Forward-Simulations of Large-Scale Structure for Cosmological Inference

Maximilian von Wietersheim-Kramsta

A dissertation submitted in fulfillment
of the requirements for the degree of
Doctor of Philosophy
of
Department of Physics and Astronomy
University College London.

Supervisors:
Prof. Benjamin Joachimi
Dr Andreu Font-Ribera

Examiners:
Prof. Ofer Lahav
Prof. David Alonso

December 8, 2023

Statement of Originality

I, Maximilian von Wietersheim-Kramsta, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work. One of my research chapters is based on my paper:

- Maximilian von Wietersheim-Kramsta, Benjamin Joachimi, Jan Luca van den Busch, Catherine Heymans, Hendrik Hildebrandt, Marika Asgari, Tilman Tröster, Sandra Unruh, and Angus H. Wright (2021). Magnification bias in galaxy surveys with complex sample selection functions. *Monthly Notices of the Royal Astronomical Society*, 504(1), pp.1452-1465.

closely following the exposition therin. I have also contributed to:

- B. Joachimi, C.-A. Lin, M. Asgari, T. Tröster, C. Heymans, H. Hildebrandt, F. Köhlinger, A. G. Sánchez, A. H. Wright, M. Bilicki, C. Blake, J. L. van den Busch, M. Crocce, A. Dvornik, T. Erben, F. Getman, B. Giblin, H. Hoekstra, A. Kannawadi, K. Kuijken, N. R. Napolitano, P. Schneider, R. Scoccimarro, E. Sellentin, H. Y. Shan, M. von Wietersheim-Kramsta, and J. Zuntz (2021). KiDS-1000 Methodology: Modelling and inference for joint weak gravitational lensing and spectroscopic galaxy clustering analysis. *Astronomy & Astrophysics*, 646, A129.

- Maria Cristina Fortuna, Henk Hoekstra, Harry Johnston, Mohammadjavad Vakili, Arun Kannawadi, Christos Georgiou, Benjamin Joachimi, Angus H. Wright, Marika Asgari, Maciej Bilicki, Catherine Heymans, Hendrik Hildebrandt, Konrad Kuijken, Maximilian Von Wietersheim-Kramsta (2021). KiDS-1000: Constraints on the intrinsic alignment of luminous red galaxies. *Astronomy & Astrophysics*, 654, A76.

- Kiyam Lin, Maximilian von Wietersheim-Kramsta, Benjamin Joachimi, and Stephen Feeney (2023). A simulation-based inference pipeline for cosmic shear with the Kilo-Degree Survey. *Monthly Notices of the Royal Astronomical Society*, 524(4), pp.6167–6180.

- Nicolas Tessore, Arthur Loureiro, Benjamin Joachimi, Maximilian von Wietersheim-Kramsta, Niall Jeffrey (2023). GLASS: Generator for Large Scale Structure. *The Open Journal of Astrophysics*, 6, p.11.

Signed: Maximilian von Wietersheim-Kramsta

Abstract

Measuring the density of fluctuations in the large-scale structure of the Universe has become a powerful tool in constraining and testing cosmological models. Current and upcoming galaxy surveys are pushing the precision limits of current physical models, while measuring unprecedented amounts of data. This thesis presents statistical simulations of large-scale structure and inference techniques for estimating cosmological parameters and calibrating systematic effects from galaxy survey data. First, I present a novel method to calibrate magnification bias observed within galaxy clustering and weak gravitational lensing measurements regardless of the selection applied to the galaxy data. This method addresses the need for estimating this systematic within the Kilo-Degree Survey's (KiDS) cosmological analysis. Secondly, I show a suite of statistical forward-simulations of large-scale structure which is designed to model galaxy survey observations, while including relevant physical and observational biases. These simulations create realistic catalogues of galaxy observations based on underlying matter density fields consistent with a given cosmological model. Next, I describe how these simulations are used to conduct the first Bayesian simulation-based inference (SBI) of cosmological parameters from weak lensing data from KiDS at a similar precision as standard analyses. This SBI analysis allows dropping the common assumption of a Gaussian likelihood, fully propagating uncertainty from data to parameter posteriors at a comparable computational cost as standard weak gravitational lensing analysis. Thus, this may facilitate the efficient extraction of information from surveys such as *Euclid* or the Vera Rubin Observatory. Lastly, I use an altered form of the forward-simulations to test signal and uncertainty modelling conducted for KiDS's analysis. This shows the sufficiency of analytical modelling when considering systematics such as variations in observational depth. In summary, in this thesis, I present novel

techniques to estimate systematic uncertainties in inferring cosmological parameters from galaxy surveys and I show how fast realistic forward-simulations may be used for SBI and model testing in current and future surveys.

Impact Statement

This thesis presents novel methods and tools to improve the inference of cosmological parameters from weak gravitational lensing and galaxy clustering measurements. With that, it makes concrete predictions and suggestions for current and upcoming galaxy surveys such as the Kilo-Degree Survey, the Dark Energy Survey, the Hyper-Suprime Cam survey, *Euclid*, and the Vera Rubin Observatory. Among these methods is the first application of simulation-based inference to a cosmic shear analysis equivalent to a traditional analysis. At the same time, the research presented in this thesis directly contributes to the analysis choices for the main cosmic shear and galaxy clustering analyses of the Kilo-Degree Survey.

With this, this work may contribute to the resolution of the greatest questions in cosmology: understanding of nature of dark matter, dark matter and gravity.

Additionally, the investigations shown in this work have required the development of multiple software packages which have subsequently been made publicly available, and which are already in use by other researchers.

To my wife and my parents

Acknowledgements

Firstly, I want to thank my supervisor, Benjamin, for his insights and guidance throughout this whole project. I am particularly grateful for the support received during the worldwide pandemic which struck shortly after the beginning of my postgraduate program. I also want thank Andreu and Stephen for their support and help as second supervisors. I would also like to express my sincerest gratitude to Kiyam and Nicolas; collaborating with them has made many of the results in this thesis possible. Of course, I am also more than grateful for all the great people with which I have had the joy to work and spent time at UCL: Lillian, Catarina, Benjamin, Francesca, Pascal, Matt, Claudia, Arthur, Alex, Dirk, Andrew, Alessio, Niall, Ellen, Marcus, Harry, Xie, Paul, Davide and Lorne. I also want to thank all the staff within the Cosmoparticle Initiative and the UCL Astrophysics Group for welcoming me. In particular, I have to highlight Raphaële and Anita for making the Cosmoparticle Hub such a pleasant workplace, John and Edd for always helping with anything computing-related, as well as Amelie, Nick and Nadia for always helping out with any admin matters.

Many thanks to the examiners, Ofer Lahav and David Alonso, for taking the time to read this thesis. I also want to thank the entire Kilo-Degree Survey consortium for giving me the opportunity to contribute to an exciting scientific endeavor. Special thanks to Marika Asgari, Angus H. Wright, Robert Reichke, Jan Luca van den Busch, Chieh-An Lin, Sven Heydenreich, Sandra Unruh, Konrad Kuijken, Catherine Heymans, Hendrik Hildebrandt, Tilman Tröster, Christos Georgiou, Maria Cristina Fortuna, Henk Hoekstra, Andrej Dvornik and Shahab Joudaki. Lastly, I want to express gratitude to the Cosmoparticle Initiative for providing its computational resources used in this research, and thanks to the Science and Technology Facilities Council for funding this research.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Background

1.1 The Standard Model of Cosmology

Cosmology concerns itself with the study of the physical laws which describe the formation, evolution and composition of the Universe at the largest scales. Over the last few decades, a consensus has emerged within the field of cosmology based on a single model being able describe most, if not all, astronomical observations that have been made so far. The model in question is known as the "Standard Model of Cosmology" or the Lambda Cold Dark Matter (ΛCDM) model (Peebles, 1982; Blumenthal et al., 1984). With the advent of precision cosmology and the proliferation of experiments probing different scales and time domains, the ΛCDM model has been remarkably successful in explaining the observations. ΛCDM is in agreement with observations at large scales in the early Universe from the cosmic microwave background (Mather et al., 1990; Bennett et al., 2013; Planck Collaboration et al., 2020), while also being in agreement with late-Universe probes at different scales: supernovae (Riess et al., 1998; Perlmutter et al., 1999), Baryonic Acoustic Oscillations (BAOs; Cole et al. 2005; Eisenstein et al. 2005; Blake et al. 2011; Beutler et al. 2011; Anderson et al. 2012), cosmic shear (Kaiser et al., 2000; Wittman et al., 2000; Van Waerbeke et al., 2000; Bacon et al., 2000; Heymans et al., 2013; Hildebrandt et al., 2017; Asgari et al., 2021; Amon et al., 2022, 2023), and many more. Having said this, it is worth noting that some potential inconsistencies within ΛCDM have been found in recent years (H_0 tension, S_8 discrepancy, etc.), but these could still be consistent with systematic effects.

The standard model of cosmology is built upon the general theory of relativity (GR; Einstein 1915; Hilbert 1915). As is shown in detail in Section 1.1.1, GR establishes the geometry of the Universe, the dynamics of this geometry and the interaction between the geometry of the Universe and its contents. Additionally, ΛCDM makes a series of assumptions within the framework of GR often referred to as the "cosmological principle". This principle states that at sufficiently large scales the Universe is spatially homogeneous and isotropic (Robertson, 1936; Walker, 1937). In other words, the cosmological principle is an assumption which implies that observations of the Universe at large scales are independent of the position of the frame-of-reference within the Universe. As a consequence, one can derive the Friedmann-Lemaître-Robertson-Walker metric (Friedmann, 1924; Lemaître, 1927) which ingrains homogeneity and isotropy within field equations of GR (see Section 1.1.2 for a detailed discussion).

With GR and the cosmological principle, the dynamics of ΛCDM are determined. To model the contents of the Universe which are subject to these dynamics, in Section 1.1.3, I present the "ingredients" of ΛCDM. These can be classed into broadly four types: baryonic matter, cold-dark matter, dark energy, and radiation. Baryonic matter includes all non-relativistic matter which can interact with itself and radiation through the forces described by the Standard Model of Particles Physics (see Mann 2010 for a review). Cold dark matter is a term used for non-relativistic matter which does not couple to any of the forces in the Standard Model of Particle Physics nor does it interact with itself. Cold dark matter is assumed to only have a causal effect on other components of the Universe through gravitational interactions. The third type of ingredient of ΛCDM is known as dark energy. Although other hypotheses exist, dark energy is generally understood as a "cosmological constant" often represented with a Λ. This implies that it is a scale-independent vacuum energy which homogeneously permeates all of space-time which can be interpreted as negative pressure in the field equations of GR. Lastly, there is the radiation or "hot matter" component of ΛCDM which is mostly made up of to two distinct types of particles: electromagnetic radiation and relativistic matter particles,

such as neutrinos at early times.

All these types of matter and energy which make up the Universe have distinct properties that influence the growth of large-scale structure in the Universe as is discussed in Section 1.2.

1.1.1 General Relativity

As an extension from special relativity (Einstein, 1905) to accelerating frames of reference, Einstein's theory of general relativity (Einstein, 1915; Hilbert, 1915) has been accepted as the prevailing theory of gravity. It is capable of accurately describing the orbits of planets (Clemence, 1947; Biswas & Mani, 2008), the gravitational lensing of light (Dyson et al., 1920), the propagation of gravitational waves (Abbott et al., 2016), the existence of black holes (from a direct image: Event Horizon Telescope Collaboration et al. 2019; from black hole mergers: Abbott et al. 2016; from the motion of stars: Eckart & Genzel 1996; from active galactic nuclei: e.g. Harms et al. 1994; from invisible companions: e.g. Rivinius et al. 2020; from X-ray bursts: e.g. Kaaret et al. 2017), etc., while also reproducing all predictions from special relativity and Newtonian mechanics.

The general theory of relativity is based on some key axioms. First, there is the Principle of the Equivalence of Gravitation and Inertia, also known as the Equivalence Principle. According to Weinberg (2008), the Equivalence Principle states that "at any space-time point in an arbitrary gravitational field there is a locally inertial coordinate system in which the effects of gravitation are absent in a sufficiently small space-time neighborhood of that point" (p. 511). Here, "locally inertial" refers to the fact that for infinitesimal intervals of space-time, a Minkowski metric applies, i.e. the laws of motion as given by special relativity apply (Einstein, 1905). This property is directly akin to an n-dimensional manifold being homeomorphic to an $\mathbb{R}^n$-space. Therefore, according to the Equivalence Principle, space-time must be described by a manifold (see Weinberg 1972 for a review of the properties of such objects). Secondly, we must assume a form for the gravitational potential to fix the constants within GR. Following the approach in Weinberg (1972), this can be done by assuming the Poisson equation

$$\nabla^2\Phi = 4\pi G\rho, \tag{1.1}$$

where ∇ is the Laplacian, Φ is the gravitational potential, G is Newton's gravitational constant and ρ is the mass density in the non-relativistic limit. Considering a weak static field, the time-component of the underlying metric of the four-dimensional space-time manifold, $g_{\mu\nu}$, is given by

$$g_{00} = 1 + 2\Phi, \tag{1.2}$$

where here and hence forth $c = 1$ and the metric is assumed to follow the convention which scales the time-like coordinate by a factor of $+1$, while scaling the three space-like coordinates by a factor of -1 (+, -, -, -). Taking into consideration that in the non-relativistic limit the time-time-component of the stress-energy-tensor, $T_{\mu\nu}$, is $T_{00} \approx \rho$, the Poisson equation becomes

$$\nabla^2 g_{00} = 8\pi G T_{00}. \tag{1.3}$$

Taking this equation as an ansatz for the general relation of $T_{\mu\nu}$, renaming the left-hand-side as a tensor and invoking the Equivalence Principle to extrapolate this relation beyond the weak field limit, one finds that

