Box-Function } \Xi(\boldsymbol{\theta}) = \begin{cases} 1 & \boldsymbol{\theta} \in \left[-\frac{L}{2}, \frac{L}{2}\right]^2 \\ 0 & \text{else} \end{cases}\,. \tag{A.3}
$$

Here, the $w_{\boldsymbol{\alpha}}$ are random variables, drawn from the random distribution describing the survey depths. For the Fourier-Transform we compute:

$$
\tilde{w}(\boldsymbol{\ell}) = \sum_{\boldsymbol{\alpha}\in\mathbb{Z}^2} w_{\boldsymbol{\alpha}}\exp(-\text{ß}\boldsymbol{\ell}L\boldsymbol{\alpha})\widetilde{\Xi}(\boldsymbol{\ell})\,, \tag{A.4}
$$

where

$$\widetilde{\Xi}(\boldsymbol{\ell}) = \frac{4\sin\left(\frac{L\ell_1}{2}\right)\sin\left(\frac{L\ell_2}{2}\right)}{\ell_1\ell_2}\,, \tag{A.5}$$

is a 2-dimensional sinc function. Assuming an uncorrelated weight-distribution $\left(\left\langle w_{\boldsymbol{\alpha}} w_{\boldsymbol{\beta}}\right\rangle = 0 \text{ for } \boldsymbol{\alpha} \neq \boldsymbol{\beta}\right)$ and setting $\left\langle w^2\right\rangle \equiv \left\langle w_{\boldsymbol{\alpha}}^2\right\rangle$ for each $\boldsymbol{\alpha}$, we get

$$\begin{aligned}
&\left\langle \int \frac{\mathrm{d}^2k}{(2\pi)^2}\, \tilde{w}(\boldsymbol{\ell}-\boldsymbol{k})\tilde{w}^*(\boldsymbol{\ell}'-\boldsymbol{k})P(\boldsymbol{k})\right\rangle \\
&\quad= \left\langle \int \frac{\mathrm{d}^2k}{(2\pi)^2} \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} w_{\boldsymbol{\alpha}} w_{\boldsymbol{\beta}} \exp[-\text{ß}(\boldsymbol{\ell}-\boldsymbol{k})L\boldsymbol{\alpha}]\,\widetilde{\Xi}(\boldsymbol{\ell}-\boldsymbol{k})\exp[\text{ß}(\boldsymbol{\ell}'-\boldsymbol{k})L\boldsymbol{\beta}]\,\widetilde{\Xi}^*(\boldsymbol{\ell}'-\boldsymbol{k})P(\boldsymbol{k})\right\rangle \\
&\quad= \int \frac{\mathrm{d}^2k}{(2\pi)^2} \sum_{\boldsymbol{\alpha}} \left\langle w^2\right\rangle \exp[-\text{ß}(\boldsymbol{\ell}-\boldsymbol{k})L\boldsymbol{\alpha} + i(\boldsymbol{\ell}'-\boldsymbol{k})L\boldsymbol{\alpha}]\,\widetilde{\Xi}(\boldsymbol{\ell}-\boldsymbol{k})\widetilde{\Xi}^*(\boldsymbol{\ell}'-\boldsymbol{k})P(\boldsymbol{k})\,.
\end{aligned} \tag{A.6}$$

Using this result, we can obtain the observed power spectrum

$$P^{\mathrm{obs}}(\boldsymbol{\ell}) = \frac{1}{(2\pi)^2}\int \mathrm{d}^2\ell' \left\langle \hat{\gamma}^{\mathrm{obs}}(\boldsymbol{\ell})\hat{\gamma}^{\mathrm{obs}*}(\boldsymbol{\ell}')\right\rangle\,, \tag{A.7}$$

by performing the $\boldsymbol{\ell}'$-integration in (A.2):

$$\begin{aligned}
P^{\mathrm{obs}}(\boldsymbol{\ell}) =& P(\boldsymbol{\ell}) + \int \frac{\mathrm{d}^2\boldsymbol{\ell}'}{(2\pi)^2}\int \frac{\mathrm{d}^2\boldsymbol{k}}{(2\pi)^2} \sum_{\boldsymbol{\alpha}} \left\langle w^2\right\rangle \exp[-\text{ß}(\boldsymbol{\ell}-\boldsymbol{k})L\boldsymbol{\alpha} + \text{ß}(\boldsymbol{\ell}'-\boldsymbol{k})L\boldsymbol{\alpha}]\,\widetilde{\Xi}(\boldsymbol{\ell}-\boldsymbol{k})\widetilde{\Xi}(\boldsymbol{\ell}'-\boldsymbol{k})P(\boldsymbol{k}) \\
=& P(\boldsymbol{\ell}) + \int \frac{\mathrm{d}^2\boldsymbol{k}}{(2\pi)^2} \sum_{\boldsymbol{\alpha}} \left\langle w^2\right\rangle \exp[-\text{ß}(\boldsymbol{\ell}-\boldsymbol{k})L\boldsymbol{\alpha}]\,\widetilde{\Xi}(\boldsymbol{\ell}-\boldsymbol{k})P(k)\int \frac{\mathrm{d}^2\boldsymbol{\ell}}{(2\pi)^2}\,\widetilde{\Xi}^*(\boldsymbol{\ell}'-\boldsymbol{k})\exp[\text{ß}(\boldsymbol{\ell}'-\boldsymbol{k})L\boldsymbol{\alpha}] \\
=& P(\boldsymbol{\ell}) + \left\langle w^2\right\rangle \int \frac{\mathrm{d}^2\boldsymbol{k}}{(2\pi)^2}\,\widetilde{\Xi}(\boldsymbol{\ell}-\boldsymbol{k})P(\boldsymbol{k}) \sum_{\boldsymbol{\alpha}} \exp[-\text{ß}(\boldsymbol{\ell}-\boldsymbol{k})L\boldsymbol{\alpha}]\,\Xi(L\boldsymbol{\alpha}) \\
=& P(\boldsymbol{\ell}) + \left\langle w^2\right\rangle \int \frac{\mathrm{d}^2\boldsymbol{k}}{(2\pi)^2}\,\widetilde{\Xi}(\boldsymbol{\ell}-\boldsymbol{k})P(\boldsymbol{k})\,,
\end{aligned} \tag{A.8}$$

which is a convolution of the power spectrum and the 2-dimensional sinc function.

Appendix A.2: The function $E(\theta)$

When computing the shear correlation between a pair of galaxies, it is of central importance whether those two galaxies lie in the same pointing or not. We want to model the probability that a pair of galaxies with separation $\boldsymbol{\theta}$ lie in the same pointing by the function $E(\boldsymbol{\theta})$, which we will derive here:

Given one square field of length L (in our case $L = 60'$) and a separation vector $\boldsymbol{\theta}$, without loss of generality we can assume $\theta_1, \theta_2 \geq 0$. As depicted in Figure A.1, the dashed square represents all possible positions that the first galaxy can take, such that the second galaxy is still within the same pointing. The volume of this square equals

$$V(|\boldsymbol{\theta}|, \phi) = [L - |\boldsymbol{\theta}|\cos(\phi)]\,[L - |\boldsymbol{\theta}|\sin(\phi)]\,, \tag{A.9}$$

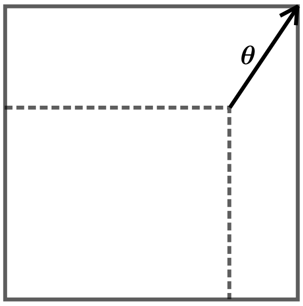


Fig. A.1: Graphic representation on how to obtain the function $E(\boldsymbol{\theta})$. For a separation vector $\boldsymbol{\theta}$, the dashed square represents the area of galaxies that have their partner in the same pointing.

where ϕ represents the angle of the vector $\boldsymbol{\theta}$. The function $E(\boldsymbol{\theta})$ then simply equals $V(|\boldsymbol{\theta}|, \phi)/L^2$. To exclude negative Volumes (which could occur when $|\boldsymbol{\theta}| > 1$ holds), we need to add the Heaviside theta function $\mathcal{H}$:

$$E(\boldsymbol{\theta}) = \left[1 - \frac{|\boldsymbol{\theta}|}{L}\cos(\phi)\right]\left[1 - \frac{|\boldsymbol{\theta}|}{L}\sin(\phi)\right]\mathcal{H}\left[1 - \frac{|\boldsymbol{\theta}|}{L}\cos(\phi)\right]\mathcal{H}\left[1 - \frac{|\boldsymbol{\theta}|}{L}\sin(\phi)\right]. \tag{A.10}$$