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{1.4}$$

where $G_{\mu\nu}$ is the Einstein tensor and $T_{\mu\nu}$ is the stress-energy tensor which for a prefect fluid in equilibrium is given as $T_{\mu\nu} = \mathrm{diag}[\rho, P, P, P]$ where P is the pressure of the fluid. In general, one can construct a field consistent with $G_{\mu\nu}$ from a Riemann curvature tensor, $R^{\nu}_{\mu\alpha\beta}$ for some covector in space-time, W_μ, which is given by

$$R^{\nu}_{\mu\alpha\beta} W_\nu \equiv [\nabla_\alpha, \nabla_\beta] W_\mu, \tag{1.5}$$

where ∇_α is the covariant derivative and $[\nabla_\alpha, \nabla_\beta]$ indicates the commutator of the covariant derivative with itself. Here is where the third axiom of GR is involved:

the "vanishing torsion assumption", i.e. that parallel-transported vectors are not rotated/do not experience torsion. Mathematically, this implies that $[\nabla_\alpha, \nabla_\beta] = \nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha$. Consequently, the Christoffel symbol, $\Gamma^\alpha_{\mu\nu}$, appearing within the covariant derivative, which is given by $\nabla_\mu W_\nu = \partial_\alpha W_\nu - \Gamma^\alpha_{\mu\nu} W_\alpha$ with $\partial_\alpha \equiv \partial/\partial x^\alpha$ for a given four-vector x^α, must be a symmetric tensor of rank 3 given by

$$\Gamma^\alpha_{\mu\nu} = \frac{1}{2} g^{\alpha\beta} \left(\partial_\nu g_{\beta\mu} + \partial_\mu g_{\beta\nu} - \partial_\beta g_{\mu\nu}\right). \tag{1.6}$$

Therefore, the Riemann curvature tensor, $R^\nu_{\mu\alpha\beta}$, is anti-symmetric, and one can contract it as $R^\alpha_{\mu\alpha\nu} \equiv R_{\mu\nu}$ to define the Ricci tensor, and in turn contract $R^\mu_\mu \equiv R$ to define the Ricci scalar. Ensuring that $G_{\mu\nu}$ is consistent with the anti-symmetric nature of $R^\nu_{\mu\alpha\beta}$, taking into consideration that $G_{\mu\nu} = T_{\mu\nu} = 0$ and $G_{00} = \nabla^2 g_{00}$ in the weak static field limit, one can obtain the Einstein field equations as follows

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu}. \tag{1.7}$$

However, Equation (1.7) still misses one axiom commonly included when writing the Einstein field equations for the whole Universe: the existence of a vacuum energy which counteracts the effect of gravity, i.e. dark energy. This can be achieved by including a term with a "cosmological constant", Λ, as follows

$$G_{\mu\nu} + g_{\mu\nu}\Lambda = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + g_{\mu\nu}\Lambda = 8\pi G T_{\mu\nu}. \tag{1.8}$$

Although this cosmological constant is nowadays used to model the vacuum energy which drives the accelerated expansion of a non-static Universe (Lemaître, 1927; Riess et al., 1998; Perlmutter et al., 1999), it was originally proposed by Einstein to balance the Einstein field equations such that the Universe could be static (Einstein, 1917). In Einstein's original work, the cosmological constant is included as an additional term added to the Einstein tensor, as shown in this work. However, Equation (1.8) allows some freedom of interpretation regarding the nature of the cosmological constant. The $g_{\mu\nu}\Lambda$ term can be freely added either to the left-hand side or to the right-hand side of the Einstein field equations, as both cases satisfy

the Bianchi identities (Voss, 1880; Bianchi, 1902). With this, the cosmological constant, when kept on the left-hand side of the Einstein field equations, can be treated as a constant scalar curvature inherent to space-time which contributes to the Einstein tensor. Equivalently, when the cosmological constant is on the right-hand side of the equation, it can be interpreted as a perfect fluid which contributes a constant energy and pressure to the stress-energy tensor. Although the latter interpretation is the most commonly used, within the context of general relativity, both are mathematically equivalent.

Equation (1.8) fully captures the dynamics of the Universe and its contents within a four-dimensional torsion-free space-time with a constant vacuum energy. At the same time, it gives the freedom to select any metric which may define the curvature of space-time, while also giving the freedom of including different types of energy and matter in the stress-energy tensor.

1.1.2 Homogeneity, Isotropy and the Friedmann-Lemaître-Robertson-Walker Metric

Building upon the cosmological Einstein field equations, one can invoke the "cosmological principle": at the largest scales, the Universe is spatially homogeneous and isotropic (Robertson, 1936; Walker, 1937). As a consequence, the metric of space-time must not have a preferred spatial direction which could allow the definition of a frame of reference that violates isotropy. Allowing the space-like coordinates of the metric to vary as a function of the time-like coordinate (i.e. allowing the Universe to evolve), we may define the following metric (Friedmann, 1924; Lemaître, 1927; Robertson, 1936; Walker, 1937),

$$\mathrm{d}s^2 = g_{\mu\nu}(x)\mathrm{d}x^\mu \mathrm{d}x^\nu = \mathrm{d}t^2 - a^2(t)\mathrm{d}\boldsymbol{\Sigma}^2, \tag{1.9}$$

where $\mathrm{d}s$ is an infinitesimal interval, x^μ is a four-vector existing on the manifold defined by $g_{\mu\nu}(x)$, t is time, $a(t)$ is the scale factor at time, t, and $\boldsymbol{\Sigma}$ is a three-vector which defines a three-dimensional space with a uniform curvature. The latter can be defined in any Lorentz invariant coordinate system. For the purposes of this thesis,

two coordinate systems are of interest: hyperspherical coordinates (χ, θ, ϕ) and reduced-circumference polar coordinates ($\mathscr{X}$, θ, ϕ). The former are given by

$$\mathrm{d}\boldsymbol{\Sigma}^2 = \mathrm{d}\chi^2 + f_k^2(\chi)\left[\mathrm{d}\theta^2 + \sin^2(\theta)\mathrm{d}\phi^2\right], \tag{1.10}$$

where χ is the comoving distance, θ and ϕ are orthogonal angular coordinates, and $f_k(\chi)$ is the transverse comoving distance defined as a function of comoving distance as follows

$$f_k(\chi) = \begin{cases} k^{-\frac{1}{2}}\sin\left(\chi k^{\frac{1}{2}}\right), & \text{if } k > 0, \\ \chi, & \text{if } k = 0, \\ |k|^{-\frac{1}{2}}\sinh\left(\chi|k|^{\frac{1}{2}}\right), & \text{if } k < 0, \end{cases} \tag{1.11}$$

where k is the Gaussian curvature of space-time at $a(t) = 1$ (in most cases, the Universe is assumed to be flat, so $k = 0$). The radial coordinate, the comoving distance, χ, is defined such that it is constant between two frames of reference which are coupled to the expansion of the Universe. The transverse comoving distance, $f_k(\chi)$, is similar to the comoving distance in that it also accounts for the expansion of the Universe. In addition, it also accounts for any deviations in the path length caused by the curvature of the underlying space-time manifold, as parametrised by k.

Alternatively, it is also useful to express the FLRW metric given in Equation (1.9) in reduced-circumference polar coordinates as follows

$$\mathrm{d}\boldsymbol{\Sigma}^2 = \frac{\mathrm{d}\mathscr{X}^2}{1 - k\mathscr{X}^2} + \mathscr{X}^2\left[\mathrm{d}\theta^2 + \sin^2(\theta)\mathrm{d}\phi^2\right], \tag{1.12}$$

where the radial coordinate, $\mathscr{X}$, is equivalent to the angular comoving distance, $\mathscr{X} \equiv f_k(\chi)$. Note that Equations (1.10) and (1.12) are equivalent in their representation of the FLRW metric. The distinguishing quality between the two metrics is the choice of radial coordinate: the comoving distance, χ, for hyperspherical coordinates, and the angular comoving distance, $\mathscr{X}$, for reduced-circumference polar coordinates.

Regardless of how we define the set of coordinates, when substituting the FLRW metric in Equation (1.9) into the cosmological Einstein field equations in Equation (1.8), we obtain the following set of independent equations known as the Friedmann equations (Friedmann, 1924),

$$\left[\frac{\dot{a}(t)}{a(t)}\right]^2 + \frac{k}{a^2} - \frac{\Lambda}{3} = \frac{8\pi G}{3}\rho, \tag{1.13}$$

$$\frac{\ddot{a}(t)}{a(t)} - \frac{\Lambda}{3} = -\frac{4\pi G}{3}\big(\rho + 3P\big), \tag{1.14}$$

where $\dot{a} \equiv \mathrm{d}a/\mathrm{d}t$. As $a(t)$ gives the relative scaling of distances at a given time, $\dot{a}/a$ from Equation (1.13) is proportional of the speed at which the Universe expands, while $\ddot{a}/a$ from Equation (1.14) is proportional to the acceleration of the expansion speed of the Universe. Since these quantities are often used in cosmological calculations, it is useful to fold them into the Hubble parameter, $H(t)$, as follows (Lemaître, 1927),

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)} = H_0 E(t), \tag{1.15}$$

where H_0 is the Hubble constant, i.e. the Hubble parameter today, and $E(t)$ is the dimensionless Hubble parameter as a function of time. H_0 is one of the main independent parameters used when constraining ΛCDM from observations as many cosmological probes are sensitive to the expansion history of the Universe. Due to its omnipresence, in the literature, it is often expressed using the so-called little h, which is defined such that $H_0 = 100h\,\mathrm{km\,s^{-1}Mpc^{-1}}$. It is also of historical significance, as the first time it was measured (Hubble, 1929) by observing a linear relation between the luminosity distance, D_L, of nearby galaxies and their associated recession speed, v, such that $H_0 = \mathrm{v}/D_\mathrm{L} > 0$; it disproved the notion of a static Universe.

This measurement by Hubble (1929) made use of an important effect to measure the luminosity distances, D_L, which is at the core of all cosmological measurements in modern cosmology: cosmological redshift. Redshift is a phenomenon

which occurs when the wavelength of light is stretched (i.e. shifted towards the red end of the electromagnetic spectrum), and it is defined as follows

$$z = \frac{\lambda(t_{\rm obs}) - \lambda(t)}{\lambda(t)}, \tag{1.16}$$

where $\lambda(t_{\rm obs})$ is the wavelength of light as seen by an observer, while $\lambda(t)$ is the wavelength of the light at the time of its emission by the source, t. Redshift can occur for multiple reasons: the light source might have a peculiar velocity with respect to the observer causing a redshift due to the Doppler effect, or the light source might be in a gravitational potential, so there is a gravitational redshift due to the light travelling through a curved space-time. Besides such cases, the type of redshift of particular interest to the work presented in this thesis is the redshift caused by the cosmological expansion of the Universe. According to the FLRW metric, the wavelengths given in Equation (1.16) scale linearly with the scale factor, $a(t)$, such that

$$z = \frac{a(t_{\rm obs})}{a(t)} - 1, \tag{1.17}$$

where $a(t_{\rm obs})$ is taken to be equal to one when the observation is made at the present day. Consequently, any instance of the scale factor, $a(t)$, can by replaced by $1/(1+z)$, so, for example, the luminosity distance can be expressed as $D_{\rm L} \equiv a(t) f_k[\chi(t)] = f_k[\chi(z)]/(1+z)$. In other words, measuring cosmological redshift allows one to directly probe the scale factor at a given time, so the expansion history of the Universe can be deciphered.

1.1.3 Ingredients of the Universe

As shown in Sections 1.1.1 and 1.1.2, the ideas of GR and the FLRW metric allow one to fully describe the expansion history of space-time. As is apparent from the dependence of the Friedmann equations (1.13) and (1.14) on the density, ρ, and pressure, P, of the fluids within the Universe, the dynamics of a given component ingredient of the Universe will depend on its equation of state given by

$$w \equiv \frac{P}{\rho}, \tag{1.18}$$

where w is a dimensionless ratio. To define the dynamics, we consider only the $T^{0\nu}$ momentum terms of the stress-energy tensor, $T^{\mu\nu} = \mathrm{diag}[\rho(t), P(t), P(t), P(t)]$, and invoke conservation of momentum, $\nabla_\nu T^{0\nu} = 0$, which gives that

$$\dot{\rho}(t) + \frac{3\dot{a}(t)}{a(t)}\big(\rho(t) + P(t)\big) = 0, \tag{1.19}$$

which is known as the continuity equation. When solving this differential equation, one finds that the density of the fluid, ρ, in general evolves according to

$$\rho(t) = \rho_0\, a(t)^{-3(1+w)}, \tag{1.20}$$

where ρ_0 is the initial density. It is often convenient to express these densities as multiples of the critical density, ρ_{crit}, which can be thought of as the total energy density contributed by the contents of a flat Universe (i.e. $k = 0$) as follows

$$\rho_{\mathrm{crit}}(t) \equiv \sum_i \rho_i(t), \tag{1.21}$$

where $\rho_i(t)$ is the energy density of any constituent form of matter, radiation, etc., which we choose to include in our model of the Universe, but we exclude any contribution from curvature ($i \neq \kappa$). For a flat Universe, the critical density is then given by

$$\rho_{\mathrm{crit}}(t) = \frac{3H^2(t)}{8\pi G}. \tag{1.22}$$

Thence, we may define a density parameter, Ω, as follows

$$\Omega(t) \equiv \frac{\rho(t)}{\rho_{\mathrm{crit}}(t)}. \tag{1.23}$$

With this, it is possible to fully characterise the properties and dynamics of a given ingredient within a FRLW space-time through its w and Ω.

1.1.3.1 Baryonic Matter

In cosmology, "baryonic" matter refers to all massive non-relativistic particles made up of quarks and charged leptons as described by the Standard Model of Particle Physics. This definition includes charged leptons which would strictly not be considered constituents of baryons in the field of particle physics. However, in cosmology, "baryons" are essentially all massive particles which couple to the electromagnetic force, so they can be observed.

With their mass, baryons possess a lot of inertia compared to their average kinetic energy such that they would not spontaneously diffuse, wherefore a baryonic fluid with a density parameter, $\Omega_{\rm b}$, would not exert any pressure on its surroundings and its $w = 0$. Given Equation (1.20), $\rho \propto a^{-3}$, so that if $\Omega_{\rm b}$ dominates the energy density of the Universe, the homogeneous evolution of the Universe is given by

$$E^2(t) \approx \Omega_{\rm b,0} a(t)^{-3} = \Omega_{b,0}[1+z(t)]^3. \tag{1.24}$$

where $\Omega_{\rm b,0}$ is the baryonic density parameter at $t = 0$.

Beyond its effect on the expansion history of the Universe, baryonic matter also impacts the formation, growth and evolution of structure. Firstly, it makes the Universe clumpier and denser by contributing $4.9 \pm 0.1\%$ to the energy density of the Universe and approximately 19% of the matter in the Universe (Planck Collaboration et al., 2020). Secondly, baryonic matter self-interacts, so at small scales, where electromagnetic forces become relevant, structure formation is limited by the pressure exerted from self-interactions (see Section 1.2.4 for a more detailed discussion of this).