As $E(\boldsymbol{\theta})$ is not isotropic, in order to obtain the function $E(\theta) = E(|\boldsymbol{\theta}|)$, we need to azimuthally average Equation (A.10) over all angles ϕ. While the case $\theta_1, \theta_2 \geq 0$ certainly does not hold for all angles ϕ, we can eliminate the other cases by simple symmetry.

$$E(\theta) = \frac{4}{2\pi}\int_0^{\frac{\pi}{2}} \mathrm{d}\phi\, E(\boldsymbol{\theta}) = \frac{2}{\pi}\begin{cases} \int_0^{\frac{\pi}{2}} \mathrm{d}\phi \left[1 - \frac{|\boldsymbol{\theta}|}{L}\cos(\phi)\right]\left[1 - \frac{|\boldsymbol{\theta}|}{L}\sin(\phi)\right], & |\boldsymbol{\theta}| \leq L \\ \int_{\cos^{-1}(L/|\boldsymbol{\theta}|)}^{\sin^{-1}(L/|\boldsymbol{\theta}|)} \mathrm{d}\phi \left[1 - \frac{|\boldsymbol{\theta}|}{L}\cos(\phi)\right]\left[1 - \frac{|\boldsymbol{\theta}|}{L}\sin(\phi)\right], & L \leq |\boldsymbol{\theta}| \leq \sqrt{2}L \\ 0, & \sqrt{2}L \leq \theta \end{cases}$$

$$= \begin{cases} \frac{1}{L^2\pi}\left[L^2\pi - (4L - \theta)\theta\right], & \theta \leq L \\ \frac{2}{\pi}\left[4\sqrt{\frac{\theta^2}{L^2} - 1} - 1 - \frac{\theta^2}{2L^2} - \cos^{-1}\left(\frac{L}{\theta}\right) + \sin^{-1}\left(\frac{L}{\theta}\right)\right], & L \leq \theta \leq \sqrt{2}L \\ 0, & \sqrt{2}L \leq \theta \end{cases}. \tag{A.11}$$

Appendix A.3: Calculation of the shear correlation functions

Following Hildebrandt et al. (2017), given a set of galaxies we calculate the shear correlation functions via

$$\xi_+^{ij}(\theta) = \frac{\sum_{a,b} w_a^i w_b^j \epsilon_a^i \epsilon_b^{j*} \Delta(|\boldsymbol{\theta}_a^i - \boldsymbol{\theta}_b^i|)}{\sum_{a,b} w_a^i w_b^j \Delta(|\boldsymbol{\theta}_a^i - \boldsymbol{\theta}_b^i|)}. \tag{A.12}$$

Here, w represents the lensing weight of the galaxy, whereas ϵ is its (complex) ellipticity and $\boldsymbol{\theta}$ its position on the sky. We have defined the function Δ as

$$\Delta(|\boldsymbol{\theta}^i_a - \boldsymbol{\theta}^i_b|) = \begin{cases} 1, & |\boldsymbol{\theta}^i_a - \boldsymbol{\theta}^j_b| \in [\theta, \theta + \mathrm{d}\theta] \\ 0, & \text{else} \end{cases} , \tag{A.13}$$

where we assume $\mathrm{d}\theta \ll \theta$. We define N as the number of pointings in the survey and F^i_k as the set of galaxies in pointing k and tomographic bin i. The numerator in Equation (A.12) then transforms to:

$$\begin{aligned} &\sum_{k,\ell=1}^{N} \sum_{a\in F^i_k} \sum_{b\in F^j_\ell} w^i_a w^j_b \epsilon^i_a \epsilon^{j*}_b \Delta(|\boldsymbol{\theta}^i_a - \boldsymbol{\theta}^i_b|) \\ &= \sum_{k=1}^{N} \sum_{a\in F^i_k} w^i_a \sum_{\ell=1}^{N} \sum_{b\in F^j_\ell} w^j_b \Delta(|\boldsymbol{\theta}^i_a - \boldsymbol{\theta}^i_b|)\, \epsilon^i_a \epsilon^{j*}_b \\ &= \sum_{k=1}^{N} \sum_{a\in F^i_k} w^i_a \left[\sum_{b\in F^j_k} w^j_b \Delta(|\boldsymbol{\theta}^i_a - \boldsymbol{\theta}^i_b|)\, \epsilon^i_a \epsilon^{j*}_b + \sum_{\ell\neq k} \sum_{b\in F^j_\ell} w^j_b \Delta(|\boldsymbol{\theta}^i_a - \boldsymbol{\theta}^i_b|)\, \epsilon^i_a \epsilon^{j*}_b \right] . \end{aligned} \tag{A.14}$$

When we denote the probability that pointing k is of percentile m by P^k_m and assume that the product $\epsilon^i_a \epsilon^{j*}_b$ always equals its expectation value, we can set the numerator as

$$\sum_{k=1}^{N} \sum_{a\in F^i_k} w^i_a \sum_m P^k_m \left[\overbrace{\sum_{b\in F^j_k} w^j_b \Delta(|\boldsymbol{\theta}^i_a - \boldsymbol{\theta}^i_b|)}^{\text{(A.15.a)}}\ \xi^{ij}_{+,mm}(\theta) + \overbrace{\sum_{\ell\neq k} \sum_{b\in F^j_\ell} w^j_b \Delta(|\boldsymbol{\theta}^i_a - \boldsymbol{\theta}^i_b|)}^{\text{(A.15.b)}} \sum_n P^\ell_n \xi^{ij}_{+,mn}(\theta) \right] . \tag{A.15}$$

The term (A.15.a) denotes all galaxies that lie within distance interval $[\theta, \theta+\mathrm{d}\theta]$ of galaxy a, and are in the same pointing as galaxy a. This term is equal to the (weighted) number density of galaxies in the pointing multiplied by $2\pi\theta\,\mathrm{d}\theta\, E(\theta)$.

The term (A.15.b) denotes all galaxies within distance interval $[\theta, \theta + \mathrm{d}\theta]$ of galaxy a, that are *not* in the same pointing as galaxy a. This is equal to the number density of galaxies in the respective pointings multiplied by $2\pi\theta\,\mathrm{d}\theta\,[1 - E(\theta)]$.

If we assume that said number density in a pointing is equal to the number density in the percentile it belongs to, N^j_n, and set $P^\ell_n = 1/10$, the numerator becomes

$$\sum_{k=1}^{N} \sum_{a\in F^i_k} w^i_a \sum_m P^k_m \left[2\pi\theta\,\mathrm{d}\theta\, E(\theta) N^j_m \xi^{ij}_{+,mm}(\theta) + 2\pi\theta\,\mathrm{d}\theta\, \frac{1 - E(\theta)}{10} \sum_n N^j_n \xi^{ij}_{+,mn}(\theta) \right] . \tag{A.16}$$

Now the term $\sum_{a\in F^i_k} w^i_a$ denotes the (weighted) number of galaxies in pointing k, which we set as the number density of galaxies in the respective percentile multiplied with the area A of the

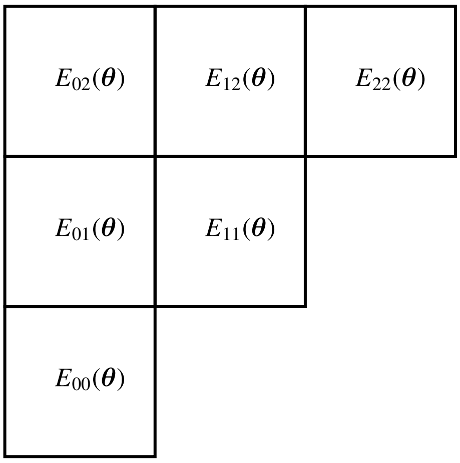


Fig. B.1: Graphic representation of the definitions of $E_{ab}(\theta)$. When the first galaxy is in the bottom left pointing, the probability to find the second galaxy in a pointing of distance (a, b) is $E_{mn}(\theta)$.