1.1.3.2 Cold Dark Matter

Cold dark matter refers to any kind of non-relativistic matter which does not self-interact, but which can be subject to gravitational interactions. Despite many different hypothesis existing for the fundamental nature of dark matter (see Chapter 27 of Particle Data Group et al. 2020 for a comprehensive review), at this time, there has not been a direct detection of any dark matter candidate particle. Nev-

ertheless, there has been convincing observational evidence for its existence from many probes. Firstly, it was originally hypothesised in order to account for the fact that the visible/baryonic matter seen in clusters of galaxies is not enough to explain the galaxies' velocity dispersion (Zwicky, 1933), while also being able to account for the lack of observed baryonic matter at the edge of galaxies in order to explain the observed rotational curves of the galaxies as a function of the distance from the centre (Rubin & Ford, 1970). Since then all modern cosmological measurements have been consistent with approximately one quarter of the energy density of the Universe being attributed to dark matter (Mather et al., 1990; Riess et al., 1998; Perlmutter et al., 1999; Kaiser et al., 2000; Wittman et al., 2000; Van Waerbeke et al., 2000; Bacon et al., 2000; Eisenstein et al., 2005; Blake et al., 2011; Bennett et al., 2013; Heymans et al., 2013; Hildebrandt et al., 2017; Planck Collaboration et al., 2020; Asgari et al., 2021; Amon et al., 2022; Li et al., 2023). Additionally, measurements of strong gravitational lensing around the Bullet Cluster (also known as 1E 0657-56) have shown with a 8σ significance level that some form of dark matter is always required to explain the observed gravitational lensing signal, even when considering alternate theories for gravity beyond GR (Clowe et al., 2006). This is because the Bullet Cluster is observed to have large spatial offset between the centre of total mass and the centre of baryonic mass which cannot be explained without any dark matter contributing to the mass of the cluster.

Although the existence of dark matter is nowadays mostly undisputed within the field of cosmology, its nature is not as certain. Within ΛCDM (Peebles, 1982; Blumenthal et al., 1984), it is assumed that dark matter is cold, i.e. that it is non-relativistic and that it does not self-interact. Other types of dark matter have been proposed which fit current observations at large-scales with a similar level of accuracy as cold dark matter, but may deviate in their predictions for observations of smaller-scale structure. Examples of such models are warm dark matter (a form of dark matter which behaves relativistically at small scales and non-relativistically at large scales; see Viel et al. 2013 for more details) and fuzzy dark matter (dark matter made up of light bosons which would Compton scatter at small scales, but

have a small probability of self-interaction at large scales; see Hu et al. 2000 for details). Having said this, for now, it has not been possible to make precise enough measurements of structure at small scales to significantly discern between cold dark matter and any alternative hypothesis.

Similarly to baryonic matter, cold dark matter does not diffuse and exert external pressure. Hence, we may add another form of matter to the stress-energy tensor with a density parameter, Ω_c, and $w = 0$, so that $\rho \propto a^{-3}$ if the Universe is dominated by dark matter such that

$$E^2(t) \approx \Omega_{\mathrm{c},0} a(t)^{-3} = \Omega_{\mathrm{c},0}[1 + z(t)]^3, \tag{1.25}$$

where $\Omega_{\mathrm{c},0}$ is the dark matter density parameter at $t = 0$. As cold dark matter's density scales in the same way with the expansion of the Universe as baryonic matter, their densities are often grouped together to define the matter density parameter of the Universe, Ω_m, as follows

$$\Omega_{\mathrm{m}}(t) \equiv \Omega_{\mathrm{c}}(t) + \Omega_{\mathrm{b}}(t). \tag{1.26}$$

Knowing that cold dark matter makes up $25.89 \pm 0.57\%$ of the energy density of the Universe and approximately 91.23% of all matter in the Universe (Planck Collaboration et al., 2020), from Equation (1.14), it becomes apparent that dark matter is the driving factor in counter-acting the acceleration of the expansion of the Universe. Simultaneously, its lack of self-interaction means that dark matter structure could grow from the early Universe without scattering. All this allows us to conclude that the signal of cosmological large-scale structure (see Section 1.2) is mostly determined by the physics of dark matter.

1.1.3.3 Dark Energy

In Section 1.1.1, we have already introduced the concept of dark energy as one of the assumptions to derive the cosmological Einstein equations of GR given by Equation (1.8). In this equation, dark energy is represented by a cosmological constant, Λ, which is assumed to be the same at all points in space and at all times. From the

Friedmann equations (1.13) and (1.14), we can see that Λ, if positive, drives the acceleration of the scale factor, $a(t)$. From Equation (1.14), it becomes apparent that the term with $\ddot{a}(t) \propto \Lambda/3$ could be folded into the stress-energy tensor by assuming that dark energy is a fluid with $w = -1$. Consequently, evaluating Equation (1.19), we trivially find that $\rho(t) = \rho_0$ for dark energy, i.e. the density of dark energy is constant over time. Analogously to matter, we may therefore also define a density parameter for dark energy, Ω_Λ, as follows

$$\Omega_\Lambda \equiv \frac{\Lambda}{3H_0^2}. \tag{1.27}$$

From that, we can model the dimensionless Hubble parameter in a dark energy dominated Universe as follows

$$E^2(t) \approx \Omega_\Lambda. \tag{1.28}$$

It is worth noting that extensions to ΛCDM exist which allow the w of dark energy to vary with time, called wCDM. These are equivalent to different modifications of Einsteinian gravity and they are usually expressed in a parametric form as follows $w(t) = w_0 + [1 - a(t)]w_a$. However, these theories are not investigated in this thesis and I will generally assume that $w = -1$. This assumption has proven to be consistent with observations which found that $w = -1.03 \pm 0.03$ (Planck Collaboration et al., 2020). Thus, it is safe to consider dark energy as a cosmological constant with a constant density equivalent to approximately 70% of the energy density of the Universe (Riess et al., 1998; Perlmutter et al., 1999; Planck Collaboration et al., 2020).

1.1.3.4 Radiation

All relativistic components in the Universe are typically classed as radiation. From the Standard Model of Particle Physics, the only fundamental particles which are typically relativistic are the bosons and the leptons without any electromagnetic charge, i.e. the neutrinos. Out of the bosons, the only one which has a large enough mean free path to be cosmologically relevant is the photon. Therefore, radiation in

cosmology can be thought of as only being constituted by photons and neutrinos.

Additionally, most of the radiation in the Universe can be attributed to background radiation, i.e. a homogeneous and isotropic emission from the early Universe during which the emission rate of radiation was extremely large when compared to today.

After the Big Bang (see 1.1.4), photons were produced everywhere in the early Universe at high rates during baryogenesis (i.e. the creation of baryonic matter such that the production of any anti-matter was outpaced) and big bang nucleosynthesis (i.e. the condensation of gluons and quarks into nucleons, such as protons and neutrons, as the Universe cooled). However, the photons' mean free path length remained shorter than a few kpc as long as the Universe was a hot plasma in which all emitted photons were quickly absorbed or scattered. Once the Universe had expanded more and cooled down, the electromagnetic binding energies were finally sufficient to allow for recombination to occur at $z \approx 1,100$ (Weinberg, 2008). During recombination, the newly formed nuclei became neutral atoms which released black body radiation until reaching their ground-state. Thanks to this, the Universe was no longer opaque and filled with hot plasma, but instead it became transparent and electrically neutral. From this point on, photons were no longer in thermal equilibrium with the rest of the Universe, i.e. photon decoupling had occurred. As a result of this, the entire Universe is permeated with photons that were only in thermal contact with matter at the time of recombination. Since these emissions have experienced more than 13 Gyrs of cosmological redshift, their central wavelength is now in the microwave regime, so it is known as the Cosmic Microwave Background (CMB). Observing the CMB thus allows one to directly probe the temperature of the Universe at recombination and any anisotropies in the emission due to primordial density fluctuations (see Section 1.2.1).

Within the first second after the Big Bang, the Universe mostly consisted of electrons, positrons, nucleons, photons and neutrinos. Electron-positron interactions through the weak nuclear force were also coupled to neutrinos which meant that they were all in thermal equilibrium. As the Universe expanded, the rate of

electron-positron interactions slowed down considerably, so neutrinos thermally decoupled from the rest of the Universe. This event created the homogeneous and isotropic Cosmic Neutrino Background (CNB; see Lesgourgues et al. 2013 for a detailed review) which is analogous to the CMB.

To define the equation of state of each type of radiation, it is useful to formalise how the energy density due to radiation may be incorporated into the stress-energy tensor. Since we know that the vast majority of radiation in the Universe comes from homogeneous and isotropic radiation backgrounds, we may assume that the distribution of radiation in phase space only depends on energy, momentum and time. For particles with a momentum, $P^\mu = (E, p^i)$ where $p^i \in \mathbb{R}^3$, at time t, the stress-energy tensor is given by (Lesgourgues et al., 2013),

$$T^\mu_\nu(t) = g \int \frac{d^3p}{(2\pi)^2} \frac{P^\mu P_\nu}{E} f(p,t), \tag{1.29}$$

where g is the number of internal degrees of freedom of the particle, $p^i_i = p$ and $f(p,t)$ is the statistical distribution of particles in phase-space. The diagonal components of $T^\mu_\nu(t)$ give

$$T^0_0(t) = \rho = g \int \frac{d^3p}{(2\pi)^2} E f(p,t), \tag{1.30}$$

$$T^i_i(t) = P(t) = -g \int \frac{d^3p}{(2\pi)^2} \frac{p^2}{3E} f(p,t). \tag{1.31}$$

Since photons are massless and neutrinos have very small masses, we may take $E = p$ in both cases (Einstein, 1905). Combining Equations (1.30) and (1.31) gives that $P = \rho/3$, so that $\rho \propto a^{-4}$ in a radiation-dominated Universe. In such a case

$$E^2(t) \approx \Omega_{\mathrm{R},0} a(t)^{-4} = \Omega_{\mathrm{R},0}[1 + z(t)]^4, \tag{1.32}$$

where $\Omega_{\mathrm{R},0}$ is the density parameter for all types of radiation at $t = 0$, and it is given by

$$\Omega_{\mathrm{R}}(t) = \Omega_\gamma(t) + \Omega_\nu(t), \tag{1.33}$$

where Ω_γ and Ω_ν are the density parameters for photons and neutrinos, respectively.

To find the equation of state of neutrinos, we may take Equations (1.30) and (1.31) and consider two cases: one where the particles in question are bosons (i.e. photons, γ) which follow Bose-Einstein statistics and another where we consider fermions (i.e. neutrinos, ν) which follow Fermi-Dirac statistics. In each case, $f(p,t)$ takes the form of the joint Bose-Einstein/Fermi-Dirac distribution of the corresponding particle-antiparticle pair. Assuming that the particles in question are relativistic, one can show that

$$\rho_i = 3P_i = \begin{cases} \frac{\pi^2}{30} g_\gamma T_\gamma^4, & \text{if } i = \gamma, \\ \frac{7}{8}\frac{\pi^2}{30} g_\nu T_\nu^4, & \text{if } i = \nu, \end{cases} \tag{1.34}$$

where T_γ and T_ν are the temperatures at thermal equilibrium of the photon and neutrino backgrounds, respectively. Given this, $\frac{g_\gamma}{g_\nu} = \frac{7}{8}(T_\gamma/T_\nu)^{-4}$. Considering that the specific entropy per comoving volume, s_i, is given by $s_i = \partial P_i/\partial T_i$, then $s_i \propto T_i^{-3}$. We also know that s_i must be conserved which can be used to show that at decoupling $T_\nu/T_\gamma = (4/11)^{1/3}$. Relating this back to Equation (1.34), we can express the density of neutrinos as a function of the photon density as follows

$$\Omega_\nu(t) = N_{\text{eff}}\,\Omega_\gamma(t)\,\frac{7}{8}\left(\frac{4}{11}\right)^{\frac{4}{3}}, \tag{1.35}$$

where N_{eff} is the effective number of neutrino species. N_{eff} is included as an additional parameter for two reasons. Firstly, historically it was not known with certainty how many active light neutrino generations there are. Nowadays, thanks to electroweak measurements from Z-boson decays (ALEPH Collaboration et al., 2006), we know that there are three different generations of neutrinos. Secondly, despite there being three species of neutrinos, it is usually assumed that $N_{\text{eff}} = 3.044$ due to higher-order corrections from neutrino mixing and due to the neutrino decoupling not being instantaneous (de Salas & Pastor, 2016; Akita & Yamaguchi, 2020).

1.1.4 Expansion History and the Big Bang

With all the ingredients of the Universe described in Section 1.1.3, we can combine their impacts on the energy content and the background expansion of the Universe to rewrite the Friedmann equations in Equations (1.13) and (1.14) as follows

$$\frac{H(t)^2}{H_0^2} = (\Omega_{\mathrm{c},0} + \Omega_{\mathrm{b},0})a(t)^{-3} + (\Omega_{\gamma,0} + \Omega_{\nu,0})a(t)^{-4} + \Omega_{k,0}a(t)^{-2} + \Omega_\Lambda, \quad (1.36)$$

where, for completeness, the density parameter due to the curvature in the FLRW metric at $t = 0$, $\Omega_{k,0} \equiv -k/H_0^2$, is also included[1]. Figure 1.1 shows how each of the terms in Equation (1.36) scales over the entire cosmic history of the Universe to define the background expansion of the Universe.

When inspecting Figure 1.1, the Friedmann equations and the FLRW metric, one may realise that the scale factor, $a(t)$, asymptotically approaches $a = 0$ as $t \to\sim 1/H_0$. Combined with the fact that $\dot{a}(t) > 0$, this means that all space-like coordinates within a FLRW metric tend to approach each other until they create a singularity at $a(t) = 0$. This singularity, known as the Big Bang, can be thought of as the origin of the time-like coordinate in the FLRW metric. Approaching this singularity going back in time, $\rho \to \infty$ and $T \to \infty$. Thus, all matter in the Universe was at some point in thermal contact.

1.2 Growth of Large-Scale Structure

The discussion of The Standard Model of Cosmology in Section 1.1 describes the background expansion of the Universe as a perfect and homogeneous fluid described by ΛCDM. However, as we know from observations (and from our own anthropological existence), the Universe is not perfectly homogeneous. In fact, it has structure even at large scales. These deviations from homogeneity can exist due to the quantum mechanical nature of matter producing random fluctuations in the density of the early Universe. Section 1.2.1 discusses the origins of these primordial mat-

[1]Note that from this point forward, I choose the shorthand notation for the density parameters which drops the zero subscript, so unless time-dependence is explicitly stated, any Ω is given at $t = 0$.

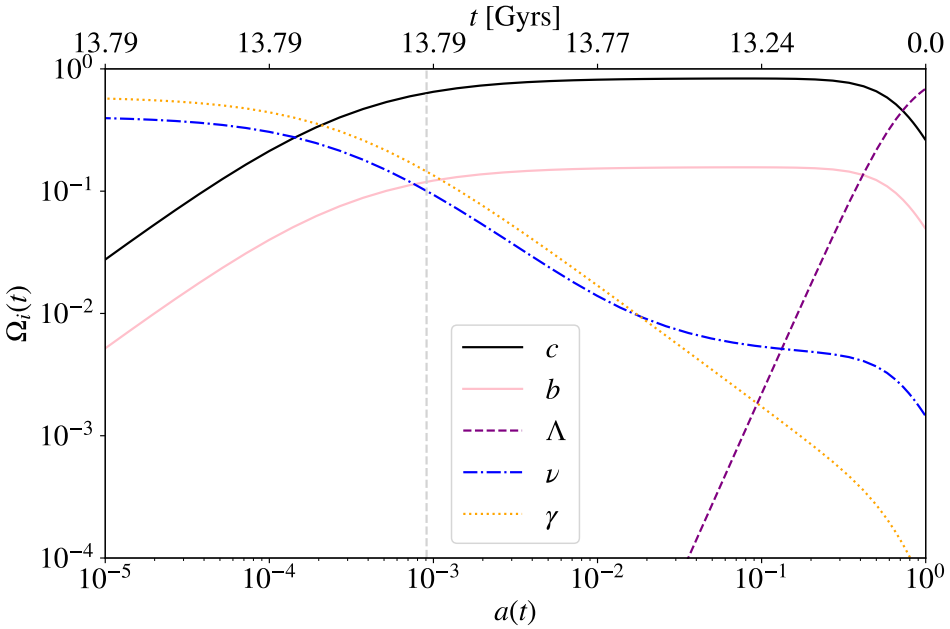

Figure 1.1: History of the density parameters, $\Omega_i(t)$, of all of the ingredients of ΛCDM discussed in Section 1.1.3 as a function of look-back time, t, in Gyrs and the associated scale factor, $a(t)$. The solid black line shows the density parameter for cold dark matter, Ω_c, the pink solid line shows the density parameter for baryonic matter, Ω_b, the purple dashed line shows the density parameter for dark energy, Ω_Λ, the blue dot-dashed line shows the density parameter of neutrinos, Ω_ν, and the orange dotted line shows the density parameter of photons/light, Ω_γ. The vertical light gray dashed line indicates recombination at $z \approx 1,100$. Figure made by the author in accordance with Planck Collaboration et al. (2020).

ter fluctuations, how they grew to cosmological scales with inflation (Guth, 1981; Linde, 1982), and shows how they are typically modelled within ΛCDM. In Section 1.2.3, I show how the primordial fluctuations grow with time within the linear regime. To complete this, in Section 1.2.4, I discuss how at small scales baryonic effects cause deviations from linear growth of structure, and how such deviations may be modelled.

These variations in structure are then traced by the galaxies and their dark matter halos which form in the late Universe, allowing us to use galaxy observations as probes of the underlying large-scale structure (see Section 1.3).

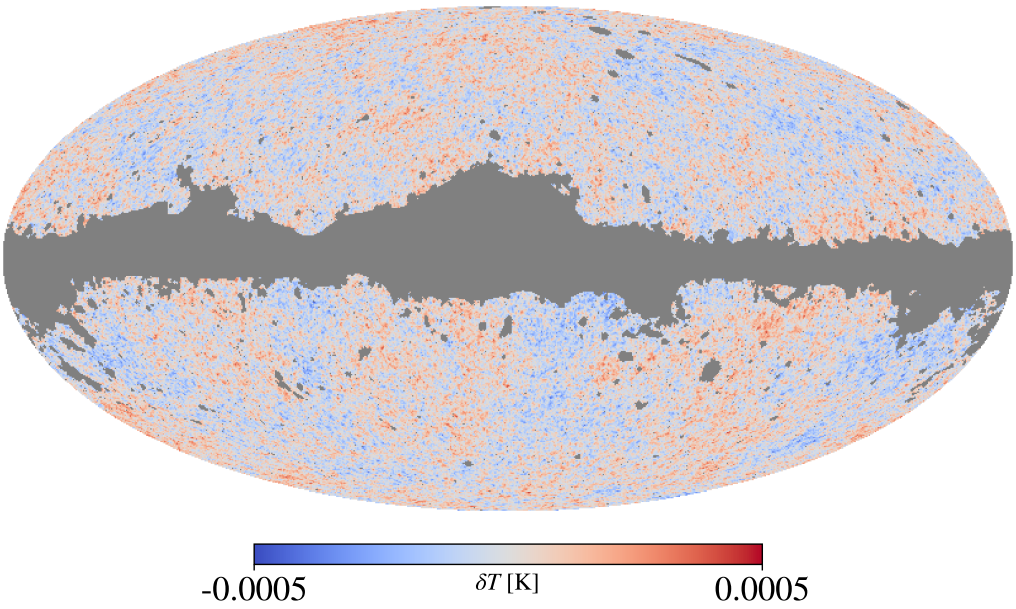

Figure 1.2: Map of the temperature anisotropies, δT, observed in the cosmic microwave background by the ESA Planck space observatory (Planck Collaboration et al., 2020). The grey areas indicate pixels which mask the galactic foreground. Figure made by the author with the Planck Collaboration et al. (2020) data.

1.2.1 Primordial Matter Fluctuations

After the Big Bang, as $a(t)$ in the FLRW metric (see Section 1.1.2) increases, the Universe expands, cools down and matter particles begin to appear, while radiation is emitted as discussed in Section 1.1.3.4. Since the creation, annihilation, emission and scattering of the fundamental particles in the early Universe are inherently stochastic processes governed by quantum mechanics, it is at this stage that random fluctuations in the underlying density of the Universe appear. The overdense regions of these fluctuations become seeds of attractors which first aggregate dark matter and which, after recombination, aggregate baryonic matter too.

At first, these fluctuations are at subatomic scales which can grow in scale as the Universe expands. However, when taking into consideration observations of large-scale structure such as the CMB shown in Figure 1.2, the expansion rate as predicted by standard ΛCDM is not enough to explain the scales of the fluctuations at recombination. Some additional problems also arise with this picture of the seeds for large-scale structure growth. From Figure 1.2, one can infer that the Universe is

remarkably homogeneous at recombination with a root-mean square of the temperature fluctuations deviating by only 0.1% from the mean temperature of $T \approx 2.725$ K (Planck Collaboration et al., 2020). However, this seems highly unlikely to be the case by the time recombination occurs at $z \approx 1,100$ when only considering the expansion history of the Universe as shown in Figure 1.1. To quantify this, we can define the particle horizon, $d_{\text{horizon}}(z)$, as the maximum angular diameter distance which a particle travelling at the speed of light could have travelled at a given age of the Universe, $t_{\text{age}}(z)$ as follows

$$d_{\text{horizon}}(z) \equiv a(z) \int_0^{t_{\text{age}}(z)} \frac{\mathrm{d}t}{a(t)} = \frac{1}{1+z} \int_z^{\infty} \frac{\mathrm{d}z}{H(z)}. \tag{1.37}$$

Since the Universe is matter-dominated during recombination, we may assume Equation (1.25) applies, such that

$$d_{\text{horizon}}(z) = \frac{2}{\sqrt{\Omega_{\text{m},0}} H_0 (1+z)} \left[1 - (1-z)^{-1/2}\right]. \tag{1.38}$$

Thus, at recombination, $d_{\text{horizon}}(z = 1,100) \approx 2$ degrees. This implies that a given point in the sky at the time of recombination only had time to be in causal contact with the matter within two degrees around it. In other words, it is highly unlikely for the entire cosmic microwave background to be approximately in thermal equilibrium as observed if the expansion history implied by ΛCDM is correct. This is known as the horizon problem and it inspired the proposition of one more ingredient to the Universe: an inflaton field (Guth, 1981; Linde, 1982). Inflation theory proposes a scalar field with negative pressure, akin to dark energy, which drives a brief period of exponentially accelerated expansion of the Universe between 10^{-36} and 10^{-32} seconds after the Big Bang. With that, the inflaton field greatly expands the particle horizon at recombination to be larger than the observable Universe, as is observed in the CMB. Simultaneously, it also solves the so-called flatness problem, which is concerned with the apparent fine-tuning needed in ΛCDM to ensure that the observed Gaussian curvature in the FLRW metric, $k \approx 0$. Inflation solves this too by expanding space-time rapidly to such an extent that any non-zero curvature

vanishes to near zero even at the scales of the observable Universe. This hypothetical scalar field is then thought to decay away into fundamental Standard Model particles and contribute to the radiation density of the Universe.

Another consequence of this rapid expansion due to inflation is that it predicts that the primordial curvature fluctuations of the Universe are distributed as a scalar near-Gaussian random field of curvature perturbations in three-dimensional space, $\mathscr{R}$. This random field is given by (Liddle & Lyth, 2000),

$$\langle \mathscr{R}(\boldsymbol{k})\mathscr{R}(\boldsymbol{k}')^* \rangle \equiv 8\pi^3 \delta(\boldsymbol{k}-\boldsymbol{k}')P_{\mathscr{R}}(|\boldsymbol{k}|), \tag{1.39}$$

where $\langle \cdot \rangle$ indicates the average over all k and k', $\boldsymbol{k}$ is the three-dimensional wavevector, $P_{\mathscr{R}}(|\boldsymbol{k}|)$ is the power spectrum of the primordial curvature fluctuations, which is typically parameterised with a power law as follows

$$P_{\mathscr{R}}(|\boldsymbol{k}|) = \frac{2\pi^2}{|\boldsymbol{k}|^3} A_{\rm s} \left(\frac{|\boldsymbol{k}|}{k_{\rm pivot}} \right)^{n_{\rm s}-1}, \tag{1.40}$$

where $A_{\rm s}$ is the primordial amplitude which correlates with the energy scale a which inflation is onset, $n_{\rm s}$ is the scalar spectral index which is a function of the duration of inflation, and $k_{\rm pivot}$ is a reference pivot wavenumber. As $n_{\rm s}$ is measured to be close to one, i.e. $n_{\rm s} = 0.965 \pm 0.004$ (Planck Collaboration et al., 2020), the power spectrum of the primordial curvature fluctuations is considered "quasi scale-invariant".

1.2.2 Perturbation Theory

The fluctuations in density in the early Universe discussed in Section 1.2.1 and their growth over time is described through the modelling of localised small perturbations from the mean of the density field in space. This can be done through expanding the Einstein field equations for the FLRW metric for small density perturbations (Lifshitz, 1946; Lifshitz & Khalatnikov, 1963) which has led to the development of gauge-invariant covariant perturbation theory (Bardeen, 1980; Kodama & Sasaki, 1984; Ellis & Bruni, 1989; Ellis et al., 1989; Mukhanov et al., 1992). In this framework, the stress-energy tensor in Equation (1.8) is perturbed as $T_{\mu\nu} \to \overline{T}_{\mu\nu} + \delta T_{\mu\nu}$. However, for the purposes of modelling observations of large scale structure made

by cosmological surveys, such a treatment is not strictly necessary and Newtonian perturbation theory is a good approximation. Newtonian perturbation theory arises out of relativistic perturbation theory when only considering the scalar perturbations to the stress-energy tensor given by $\delta\rho$ and δP. The basis of Newtonian perturbation theory is the so-called Newtonian or shear-free gauge which makes any non-Newtonian effects on the density perturbations disappear, such that the Poisson equation as given in Equation (1.1) fully describes gravitational interactions. Additionally, we also assume that the energy and matter in the Universe can be modelled as a perfect fluid such that the continuity equation given in Equation (1.19) and Euler's equation (Euler, 1757) apply. The latter is given by

$$\left.\frac{\partial \boldsymbol{u}}{\partial t}\right|_{\boldsymbol{r}} + (\boldsymbol{u}\cdot\boldsymbol{\nabla}_{\boldsymbol{r}})\boldsymbol{u} = -\frac{1}{\rho}\boldsymbol{\nabla}_{\boldsymbol{r}}P - \boldsymbol{\nabla}_{\boldsymbol{r}}\Phi, \tag{1.41}$$

where $\boldsymbol{r}$ and t represent the physical space and time coordinates and $\boldsymbol{u}$ is the velocity distribution of the perfect fluid (which is assumed to be $|\boldsymbol{u}| \ll c$).

Firstly, we transform these equations into comoving spatial coordinates, $\boldsymbol{\Sigma}$, as $\boldsymbol{r} = a(t)\boldsymbol{\Sigma}$. Therefore, Equations (1.1), (1.19) and (1.41) become

$$\nabla^2_{\boldsymbol{\Sigma}}\Phi = 4\pi G a^2 \rho, \tag{1.42}$$

$$\left.\frac{\partial \rho}{\partial t}\right|_{\boldsymbol{\Sigma}} - \frac{\dot{a}}{a}\boldsymbol{\Sigma}\cdot\boldsymbol{\nabla}_{\boldsymbol{\Sigma}}\rho + \boldsymbol{\nabla}_{\boldsymbol{\Sigma}}\cdot(\rho\boldsymbol{u}) = 0, \tag{1.43}$$

$$a\left.\frac{\partial \boldsymbol{u}}{\partial t}\right|_{\boldsymbol{\Sigma}} - \dot{a}\boldsymbol{\Sigma}\cdot\boldsymbol{\nabla}_{\boldsymbol{\Sigma}}\boldsymbol{u} + \boldsymbol{u}\cdot\boldsymbol{\nabla}_{\boldsymbol{\Sigma}}\boldsymbol{u} = -\frac{1}{\rho}\boldsymbol{\nabla}_{\boldsymbol{\Sigma}}P - \boldsymbol{\nabla}_{\boldsymbol{\Sigma}}\Phi, \tag{1.44}$$

where $\boldsymbol{u}(\boldsymbol{\Sigma},t) = \dot{a}(t)\boldsymbol{\Sigma}(t) + a(t)\boldsymbol{v}(\boldsymbol{\Sigma},t)$ and $\boldsymbol{v} \equiv \dot{\boldsymbol{\Sigma}}$ is the peculiar velocity of the fluid. Next, we apply a small perturbation in the density as follows

$$\begin{aligned}
\rho(\boldsymbol{\Sigma},t) &\to \overline{\rho}(t) + \delta\rho(\boldsymbol{\Sigma},t) \equiv [1+\delta(\boldsymbol{\Sigma},t)]\overline{\rho}(t),\\
P(\boldsymbol{\Sigma},t) &\to \overline{P}(t) + \delta P(\boldsymbol{\Sigma},t),\\
\Phi(\boldsymbol{\Sigma},t) &\to \overline{\Phi}(t) + \delta\Phi(\boldsymbol{\Sigma},t),
\end{aligned} \tag{1.45}$$

where $\delta(\boldsymbol{\Sigma},t)$ is the matter/energy density contrast from the mean density, $\overline{\rho}(t)$,

which is associated with the mean pressure, $\overline{P}(t)$, and the mean gravitational potential, $\overline{\Phi}(t)$. Upon making these substitutions and subtracting the zeroth-order terms, the perturbed fluid equations are given by

$$\nabla^2_{\boldsymbol{\Sigma}}\delta\Phi = 4\pi G a^2 \overline{\rho}\,\delta, \tag{1.46}$$

$$\dot{\delta} + \frac{1}{a}\boldsymbol{\nabla}_{\boldsymbol{\Sigma}}\cdot[(1+\delta)\boldsymbol{v}] = 0, \tag{1.47}$$

$$\dot{\boldsymbol{v}} + \frac{\dot{a}}{a}\boldsymbol{v} + \frac{1}{a}(\boldsymbol{v}\cdot\boldsymbol{\nabla}_{\boldsymbol{\Sigma}})\boldsymbol{v} = -\frac{1}{a\overline{\rho}(1+\delta)}\boldsymbol{\nabla}_{\boldsymbol{\Sigma}}\delta P - \frac{1}{a}\boldsymbol{\nabla}_{\boldsymbol{\Sigma}}\delta\Phi. \tag{1.48}$$