pointing. Applying this and setting $P^k_m = 1/10$, the numerator reads

$$\begin{aligned}
&\frac{2\pi\theta\,\mathrm{d}\theta}{10}\sum_{k=1}^{N}\sum_m N^i_m A\left[E(\theta)N^j_m\xi^{ij}_{+,mm}(\theta) + \frac{1-E(\theta)}{10}\sum_n N^j_n\xi^{ij}_{+,mn}(\theta)\right] \\
&= \frac{2\pi\theta\,\mathrm{d}\theta\,NA}{10}\sum_m N^i_m\left[E(\theta)N^j_m\xi^{ij}_{+,mm}(\theta) + \frac{1-E(\theta)}{10}\sum_n N^j_n\xi^{ij}_{+,mn}(\theta)\right].
\end{aligned} \tag{A.17}$$

The same line of argumentation can be applied to the denominator, which then reads:

$$\frac{2\pi\theta\,\mathrm{d}\theta\,NA}{10}\sum_m N^i_m\left[E(\theta)N^j_m + \frac{1-E(\theta)}{10}\sum_n N^j_n\right]. \tag{A.18}$$

Taking the ratio of the two quantities, we see that Equations (A.12) and (19) are the same[8].

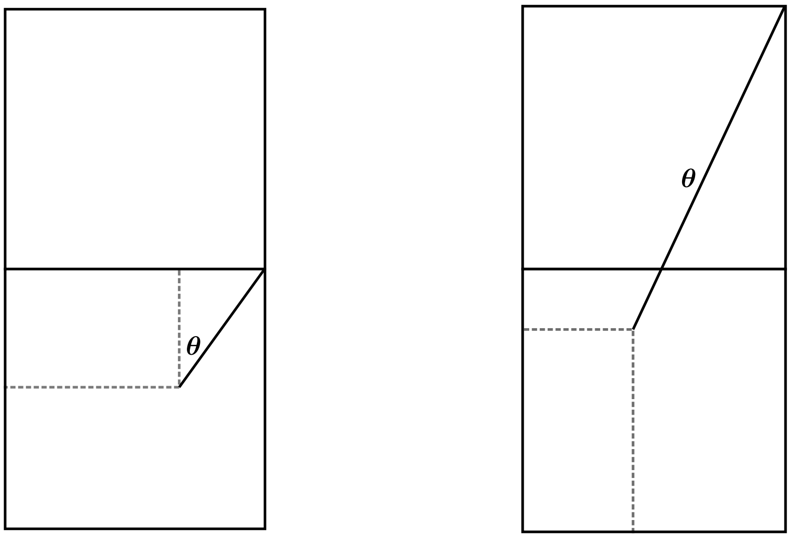
Appendix B: Outlook: Finite field effects

In this chapter we will outline how to calculate the correction of the correlation functions for a finite survey with a potentially correlated distribution of depth between pointings. Essentially, this boils down to the calculation of $P^{ij}_{mn}(\theta)$ from Equation (15). We calculate this weighting by the geometrical probability that a pair of galaxies of separation θ is of percentiles m and n, $P(m, n|\theta)$, weighted by the respective number of galaxies in the percentiles N^i_m, N^j_n:

$$P^{ij}_{mn}(\theta) = N^i_m N^j_n P(m, n|\theta). \tag{B.1}$$

At first we define Functions $E_{ab}(\boldsymbol{\theta})$ as the probabilty that a galaxy pair of separation $\boldsymbol{\theta}$ is in pointings of distance (a, b). This situation is depicted in Figure B.1. Due to symmetry, for the azimuthal average of the functions, $E_{ab}(\theta) = E_{-ab}(\theta) = E_{ba}(\theta)$ holds for all combinations of a and b. Note that $E_{00}(\theta) = E(\theta)$ and $\sum_{a,b} E_{ab}(\theta) \equiv 1$.

[8] Note that while here N^i_m denotes a number density, in Equations (A.12) and (19) it denotes the total (weighted) number of galaxies. However, the difference is just a multiplication with the area A of the pointings, which appears both in the numerator and the denominator and is thus cancelled out.



(a) For $\theta \sin(\phi) < L$ the volume of the dashed rectangle is $V(\theta, \phi) = \theta \sin(\phi)[L - \theta \cos(\phi)]$.

(b) For $\theta \sin(\phi) > L$ the volume of the dashed rectangle is $V(\theta, \phi) = [2L - \theta \sin(\phi)]\,[L - \theta \cos(\phi)]$.

Fig. B.2: How to calculate $E_{01}(\theta)$ for different values of θ.