This system of non-linear differential equations can be combined by differentiating Equation (1.47) with respect to time and combining it with Equations (1.46) and (1.48) such that

$$\ddot{\delta} + \frac{2\dot{a}}{a}\dot{\delta} - 4\pi G\overline{\rho}\,\delta = \frac{1}{a^2}\boldsymbol{\nabla}_{\boldsymbol{\Sigma}}\cdot(\boldsymbol{v}\cdot\boldsymbol{\nabla}_{\boldsymbol{\Sigma}}\boldsymbol{v}) - \frac{1}{a^2}[a\boldsymbol{\nabla}_{\boldsymbol{\Sigma}}\cdot(\delta\boldsymbol{v})] + \frac{1}{a^2\overline{\rho}}\boldsymbol{\nabla}_{\boldsymbol{\Sigma}}\cdot\frac{\boldsymbol{\nabla}_{\boldsymbol{\Sigma}}\delta P}{1+\delta}. \tag{1.49}$$

This equation is valid for all orders of δ including non-linear orders which take into consideration the Hubble friction, but higher than second-order terms only become relevant for large values of δP.

1.2.3 Linear Structure Growth

Assuming that the matter/energy density contrast of the Universe is small, $\delta \ll 1$, and neglecting any other forces which may be at play other than fluid pressure and gravity, we may linearise Equation (1.49) such that

$$\ddot{\delta} + \frac{2\dot{a}}{a}\dot{\delta} - 4\pi G\overline{\rho}\,\delta = \frac{c_{\mathrm{s}}^2}{a^2}\nabla^2_{\boldsymbol{\Sigma}}\delta, \tag{1.50}$$

where c_{s} is the speed of sound within the perfect fluid defined as $c_{\mathrm{s}}^2 \equiv \delta P/\overline{\rho}\,\delta$. Rewriting this equation for the Fourier transform of $\delta(\boldsymbol{\Sigma},t)$, we obtain

$$\ddot{\tilde{\delta}}(\boldsymbol{k},t) + \frac{2\dot{a}(t)}{a(t)}\dot{\tilde{\delta}}(\boldsymbol{k},t) + \left[\frac{|\boldsymbol{k}|^2 c_{\mathrm{s}}^2}{a(t)^2} - 4\pi G\overline{\rho}(t)\right]\tilde{\delta}(\boldsymbol{k},t) = 0, \tag{1.51}$$

where $\boldsymbol{k}$ is the comoving wavevector in $\boldsymbol{\Sigma}$-space and $\tilde{\delta}(\boldsymbol{k},t)$ is defined as

$$\tilde{\delta}(\boldsymbol{k},t) \equiv \int \mathrm{d}^3\Sigma e^{-i\boldsymbol{k}\cdot\boldsymbol{\Sigma}}\delta(\boldsymbol{\Sigma},t). \tag{1.52}$$

If $4\pi G\overline{\rho}(t) \gg |\boldsymbol{k}|^2 c_\mathrm{s}^2/a(t)^2$, the form of Equation (1.51) allows one to separate the spatial and temporal components of $\tilde{\delta}(\boldsymbol{k},t)$, so one can define

$$\ddot{D}(t) + 2H(t)\dot{D}(t) - 4\pi G\overline{\rho}(t)D(t) = 0, \tag{1.53}$$

where $H(t)$ is again the Hubble parameter and $D(t)$ is the linear growth rate of structure. Such a separation implies that, if the initial conditions for the density contrast field are known, it is possible to get the density contrast at any time, t, to first order just by applying the linear growth rate, $D(t)$, such that $\tilde{\delta}(\boldsymbol{k},t) = D(t)\tilde{\delta}(\boldsymbol{k})$. This assumption of scale-independent linear structure growth breaks down in the small-scale regime (i.e. large $|\boldsymbol{k}|$). From Equation (1.51), we find that the limit of this regime is given by the Jeans scale, k_J, as follows (Jeans, 1902),

$$k_\mathrm{J}(t) = [4\pi G\overline{\rho}(t)]^{\frac{1}{2}}\frac{a(t)}{c_\mathrm{s}}. \tag{1.54}$$

As $|\boldsymbol{k}| \to k_\mathrm{J}$, the scale-dependent $\tilde{\delta}(\boldsymbol{k},t)$ term in Equation (1.51) becomes dominant, such that any oscillations in the matter fluid at such small scales would grow in a scale-dependent manner. Subsequently, for $|\boldsymbol{k}| > k_\mathrm{J}$, the oscillations cease growing at all. During the radiation-dominated era and assuming a baryonic matter density of $\omega_\mathrm{b} = 0.04$, it can be shown that $k_\mathrm{J} \approx (1+z)^2/(2\times 10^6\,\mathrm{Mpc})$ (Longair, 1989). k_J is minimal at matter-radiation equality, $z_\mathrm{eq} \approx 3,400$, giving $k_\mathrm{J} \sim 1\,\mathrm{Mpc}^{-1}$. This implies that, during the radiation-dominated era, structure at extragalactic scales grows in a near scale-independent manner. This applies for even smaller scales during the matter-dominated era, as the speed of sound of baryonic matter is substantially smaller than the speed of sound of a radiation dominated Universe. Assuming again that $\omega_\mathrm{b} = 0.04$, one finds that $k_\mathrm{J} \sim 10^5\,\mathrm{Mpc}^{-1}$ at recombination, $z \approx 1,100$ (Longair, 1989). Consequently, during matter domination, any density fluctuations at sub-galactic scales can grow. Hence, we can conclude that scale-invariant structure growth is a valid assumption at extragalactic scales for most of cosmic history.

Given this, as different types of energy dominate the energy density of the Universe, the linear growth rate of structure will vary accordingly. In the early Universe, during the radiation-dominated era, as discussed in Section 1.1.3.4, $P = \rho/3$ and $\rho \propto a^{-4}$, such that $a \propto t^{1/2}$. In this case, Equation (1.53) can be rewritten as

$$\ddot{D}_{\rm c}(t) + \frac{1}{t}\dot{D}_{\rm c}(t) = 0, \tag{1.55}$$

where $D_{\rm c}(t)$ is the linear structure growth with time affecting cold dark matter (this excludes baryonic matter as this equation does not take into consideration any electromagnetic interactions between the radiation and the baryonic matter). Therefore, in the radiation-dominated era, there are two possible solutions for the structure growth given by

$$D_{\rm c}(t) = D_{\rm c,0} + D'_{\rm c,0}\ln(t/t_0), \tag{1.56}$$

where $D_{\rm c,0}$, $D'_{\rm c,0}$ and t_0 are constants which depend on the initial conditions. From this we can conclude that during the radiation-dominated era before recombination any inhomogeneities in the density would either stay static or grow slowly with time.

Once (dark) matter becomes dominant in the Universe, we know from Sections 1.1.3.1 and 1.1.3.2 that $\rho \propto a^{-3}$ and therefore $a \propto t^{2/3}$. Altering Equation (1.53) accordingly, and taking matter to be pressureless ($c_{\rm s} = 0$),, we find that

$$\ddot{D}_{\rm m}(t) + \frac{3}{4t}\dot{D}_{\rm m}(t) - \frac{2}{3t^2}D_{\rm m}(t) = 0, \tag{1.57}$$

where $D_{\rm m}(t)$ is the linear structure growth rate of all matter. The matter-dominated era also allows two possible solutions for the structure growth which are given by

$$D_{\rm m}(t) = D_{\rm m,0}\left(\frac{t}{t_0}\right)^{-1} + D'_{\rm m,0}\left(\frac{t}{t_0}\right)^{2/3}. \tag{1.58}$$

Assuming that density fluctuations appear at early times, the t^{-1} mode in $D_{\rm m}(t)$ is suppressed at late times, so the realised mode is the $t^{2/3}$ term. Thus, for most of the

cosmic history of the Universe, the perturbations to the mean matter density grow as a power law with time.

In the late Universe, as the era of dark energy dominance begins, the background expansion of the Universe is largely determined by the cosmological constant discussed in Section 1.1.3.3. In this scenario, $w = -1$, so ρ is independent of the scale factor, $a(t)$, so $a \propto e^{Ht}$. This then allows one to make the following approximation of Equation (1.53),

$$\lim_{\Omega_\Lambda \to 1} \left[\ddot{D}_\mathrm{m}(t) + 2H(t)\dot{D}_\mathrm{m}(t) \right] \to 0. \tag{1.59}$$

In this limit, the equation has the two following solutions

$$D_\mathrm{m}(t) = D_{\mathrm{m},0} + D'_{\mathrm{m},0}\, e^{-2H(t)t}. \tag{1.60}$$

Consequently, dark energy suppresses the growth of linear structure as time progresses.

1.2.3.1 Linear Matter Power Spectrum

Neither the density contrast, δ, nor the structure growth rate, D, can be directly observed by cosmological surveys. Instead we observe tracers of the underlying large-scale structure (such as galaxies or the cosmic microwave background). To model this, we are interested in understanding the statistics of the fluctuations in the matter density through its three-dimensional matter power spectrum, $P_\delta(|\boldsymbol{k}|, t, t')$, defined as

$$\langle \tilde{\delta}_\mathrm{m}(\boldsymbol{k}, t) \tilde{\delta}_\mathrm{m}(\boldsymbol{k}', t') \rangle = (2\pi)^3 \delta_\mathrm{D}^{(3)}(\boldsymbol{k} + \boldsymbol{k}') P_\delta(|\boldsymbol{k}|, t, t'), \tag{1.61}$$

where $\delta_\mathrm{D}^{(3)}$ is the three-dimensional Dirac delta function.

As the initial conditions at $t \approx 0$ are set by the primordial curvature fluctuations which remain after inflation, as discussed in Section 1.1.4, the equal-time three-dimensional matter power spectrum, $P_\delta(|\boldsymbol{k}|, t)$, which gives the correlations between matter field fluctuations at a given look-back time ($t = t'$), can be modelled as

$$P_\delta(|\boldsymbol{k}|,t) = T^2(|\boldsymbol{k}|)D^2(t)P_{\mathscr{R}}(|\boldsymbol{k}|), \tag{1.62}$$

where $P_{\mathscr{R}}(|\boldsymbol{k}|)$ is the three-dimensional power spectrum of primordial curvature fluctuations as defined in Equation (1.40) and $T(|\boldsymbol{k}|)$ is the transfer function which depends on the solutions to the spatial part of Equation (1.51). To parametrise the linear matter power spectrum, it is common to use the parameter σ_8 which is defined by

$$\sigma_R^2 \equiv \int \mathrm{d}k\, P_\delta(k, t=0)\, W(k,R), \tag{1.63}$$

where R is the characteristic scale of the kernel function, $W(k,R)$, which is given by

$$W(k,R) = \frac{3k^2}{2\pi^2(kR)^3}\left[\sin(kR) - kR\cos(kR)\right], \tag{1.64}$$

where R is usually set to $8\,h^{-1}$Mpc, so σ_8 can the interpreted as the amplitude of the matter overdensity fluctuations at scales of $8\,h^{-1}$Mpc. This scale is chosen as a convention as σ_8 is roughly of order unity for realistic cosmological models.

As with the growth rate, $D(t)$, in Section 1.2.3, the transfer function depends on the type of matter/energy which dominates the expansion of the Universe at a given time. For a given cosmic epoch, $T(|\boldsymbol{k}|)$ is given by the Einstein-Boltzmann equations which take into consideration radiative coupling, the fluid equations given in Equation (1.49) and the evolution of the metric (see e.g. Kodama & Sasaki 1984; Sugiyama 1989).

These equations are typically solved numerically using tools such as CAMB (Code for Anisotropies in the Microwave Background; Lewis et al. 2000) or CLASS (Cosmic Linear Anisotropy Solving System; Lesgourgues 2011). For example, Figure 1.3 shows the matter power spectrum as calculated with CAMB assuming a ΛCDM cosmology consistent with Planck Collaboration et al. (2020) at three different look-back times. In Figure 1.3, the power spectrum at large scales (small k), is largely determined by the primordial power spectrum such that

$P(k) \propto k^{n_s-1}$. This trend continues until the peak of all three power spectra which gives the scale of the horizon of Universe when matter-radiation equality occurs, $k_{\rm eq}$, where $\Omega_{\rm m}(t_{\rm eq}) = \Omega_{\rm R}(t_{\rm eq})$. When $k < k_{\rm eq}$, the scale of the fluctuations exceeds the horizon at $k_{\rm eq}$, so any such modes would only enter the horizon during the matter-dominated epoch, while for $k > k_{\rm eq}$, the perturbations enter the horizon during the radiation-dominated epoch. At $k > k_{\rm eq}$, one can also notice some wiggles in the power spectrum. These occur at the characteristic scale of Baryonic Acoustic Oscillations (BAOs). These BAOs occur because before recombination baryonic matter and photons are coupled through Thomson scattering. After matter-radiation equality, as the dark matter inhomogeneities begin to grow. Any inhomogeneities in the baryonic matter can only grow until Thomson scattering smooths away any inhomogeneities, in other words, the perturbations to baryonic density oscillate at a scale set by speed of sound in the baryonic plasma. As soon as recombination occurs and baryonic matter becomes electrically neutral; photons and baryons decouple and these oscillations are imprinted into the matter density perturbations. Figure 1.3 also shows how the matter power spectrum evolves with look-back time (or redshift): as the particle horizon grows with decreasing look-back time, the overall power of the power spectrum at all modes increases, because a larger comoving volume is causally coupled with itself.

To couple the matter power spectrum to observable quantities in galaxy clustering surveys, we assume that the number density of galaxies traces the underlying matter density contrast, δ, as a function of some linear bias, given by $b(|\boldsymbol{k}|)$ as follows

$$P_{\rm gg}(|\boldsymbol{k}|,t) = b^2(|\boldsymbol{k}|)P_\delta(|\boldsymbol{k}|,t), \tag{1.65}$$

where $P_{\rm gg}(|\boldsymbol{k}|,t)$ is the three-dimensional galaxy power spectrum which is defined as

$$\langle \tilde{\delta}_{\rm g}(\boldsymbol{k},t)\tilde{\delta}_{\rm g}(\boldsymbol{k}',t)\rangle = (2\pi)^3\delta_{\rm D}^{(3)}(\boldsymbol{k}+\boldsymbol{k}')P_{\rm gg}(|\boldsymbol{k}|,t), \tag{1.66}$$

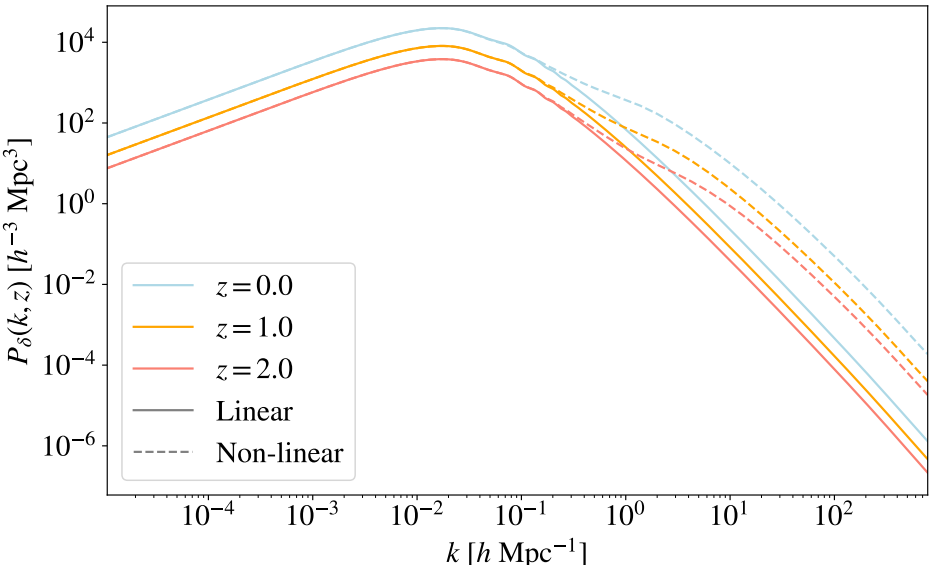

Figure 1.3: Plot of the three-dimensional matter power spectrum, $P_\delta(k,z)$, as a function of both wavenumber, k, and redshift, z. The solid lines show the linear matter power spectrum, while the dashed lines show the non-linear matter power spectrum including baryonic effects at small scales. The light blue line shows the power spectra at $z = 0$, the orange lines at $z = 1.0$ and the red lines at $z = 2.0$. All power spectra were calculated using CAMB (Code for Anisotropies in the Microwave Background; Lewis et al. 2000) and assuming a flat ΛCDM cosmology consistent with Planck Collaboration et al. (2020) and a halofit model (Smith et al., 2003; Takahashi et al., 2012; Bird et al., 2012) for the non-linear matter power spectrum with $A_{\rm bary} = 3.13$, $\eta_{\rm bary} = 0.603$ and $\log(T_{\rm AGN}) = 7.8$. Figure made by the author.

where $\tilde{\delta}_{\rm g}(\boldsymbol{k},t)$ is the Fourier transform of $\delta_{\rm g}(\boldsymbol{\Sigma},t) \equiv \delta n(\boldsymbol{\Sigma},t)/\overline{n}(t)$ where the number distribution of galaxies, $n(t)$, is often assumed to be point-like as follows

$$n(t) = \overline{n}(t) + \delta n(\boldsymbol{\Sigma}) = \sum_i \delta_{\rm D}^{(3)}(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_i). \tag{1.67}$$

This relationship is usually assumed to hold within the linear regime where $\delta \ll 1$, the structure due to galaxy sub-halos is not resolved, non-linearities due to baryonic feedback are small and gravitational non-linearities are not dominant.

1.2.4 Non-Linear Structure Growth

Throughout Section 1.2.3, there have been repeated assumptions of linearity: only considering linear perturbations to the fluid equation, neglecting any additional in-

teractions due to baryonic physics (particularly at small scales) and assuming a linear galaxy bias. At large scales, these assumptions have been shown to be accurate, but as we aim to model smaller scales, scales smaller than $|\boldsymbol{k}| \sim 10^{-1}\, h\mathrm{Mpc}^{-1}$, there can be substantial deviations from linearity in the observed matter power spectrum (see Figure 1.3).

As δ approaches unity at smaller scales, the non-linear terms in Equation (1.49) become relevant. In this, so-called "weak" non-linear regime, the allowed solutions for $\delta(\boldsymbol{k})$ have phases which are substantially shifted from their initial values. As a consequence, the distribution of the density fluctuations become increasingly non-Gaussian as they deviate from the Gaussian primordial field (Davis et al., 1985; Sahni & Coles, 1995; Jenkins et al., 1998).

At even smaller scales, additional non-linearities may become relevant as the matter distribution deviates even further from Gaussianity and the strength of baryonic interactions becomes comparable to the magnitude of gravitational potentials. This is know as the regime of "strong" non-linearity.

At scales near $|\boldsymbol{k}| \sim 2\, h\mathrm{Mpc}^{-1}$, as shown in Figure 1.3, non-linearities in the matter power spectrum start contributing a considerable fraction of the overall power, as gravitationally bound structures subject to baryonic physics start to form at these scales. The dynamics of the contents of such objects are non-trivial and highly non-linear as hydrodynamical effects, magnetodynamics, star formation and the thermodynamics of intergalactic gas start influencing the growth of structure of baryons. Therefore, these effects will not only alter the overall matter power spectrum, but also complicate the correlation between the distribution of galaxies and the underlying dark matter distribution.

Regarding the matter power spectrum, the most accurate way to track any non-linear effects is the use of large numerical simulations which include any baryonic interactions at small scales. However, since these simulations are difficult and expensive to compute, it is common to opt for semi-analytical models which are calibrated with simulations. One popular such model is the halo model (Seljak, 2000; Peacock & Smith, 2000; Ma & Fry, 2000) which assumes that all matter is bound

within spherical halos and their clustering can be explained by considering the clustering between halo pairs and the clustering within halos themselves. Nevertheless, it has been found that a purely analytical halo model does not accurately recover the non-linearities in the matter power spectrum when compared to simulations (Tinker et al., 2005; Valageas & Nishimichi, 2011; Mead et al., 2015; van Daalen & Schaye, 2015). For this reason, as is the case in this thesis, it is common to opt for a fitting function approach based on the halo model which is calibrated by simulations such as halofit (Smith et al., 2003; Takahashi et al., 2012; Bird et al., 2012) or HMCODE (Halo Model code; Mead et al. 2015, 2016, 2021). These models either create a functional form for the halo model or they alter the halo model as a function of some parameters. The latter approach is the one used in HMCODE and it has been found to yield better agreements with simulations (Mead et al., 2021). The main free parameters typically varied in these models is the baryonic feedback amplitude, $A_{\rm bary}$, which scales the halo-mass concentration with redshift, and the halo bloating parameter, η, which scales the size of the halo mass density profile (Navarro et al., 1997).

The calibration of halo model fitting functions is typically based on hydrodynamical simulations, such as COSMO-OWLS (Le Brun et al., 2014) or BAHAMAS (McCarthy et al., 2017), which include accurate modelling for gravitational collapse, star formation, thermodynamics, etc. Among these effects, one of the dominant factors for the suppression of structure formation at small scales in baryonic feedback due to Active Galactic Nuclei (AGN). AGN are thought to be supermassive black holes at the centres of galaxies which are actively accreting interstellar gas which subsequently becomes heated as it is accelerated by the black holes' gravity. This effect can heat the gas surrounding a supermassive black hole to such an extend that the gas expands outside of the virial radius of the galaxy's halo (Schaye et al., 2010; van Daalen et al., 2011; Martizzi et al., 2014; van Daalen & Schaye, 2015). This effect can therefore significantly disturb the density profile of the baryonic matter in a halo which affects $A_{\rm bary}$ and η (Duffy et al., 2010). Thus, AGN feedback is often implicitly encoded in the semi-analytical fits in halofit or HM-

CODE, while the most recent iteration of HMCODE (Mead et al., 2021) allows to explicitly vary the amplitude of AGN feedback as a function of the average temperature of AGN ($T_{\rm AGN}$).

If we are interested in modelling the three-dimensional galaxy power spectrum from the matter power spectrum, as defined in Equation (1.66), baryonic physics add an additional layer of complication as Equation (1.65) may not hold for large values of $|\boldsymbol{k}|$ as the galaxy bias, b, deviates from linearity too. Within the context of the halo model, firstly the clustering of halos may not necessarily follow the same power spectrum as dark matter which may be quantified by a halo bias (Cole & Kaiser, 1989). Additionally, it is typical to assume a linear galaxy bias with $b = 1$ is equivalent to assuming that each dark matter halo has one galaxy at its centre. Although to first-order, this is roughly the case, we know that this cannot be the full picture, since systems of multiple galaxies are known to exist frequently, while it is also known that most galaxies have multiple dwarf galaxies surrounding them. There are many models which address this scale dependence in the halo and galaxy bias through parametric alterations of the relation shown in Equation (1.65) (Cole & Kaiser, 1989; Mo & White, 1996; Jing, 1998; Sheth et al., 2001; Seljak, 2000; Pillepich et al., 2010; Tinker et al., 2010; Bhattacharya et al., 2011).

1.3 Galaxy Surveys as Probes of Large-Scale Structure

In Section 1.2, I have outlined how perturbations in the energy/matter density of the Universe evolve and grow under ΛCDM, how we may quantify them and how they may be traced by galaxies. These concrete predictions of ΛCDM for the statistical distribution of dark matter and galaxies allows us to test the model through galaxy surveys which constrain these statistics in the late Universe. There are three main probes of this which can be constructed from observables: galaxy clustering, weak gravitational lensing and galaxy-galaxy-lensing.

In Section 1.3.1, I discuss galaxy clustering which refers to the study of the spatial distribution of galaxies in the sky and along the line of sight. Section 1.3.2 presents weak-gravitational lensing and how this observable measures cosmic shear, which is a direct probe of the underlying matter power spectrum. Section 1.3.3 shows how both of these ideas can be combined to measure how galaxies in the foreground gravitationally lens galaxies in the background. Lastly, Section 1.3.4 discusses the current state-of-the-art in the field of cosmological galaxy surveys as well as future prospects.

1.3.1 Galaxy Clustering

Galaxy clustering is the phenomenon of galaxies spatially coalescing at certain scales in accordance with the underlying distribution of matter and the dynamics of galaxy evolution. This can be measured by galaxy surveys through the mapping of galaxy positions on the sky, $\boldsymbol{x}$, and along the line of sight by measuring the galaxies' redshift, z. Once this is measured, one can define summary statistics which compress the information in the galaxy positions. The most commonly used one is a two-point correlation function, $\xi_{\rm gg}$, which is defined from theory as

$$\xi_{\rm gg}(\Delta x \equiv |\boldsymbol{x}_i - \boldsymbol{x}_j|, z) \equiv \langle \delta_{\rm g}(\boldsymbol{x}_i, z)\, \delta_{\rm g}(\boldsymbol{x}_j, z) \rangle, \tag{1.68}$$

where i and j are indices representing different galaxy positions, $\Delta x \equiv |\boldsymbol{x}_i - \boldsymbol{x}_j|$ since we assume isotropy in ΛCDM which implies that the structure in the Universe only

depends on the scalar distance, and $\delta_{\rm g}$ are the deviations from the mean in the number counts of galaxies as shown in Equation (1.66). Equation (1.66) can be rewritten in spatial coordinates as follows

$$\xi_{gg}(\Delta x, z) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} P_{\rm gg}(|\boldsymbol{k}|, z)\, e^{-i|\boldsymbol{k}|\Delta x}. \tag{1.69}$$

To estimate $\xi_{\rm gg}(\Delta x, z)$ from observations, we must first discretise the domain: Δx is binned into separation bins with the index $\imath$, while z is binned into redshift bins (also known as lens bins or tomographic bins) indexed by i and j. Then, for a given redshift bin pair, we can estimate $\xi_{\rm gg}(\Delta x, z)$ by comparing the spatial distribution of the observed galaxies to the spatial distribution of randomly sampled points. To do so, a commonly used estimator is the Landy-Szalay estimator (Landy & Szalay, 1993) for the observed two-point correlation function, $\hat{\xi}_{\rm gg}^{(ij)}$, given by

$$\hat{\xi}_{\rm gg}^{(ij)}(\Delta x_\imath) = 1 + \left(\frac{N_{\rm rand}}{N}\right)^2 \frac{DD^{(ij)}(\Delta x_\imath)}{RR(\Delta x_\imath)} - \frac{N_{\rm rand}}{N}\left(\frac{DR^i(\Delta x_\imath)}{RR(\Delta x_\imath)} + \frac{DR^j(\Delta x_\imath)}{RR(\Delta x_\imath)}\right), \tag{1.70}$$

where N is the number of galaxies in i and j, $N_{\rm rand}$ is the number of random points sampled usually from a Poisson distribution or form a uniform distribution for a perfect survey, $DD^{(ij)}(\Delta x_\imath)$ is the number of galaxy pairs between sample i and sample j which are at a spatial separation such that they fall within the $\imath^{\rm th}$ bin, $DR^{(ij)}(\Delta x_\imath)$ is the number of pairs between galaxy sample i and the random sample of points which are at a spatial separation such that they fall within the $\imath^{\rm th}$ bin, and $RR(\Delta x_\imath)$ is the number of point pairs from the random sample which are at a distance n such that they fall within the $\imath^{\rm th}$ bin. In practice, as we are free to sample an arbitrary number of random points, it is common to set $N_{\rm rand} \gg N$ for better numerical accuracy.