Let $P^*(m, n|a, b)$ denote the probability that two pointings of distance (a, b) are of percentile m and n (which is directly calculable from a given survey footprint). Then the following equation holds:

$$P(m, n|\theta) = \sum_{a,b} E_{ab}(\theta) P^*(m, n|a, b)\,. \tag{B.2}$$

Note that the expectation value of $P^*(m, n|a, b)$ for uncorrelated distributions is

$$\langle P^*(m, n|a, b)\rangle = \begin{cases} 0.1\,\delta_{mn}, & \text{for } (a, b) = (0, 0) \\ 0.01, & \text{else} \end{cases}\,, \tag{B.3}$$

where δ_{mn} denotes the Kronecker delta. Keeping in mind that

$$\sum_{(a,b)\neq(0,0)} E_{ab}(\theta) = 1 - E(\theta)\,, \tag{B.4}$$

we can use the expectation value (B.3) to calculate (B.2) as a consistency check. In that case, we receive the same value for the coefficients in (B.1) as we have in Equation (18) in Chapter **??** for the case of an infinite footprint and uncorrelated distribution of depth.

The E_{ab} can all be calculated analytically, similar to our method in Chapter A.2. We again assume a selection of square fields with side length L, and later set $L = 60'$ to adapt to the KV450 survey. As an example, for E_{01} we have several possible situations, depicted in Figure B.2. For a separation vector $\boldsymbol{\theta}$ with modulus $\theta < L$ and inclination angle ϕ, the dashed rectangle in Figure B.2a depicts the galaxies that have a partner in the upper pointing. For $L < \theta < \sqrt{2}L$, the dashed rectangle in Figure B.2a represents the area of galaxies with a partner in the upper square under the condition that $\theta \sin(\phi) < L$. For $\theta \sin(\phi) > L$, the situation is depicted by Figure B.2b. As soon as $\theta > \sqrt{2}L$ holds, the situation of Figure B.2a is impossible. When $\theta < 2L$ holds, the vector is only constrained by $\theta\cos(\phi) < L$, but as soon as $\theta > 2L$ holds, we additionally need to impose $\theta\sin(\phi) < 2L$. Again setting $E_{ab}(\boldsymbol{\theta}) = V(\theta,\phi)/L^2$, we define

$$\begin{aligned}
E_{01}^{(a)}(\boldsymbol{\theta}) &\equiv \frac{\theta}{L}\sin(\phi)[1-\frac{\theta}{L}\cos(\phi)] \\
E_{01}^{(b)}(\boldsymbol{\theta}) &\equiv [2-\frac{\theta}{L}\sin(\phi)]\,[1-\frac{\theta}{L}\cos(\phi)]
\end{aligned} \tag{B.5}$$

Taking the azimuthal average, we compute:

$$\begin{aligned}
E_{01}(\theta) &= \begin{cases}
\frac{1}{\pi}\int_0^{\frac{\pi}{2}} \mathrm{d}\phi\, E_{01}^{(a)}(\boldsymbol{\theta}) & \frac{\theta}{L} < 1 \\
\frac{1}{\pi}\left[\int_{\cos^{-1}(L/\theta)}^{\sin^{-1}(L/\theta)} \mathrm{d}\phi\, E_{01}^{(a)}(\boldsymbol{\theta}) + \int_{\sin^{-1}(L/\theta)}^{\frac{\pi}{2}} \mathrm{d}\phi\, E_{01}^{(b)}(\boldsymbol{\theta})\right] & 1 < \frac{\theta}{L} < \sqrt{2} \\
\frac{1}{\pi}\int_{\cos^{-1}(L/\theta)}^{\frac{\pi}{2}} \mathrm{d}\phi\, E_{01}^{(b)}(\boldsymbol{\theta}) & \sqrt{2} < \frac{\theta}{L} < 2 \\
\frac{1}{\pi}\int_{\cos^{-1}(L/\theta)}^{\sin^{-1}(2L/\theta)} \mathrm{d}\phi\, E_{01}^{(b)}(\boldsymbol{\theta}) & 2 < \frac{\theta}{L} < \sqrt{5} \\
0 & \sqrt{5} < \frac{\theta}{L}
\end{cases} \\
&= \begin{cases}
\frac{(2L-\theta)\theta}{2\pi L^2} & \frac{\theta}{L} < 1 \\
\frac{1}{\pi}\left[\frac{3}{2} - 2\frac{\theta}{L} + \frac{\theta^2}{L^2} + 2\sqrt{\frac{\theta^2}{L^2}-1} + 2\sec^{-1}\left(\frac{\theta}{L}\right)\right] & 1 < \frac{\theta}{L} < \sqrt{2} \\
\frac{1}{2\pi}\left[-1 - 4\frac{\theta}{L} + 4\sqrt{\frac{\theta^2}{L^2}-1} + 4\csc^{-1}\left(\frac{\theta}{L}\right)\right] & \sqrt{2} < \frac{\theta}{L} < 2 \\
\frac{1}{2\pi}\left[-5 - \frac{\theta^2}{L^2} + 2\sqrt{\frac{\theta^2}{L^2}-4} + 4\sqrt{\frac{\theta^2}{L^2}-1} - 4\sec^{-1}\left(\frac{\theta}{L}\right) + 4\sin^{-1}\left(\frac{2L}{\theta}\right)\right] & 2 < \frac{\theta}{L} < \sqrt{5} \\
0 & \sqrt{5} < \frac{\theta}{L}
\end{cases} .
\end{aligned} \tag{B.6}$$

Naturally, to calculate those functions for all possible combinations would be rather tedious, however they are simple to determine numerically (compare Figure B.3). A plot of these functions can be found in Figure B.4.

When we now simulate random distributions of the depth-function for a $100\,\mathrm{deg}^2$-field, a $450\,\mathrm{deg}^2$-field and a $1000\,\mathrm{deg}^2$-field, we can compare how they differ from each other and estimate how important finite-field effects are. As can be seen from Figures **??**, **??** and **??**, the effect is quite significant for a $100\,\mathrm{deg}^2$-field, but almost negligible for a $1000\,\mathrm{deg}^2$-field. This leads to the assumption that both for the KiDS- as for the Euclid-survey, finite field effects do not need to be accounted for. However, if the distribution of depth is correlated in the surveys, that might have a noticeable impact on the results.

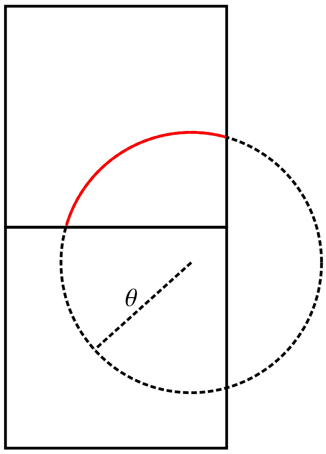


Fig. B.3: Visualisation of the numerical computation for $E_{01}(\theta)$. For a circle of radius θ, the length of the red arc divided by 2π represents the fraction of galaxies within the respective pointing. This value needs to be integrated for all possible centers of the circle in the pointing. That procedure is straightforward to expand for other $E_{ab}(\theta)$.

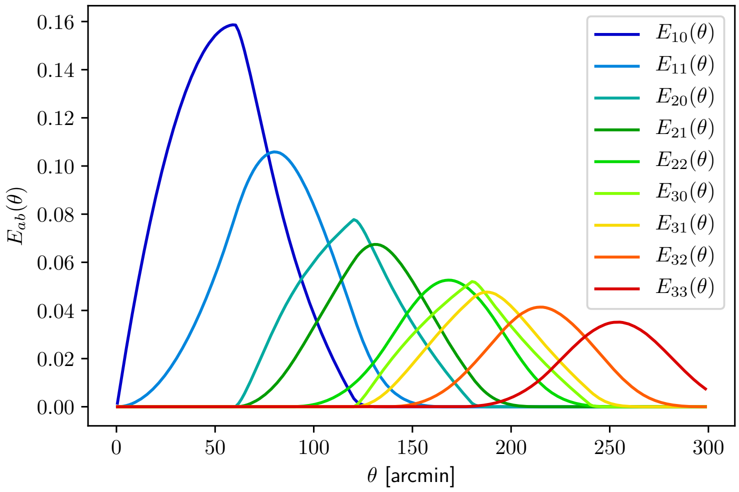


Fig. B.4: The functions $E_{ab}(\theta)$ for the first few possible combinations.