It is important to note that $\hat{\xi}_{\rm gg}^{(ij)}(\Delta x_\imath)$ may be biased due to selection effects in the galaxy survey in question. Along the line of sight, the flux limit of the survey implies that further away galaxies are less likely to be observed as they tend to be fainter. At the same time, if the survey is photometric, the redshift estimates may

be biased, so the galaxies might end up in the incorrect bin. This can in turn bias the $\hat{\xi}_{\rm gg}^{(ij)}(\Delta x_\iota)$ measurements if this bias is not reflected in the redshift distribution of the lens galaxy sample in question. In the angular direction of the sky, the galaxy counts could be biased due to a variety of reasons which are non-isotropic: galactic absorption, variations in the seeing from observation to observation (changing the apparent magnitude limit as function of the direction) or masking of certain areas due to foreground light sources such as stars. At the same time, due to the magnitude limit (which itself may not be isotropic), an additional source of non-isotropic bias is added called the magnification bias. As is discussed in detail in Section 1.3.2, gravitational lensing due to matter in the foreground of a galactic light source may focuses light such that the observed flux from the light source is boosted to the point where it surpasses the flux limit of the galaxy survey. Therefore, a galaxy which should not have been observed according to the selection function of the survey will be observed, which may lead to the overcounting of galaxies. The reverse can occur when the foreground matter changes the solid angle from the source such that flux decreases. As this effect is correlated with the foreground dark matter distribution, it may add additional structure to the observed galaxy count fields.

Additionally, as we consider the spatial distribution of galaxies, a random realisation of many possible ones given the underlying galaxy power spectrum, $\hat{\xi}_{\rm gg}^{(ij)}(\Delta x_\iota)$ is subject to shot noise. This is often modelled as white Poisson noise such that

$$\hat{P}_{\rm gg}(|\boldsymbol{k}|,z) = P_{\rm gg}(|\boldsymbol{k}|,z) + \frac{1}{\overline{n}_{\rm gal}(z)}, \tag{1.71}$$

where $\hat{P}_{\rm gg}(|\boldsymbol{k}|,z)$ is the inferred galaxy power spectrum and $\overline{n}_{\rm gal}(z)$ is the average three-dimensional galaxy number density at redshift, z. Simultaneously, the shot noise is increasing the statistical uncertainty of the power spectrum (see e.g. Wadekar & Scoccimarro 2020 for detailed modelling of the shot noise in the galaxy clustering covariance).

1.3.2 Weak Gravitational Lensing and Cosmic Shear

Gravitational lensing is a phenomenon predicted by General Relativity which involves light from distant sources being deflected from its geodesic in Minkowski space due to matter in the foreground curving space-time and light travelling along the geodesic in this curved space-time. This phenomenon is typically split into three regimes: strong gravitational lensing (the mass of the gravitational lens is large enough for multiple images of the source to appear), weak gravitational lenses (the mass of the lens slightly distorts the shape and the size of a single image) and microlensing (the mass of the lens is small, so it does not measurably distort the shape and size of the source, but alters its brightness, while potentially varying over observable time scales). In this thesis, I am exclusively interested in the weak gravitational lensing regime on cosmological scales. To introduce how this phenomenon can be modelled and observed, this section follows the reviews given in Bartelmann & Schneider (2001), Schneider (2003), Van Waerbeke & Mellier (2003), Schneider (2006), Heavens (2009) and Kilbinger (2015).

On cosmological scales, all light emitted by distant sources, such as galaxies, experiences weak gravitational lensing due to the matter density perturbations in the foreground. Therefore, the weak gravitational lensing signal ought to be coupled directly to the matter power spectrum. This signal, when considering only the shape distortions of galaxies and removing any contaminating signals, is known as cosmic shear. To model this, we alter the FLRW metric given in Equation (1.9) by considering a nearly flat metric ($k = 0$) with a gravitational potential field, Φ, given by

$$\mathrm{d}s^2 = \big(1+\Phi\big)\mathrm{d}t^2 - \big(1-\Phi\big)a^2(t)\mathrm{d}\boldsymbol{\Sigma}^2, \tag{1.72}$$

where $\boldsymbol{\Sigma}$ is given in hyperspherical coordinates as defined in Equation (1.10) and the gravitational potential field, Φ, is related to the overdensity field, δ, through the Poisson equation as follows

$$\nabla^2\Phi = \frac{3H_0^2\Omega_\mathrm{m}}{2a(t)}\delta. \tag{1.73}$$

In general, any particle in motion on the worldline, x^α, within space-time follows a geodesic described by the geodesic equation as follows

$$\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\lambda^2} + \Gamma^\alpha_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} = 0, \tag{1.74}$$

where λ is the affine parameter and $\Gamma^\alpha_{\mu\nu}$ is the Christoffel symbol as given by Equation (1.6). Considering a photon travelling within a manifold given by Equation (1.72), it must be moving on a light-like geodesic such that $\mathrm{d}s^2 = 0$ and therefore a comoving distance interval is given by $\mathrm{d}\chi = \pm a^{-1}(t)\mathrm{d}t \equiv \mathrm{d}\eta$ (where the positive root of the proper time interval, $\mathrm{d}\eta$, represents an outgoing photon and the negative root an incoming one). Assuming that the photon is incoming into the gravitational potential, it can be shown that the geodesic for the time-like coordinate, η, to zeroth order in Φ reduces to the following (Heavens, 2009),

$$\frac{\mathrm{d}\eta}{\mathrm{d}\lambda} = \frac{1}{a^2}. \tag{1.75}$$

We can also change the coordinate frame such that the gravitational potential, Φ, is defined as a function of the transverse comoving coordinate along the line of sight, $f_k(\chi)$, and two transverse coordinates which are perpendicular to the line of sight, x_1 and x_2. Combining Equation (1.75) with the geodesic equation in the zeroth order of Φ for x_1 and x_2, one can find that

$$\frac{\mathrm{d}^2 x_i}{\mathrm{d}\eta^2} = \frac{-2\,\partial\Phi(x_1, x_2, f_k(\chi))}{\partial x_i}, \tag{1.76}$$

where $x_i \in \{1,2\}$. Integrating this with respect to $\mathrm{d}\eta = -\mathrm{d}\chi$ twice,

$$x_i = f_k(\chi)\theta_i - 2\int_0^\chi \mathrm{d}\chi'[f_k(\chi) - f_k(\chi')]\frac{\partial\Phi(x_1', x_2', f_k(\chi'))}{\partial x_i'}, \tag{1.77}$$

where θ_i is an integration constant which through a dimensionality argument must represent a small angle deviation in the transverse direction, such that $x_i = f_k(\chi)\theta_i$ in the absence of a gravitational potential. It is important to note here that taking the

integral along the radial direction where $f_k(\chi) \parallel f_k(\chi')$ is an approximation, as in reality the path length travelled by the photon is altered by small deviations caused by fluctuations in the metric. Having said this, we may safely assume this so-called Born approximation, as it is remarkably accurate in the weak lensing regime (Petri et al., 2017). Taylor expanding $\frac{\partial \Phi}{\partial x_i'}$ to linear order in both x_1 and x_2 and transforming the coordinates to $\theta_i \equiv x_i / f_k(\chi)$, one finds that

$$x_i = f_k(\chi)\theta_i - 2\int_0^\chi \mathrm{d}\chi' [f_k(\chi) - f_k(\chi')] \left(\frac{\partial \Phi(\theta_i' = 0, \theta_j', f_k(\chi'))}{\partial \theta_i'} + \theta_j f_k(\chi') \frac{\partial^2 \Phi(\theta_1', \theta_2', f_k(\chi'))}{\partial \theta_i' \partial \theta_j'} \right). \quad (1.78)$$

Thus, the deviation in x_i between two nearby photons that have been lensed by Φ is given by

$$\Delta x_i(\mathbf{\Sigma}) = f_k(\chi)\Delta\theta_j(\delta_{ij} - \phi_{ij}(\mathbf{\Sigma})), \quad (1.79)$$

where δ_{ij} is the Kronecker delta and ϕ_{ij} is defined as

$$\phi_{ij}(\mathbf{\Sigma}) \equiv 2\int_0^\chi \mathrm{d}\chi' \frac{[f_k(\chi) - f_k(\chi')]}{f_k(\chi) f_k(\chi')} \frac{\partial^2 \Phi(\mathbf{\Sigma})}{\partial \theta_i' \partial \theta_j'}. \quad (1.80)$$

In the case where the observations are of many sources in population ι which are distributed along the line of sight according to $p^{(\iota)}(\chi)$, we can sum over Equation (1.80) for all sources in the distribution, giving that

$$\phi_{ij}(\mathbf{\Sigma}) \equiv 2\int_0^\chi \mathrm{d}\chi' \frac{g^{(\iota)}(\chi')}{f_k(\chi)} \frac{\partial^2 \Phi(\mathbf{\Sigma})}{\partial \theta_i' \partial \theta_j'}. \quad (1.81)$$

where $g^{(\iota)}(\chi)$ is given by

$$g^{(\iota)}(\chi) \equiv \int_\chi^\infty \mathrm{d}\chi' p^{(\iota)}(\chi') \frac{f_k(\chi') - f_k(\chi)}{f_k(\chi')}. \quad (1.82)$$

Equation (1.79) is a linear mapping from the source to the image as it reaches

the observer. This can be expressed with the amplification matrix, $\mathscr{A}$, which is defined as follows

$$\mathscr{A}_{ij} \equiv \delta_{ij} - \phi_{ij}. \tag{1.83}$$

As θ_i and θ_j are interchangeable, $\mathscr{A}_{ij}$ must be symmetric and it has only 3 degrees of freedom. With that, it is convenient to expand the amplification matrix into an isotropic expansion term and a shear term as follows

$$\mathscr{A}_{ij} = \begin{pmatrix} 1-\kappa & 0 \\ 0 & 1-\kappa \end{pmatrix} + \begin{pmatrix} -\gamma_1 & -\gamma_2 \\ -\gamma_2 & \gamma_1 \end{pmatrix}, \tag{1.84}$$

where κ is the convergence and $\gamma \equiv \gamma_1 + \mathrm{i}\gamma_2$ is the complex shear. These terms relate back to Equation (1.83) as follows

$$\begin{aligned} \kappa &\equiv \frac{1}{2}(\phi_{11} + \phi_{22}), \\ \gamma &\equiv \frac{1}{2}(\phi_{11} - \phi_{22}) + \mathrm{i}\phi_{12}. \end{aligned} \tag{1.85}$$

Upon inspection of Equation (1.84), it becomes apparent that the convergence, κ, isotropically magnifies the size of the observed image for $\kappa > 0$, while γ_1 stretches the image along the θ_1 axis and γ_2 stretches the image of the source along the $\theta_2 = \pm\theta_1$ diagonal axis. The amplification matrix is also often expressed in terms of the reduced shear, g, as follows

$$\mathscr{A}_{ij} = (1-\kappa) \begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix}, \tag{1.86}$$

where $g \equiv \gamma/(1-\kappa) = g_1 + \mathrm{i}g_2$. To map the image of the source from the observed image, we must invert the amplification matrix. This is possible in the weak lensing limit, where $|\kappa| \ll 1$ and $|\gamma| \ll 1$, and as with any invertable matrix, $\mathscr{A}_{ij}^{-1} \propto 1/\det(\mathscr{A})$ holds. Therefore, the size of the image is scaled with respect to the source by the magnification, μ, given by

$$\mu \equiv \frac{1}{\det(\mathscr{A})} = \frac{1}{(1-\kappa)^2 - \gamma^2} \approx 1 + 2\kappa. \tag{1.87}$$

Here, μ only affects the overall size of the source, as gravitational lensing conserves surface brightness, the flux of the source may be magnified. Hence, in flux-limited galaxy surveys, there may be a non-isotropic bias in the galaxy counts which is known as the magnification bias. Although this is not a problem for current and upcoming cosmic shear measurements (Deshpande et al., 2020; Duncan et al., 2022), the magnification bias can be considerable in the galaxy clustering (see Section 1.3.1) and in the galaxy-galaxy lensing signals (see Section 1.3.3). When measuring galaxy shapes, the measurement is only sensitive to changes in the shape, so one can only determine the reduced shear from observations. To illustrate how such observations allow to probe the underlying large-scale structure, Figure 1.4 exemplifies the shear induced by the convergence field given by the large-scale structure along a given line of sight. The left panel shows convergence along the line of sight over a two-dimensional plane. On the right of Figure 1.4, we can see that the associated shear vector field is correlated with the convergence field such that regions with larger convergence induce a coherent shear tangential to the overdense region. Therefore, the observed shear can be used to reconstruct the properties of the underlying convergence.

To make the relation between the convergence, κ, of galaxy population i with the underlying matter overdensity, δ, we can combine Equation (1.85) with the Poisson equation given by Equation (1.73), such that

$$\kappa^{(i)}(\mathbf{\Sigma}) = \frac{3H_0^2\Omega_\mathrm{m}}{2} \int_0^\chi \mathrm{d}\chi' \frac{f_k(\chi')\, g^{(i)}(\chi')}{f_k(\chi)} \frac{\delta(\mathbf{\Sigma})}{a(\chi')}. \tag{1.88}$$

The integral on the right-hand side of the equation means that $\kappa^{(i)}(\mathbf{\Sigma})$ at a given look-back time, z or χ, is two-dimensional. Therefore, the statistics of the convergence field, κ, can be captured by a two-dimensional angular power spectrum, $C_{\kappa\kappa}$, which is defined through

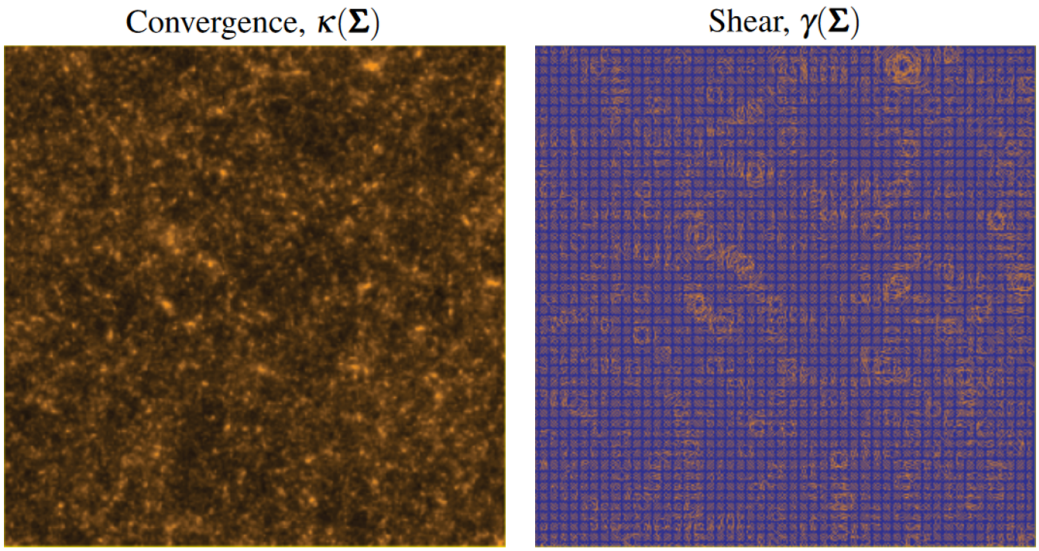

Figure 1.4: Spatial maps showing convergence (left panel) and the associated shear field (right panel) over a two-dimensional plane. On the left, the lighter orange regions represent areas with large convergence, while the dark regions represent areas with small values of convergence, i.e. less matter along the line of sight. On the right, the lines of the vector field represent the average direction and magnitude of the observed shear which correlates to the convergence shown on the left panel. Figure based on White & Hu (2000).

$$\langle \tilde{\kappa}^{(i)}(\boldsymbol{\ell},\chi)\tilde{\kappa}^{(j)*}(\boldsymbol{\ell}',\chi)\rangle = (2\pi)^2\delta_{\rm D}^{(2)}(\boldsymbol{\ell}-\boldsymbol{\ell}')C_{\kappa\kappa}^{(ij)}(|\boldsymbol{\ell}|), \tag{1.89}$$

where $\boldsymbol{\ell}$ is the two-dimensional wavevector, $\delta_{\rm D}^{(2)}$ is the two-dimensional Dirac delta function and $\tilde{\kappa}^{(i)}(\boldsymbol{\ell})$ is the two-dimensional Fourier transform of $\kappa^{(i)}(\mathbf{\Sigma})$ defined by

$$\begin{aligned}\tilde{\kappa}^{(i)}(\boldsymbol{\ell},\chi) &= \int \mathrm{d}^2\mathbf{\Sigma}\,\kappa^{(i)}(\mathbf{\Sigma})\,e^{-i\boldsymbol{\ell}\cdot\mathbf{\Sigma}} \\ &= \int_0^\infty \mathrm{d}\chi W_\kappa^{(i)}(\chi)\int \mathrm{d}^2\mathbf{\Sigma}\,\delta^{(i)}(\mathbf{\Sigma},\chi)\,e^{-i\boldsymbol{\ell}\cdot\mathbf{\Sigma}},\end{aligned} \tag{1.90}$$

where $W_\kappa^{(i)}(\chi) \equiv \frac{3H_0^2\Omega_{\rm m}}{2} f_k(\chi)\,g^{(i)}(\chi)/a(\chi)$ and $\delta^{(i)}(\mathbf{\Sigma},\chi)$ is the three-dimensional overdensity field as defined in Equation (1.45). Substituting Equation (1.90) into Equation (1.89), while replacing $\delta^{(i)}$ with its Fourier transform, $\tilde{\delta}^{(i)}$ using Equation (1.52), and using Equation (1.61), one finds that

$$C_{\kappa\kappa}^{(ij)}(\ell) = \int_0^\infty \mathrm{d}\chi W_\kappa^{(i)}(\chi) \int_0^\infty \mathrm{d}\chi' W_\kappa^{(j)}(\chi') \int \mathrm{d}k\, k^2 P_\delta(k,\chi,\chi') \frac{j_\ell(k\chi)}{[k f_k(\chi)]^2} \frac{j_\ell(k\chi')}{[k f_k(\chi')]^2}, \tag{1.91}$$

where j_ℓ is the ℓ^{th} spherical Bessel function and $P_\delta(k,\chi,\chi')$ is the unequal-time three-dimensional power spectrum (i.e. the power spectrum between matter fields at different look-back times/comoving distances from the observer) which is commonly approximated taking the geometric mean of the matter power spectrum at each time as (Castro et al., 2005; Kitching & Heavens, 2017),

$$P_\delta(k,\chi,\chi') \approx \sqrt{P_\delta(k,\chi) P_\delta(k,\chi')}. \tag{1.92}$$

In addition, there is the so-called Limber approximation (Limber, 1953; Kaiser, 1992) which consists in a Taylor expansion around the approximate maximum of the Bessel function at $k f_k(\chi) = \ell + 1/2$, and only considering the first-order term. In this case, Equation (1.91) reduces to

$$C_{\kappa\kappa}^{(ij)}(\ell) \approx \int_0^\infty \frac{d\chi}{f_k^2(\chi)} W_\kappa^{(i)}(\chi) W_\kappa^{(j)}(\chi) P_\delta\left(\frac{\ell+1/2}{f_k(\chi)}, \chi\right). \tag{1.93}$$

As can be seen from Equation (1.93), the Limber approximation reduces the number of integrals needed from three in Equation (1.91) to one, while making the Bessel function in Equation (1.91) vanish which are highly oscillatory and numerically expensive to integrate. At the same time, Limber approximation has proven to be accurate for cosmic shear particularly for large ℓ. Still, the bias in the signal due to the Limber approximation for $\ell < 40$ is $< 1\%$ (Kitching et al., 2017; Lemos et al., 2017). To get the cosmic shear angular power spectrum, $C_{\text{GG}}^{(ij)}(\ell)$, on the curved sky, we simply take (Lemos et al., 2017),

$$C_{\text{GG}}^{(ij)}(\ell) = \frac{(\ell+2)!}{(\ell-2)!} \frac{1}{\ell^2(\ell+1)^2} C_{\kappa\kappa}^{(ij)}(\ell), \tag{1.94}$$

where ! indicates the factorial.

When constraining cosmological parameters using weak gravitational lensing,

the signal from two-point statistics given by Equation (1.94) is degenerate for certain combinations of σ_8 (as defined in Equation (1.63)) and Ω_m. To account for this when reporting cosmological results from weak gravitational lensing, it is common to use the following reparametrisation given by

$$S_8 \equiv \sigma_8 \sqrt{\frac{\Omega_\mathrm{m}}{0.3}}. \tag{1.95}$$

When conducting a (photometric) weak gravitational lensing survey, the shapes, photometric redshifts and positions of galaxies are measured. The shapes of the galaxies can be measured using parametric fitting or free-form algorithms. Either methodology requires corrections for instrumental biases in the shapes due to the point-spread function of the telescope, charge-transfer inefficiencies in the charged-coupled device (CCD), etc. (see Heymans et al. 2006; Massey et al. 2007; Miller et al. 2007; Kitching et al. 2010 for details on shape calibration). After obtaining shape measurements from images and applying the necessary corrections, the resulting galaxy ellipticity, ε, is an estimator of the galaxy's reduced shear applied to the intrinsic galaxy ellipticities, i.e. $\varepsilon = \varepsilon_1 + \mathrm{i}\varepsilon_2 = \hat{\gamma} + \hat{\varepsilon}_\mathrm{int}$. In addition, the photometric redshifts are estimated from multiband photometry. This can be achieved with a plethora of algorithms, but they are always based on calibrating the algorithm with a spectroscopic reference sample (see e.g. Lima et al. 2008; Wright et al. 2020a). Depending on the photometric redshift estimate, the source galaxies are binned in a given tomographic redshift bin i. With these measurements made, the data is commonly compressed into a two-point correlation function over galaxy pairs, α and β, as a summary statistic defined as

$$\hat{\xi}^{(ij)}_{\varepsilon\varepsilon,\pm}(\theta) = \frac{\sum_{\alpha\in i,\beta\in j} w_\alpha w_\beta [\varepsilon_1^{(i)}(\boldsymbol{\theta}_\alpha)\varepsilon_1^{(j)}(\boldsymbol{\theta}_\beta) \pm \varepsilon_2^{(i)}(\boldsymbol{\theta}_\alpha)\varepsilon_2^{(j)}(\boldsymbol{\theta}_\beta)]}{\sum_{\alpha\in i,\beta\in j} w_\alpha w_\beta}, \tag{1.96}$$

where w_α is the galaxy weight for a galaxy α, which accounts for selection effects and systematic biases, and $\hat{\xi}^{(ij)}_{\varepsilon\varepsilon,\pm}(\theta)$ is the estimated spatial two-point correlation function as a function of the angular separation, θ between a given galaxy pair

from tomographic bins i and j. Note that Equation (1.96) does not consider any systematic effects which may shift or scale the measured ellipticities. $\hat{\xi}^{(ij)}_{\varepsilon\varepsilon,+}$ gives the correlation between the shapes as a function of separation, while $\hat{\xi}^{(ij)}_{\varepsilon\varepsilon,-}$ yields the anti-correlation between the shapes. Since the measured ε also includes the intrinsic galaxy ellipticities, this induces auto-correlations in the galaxy shapes when $i = j$ known as the shape noise. To account for this, one can subtract the shape noise from the autocorrelations as follows (Schneider et al., 2002),

$$\langle \varepsilon_1^{(i)}(\boldsymbol{\theta}_\alpha)\varepsilon_1^{(j)}(\boldsymbol{\theta}_\beta) + \varepsilon_2^{(i)}(\boldsymbol{\theta}_\alpha)\varepsilon_2^{(j)}(\boldsymbol{\theta}_\beta)\rangle = \hat{\xi}^{(ij)}_{\gamma\gamma,+}(\theta) + \delta_{ij}\sigma_\varepsilon^{(i)2}, \tag{1.97}$$

where δ_{ij} is the Kronecker delta and $\sigma_\varepsilon^{(i)}$ is the shape noise due to the galaxies' intrinsic ellipticities as given by

$$\sigma_\varepsilon^{(i)2} \equiv \frac{\sum_{\alpha\in i} w_\alpha^2 \left[\varepsilon_1^{(i)2} + \varepsilon_2^{(i)2}\right]}{\sum_{\alpha\in i} w_\alpha^2}. \tag{1.98}$$

$\hat{\xi}^{(ij)}_{\gamma\gamma,\pm}(\theta)$ can be related to the curl-free weak lensing angular power spectrum, $\hat{C}^{(ij)}_{\gamma\gamma}$ by performing an inverse Hankel transform as follows

$$\hat{C}^{(ij)}_{\gamma\gamma}(\ell) = 2\pi \int_0^\infty \mathrm{d}\theta\, \theta\, \xi_+(\theta) J_0(\ell\theta) = 2\pi \int_0^\infty \mathrm{d}\theta\, \theta\, \xi_-(\theta) J_4(\ell\theta), \tag{1.99}$$

where $J_0(\ell\theta)$ and $J_4(\ell\theta)$ are the zeroth and fourth order Bessel functions, respectively.

To relate the weak lensing angular power spectrum, $\hat{C}^{(ij)}_{\gamma\gamma}(\ell)$, to the cosmic shear angular power spectrum, $C^{(ij)}_{\mathrm{GG}}(\ell)$, one must take into consideration that cosmic shear is not the only physical effect which may coherently shear galaxy shapes as a function of scale. There is also the phenomenon of intrinsic alignments (IAs). IAs can be thought of as occurring in galaxies which formed under the influence of the same tidal fields such that their orientation and shape may be intrinsically correlated at certain scales (see e.g. Troxel & Ishak 2015 for a review). This effect can

be degenerate with cosmic shear and contributes to the weak gravitational lensing signal as follows

$$\hat{C}_{\gamma\gamma}^{(ij)}(\ell) = C_{\rm GG}^{(ij)}(\ell) + C_{\rm II}^{(ij)}(\ell) + C_{\rm GI}^{(ij)}(\ell) + C_{\rm IG}^{(ij)}(\ell), \tag{1.100}$$

where $C_{\rm II}^{(ij)}(\ell)$ is the angular power spectrum of the intrinsic alignment perturbations with themselves, while $C_{\rm GI}^{(ij)}(\ell)$ and $C_{\rm IG}^{(ij)}(\ell)$ are the angular power spectra of the intrinsic alignments perturbations correlated with the shear field. These summary statistics of the intrinsic alignment are typically modelled parametrically as a function of the underlying matter power spectrum. An example of such a model is the non-linear alignment model (Hirata & Seljak, 2004) which is discussed in more detail in Chapters 2 and 3.

1.3.3 Galaxy-Galaxy Lensing

In the weak gravitational lensing regime described in Section 1.3.2, the foreground density inhomogeneity which leads to a lensing potential does not necessarily have to be limited to large-scale structure in δ. Galaxies in the foreground (and the dark matter around them) may lens background source galaxies at higher redshifts in the same way; a phenomenon which is known as galaxy-galaxy lensing (GGL). This effect causes the observed shapes of distant source galaxies to correlate with the positions of observed galaxies in the foreground along the line of sight (Tyson et al., 1984; Brainerd et al., 1996; Fischer et al., 2000). As the foreground galaxies (often called lenses) also generally trace the overdensities due to large-scale structure in the Universe, the observed signal from galaxy-galaxy lensing is also a powerful cosmological probe in combination with galaxy clustering and galaxy-galaxy lensing.

Galaxy-galaxy lensing can be observed through the tangential shear, γ_+, of background sources relative to the line between the lens galaxy and the source galaxy. For a given source-lens pair, this is defined as

$$\gamma_+ \equiv -{\rm Re}(\gamma e^{-2i\phi}), \tag{1.101}$$

where γ is the shear measured on the source and ϕ is the angular position of the source galaxy with respect to the horizontal axis centred at the lens.

Over many source-lens pairs, one can define the average tangential shear, $\overline{\gamma}_+$, which depends on cosmology as follows

$$\overline{\gamma}_+^{(ij)}(\theta) \equiv \langle \gamma_+^{(ij)}(\theta) \rangle = \frac{1}{2\pi} \int_0^\infty \mathrm{d}\ell\, \ell C_{\mathrm{g}\kappa}^{(ij)}(\ell) J_2(\ell\theta), \tag{1.102}$$

where $C_{\mathrm{g}\kappa}^{(ij)}(\ell)$ is the galaxy-shear angular power spectrum for lens bin, i, and source bin, j, which in the Limber approximation is given by

$$C_{\mathrm{g}\kappa}^{(ij)}(\ell) \approx \int_0^\infty \frac{\mathrm{d}\chi}{f_k^2(\chi)} W_{\mathrm{g}}^{(i)}\left(\frac{\ell+1/2}{f_k(\chi)}, \chi\right) W_\kappa^{(j)}(\chi) P_\delta\left(\frac{\ell+1/2}{f_k(\chi)}\right), \tag{1.103}$$

where $W_{\mathrm{g}}^{(i)}(k,\chi)$ is the geometric weight function for galaxy clustering of the lens sample, i, given by

$$W_{\mathrm{g}}^{(i)}(k,\chi) \equiv b^{(i)}(k) p^{(i)}(\chi), \tag{1.104}$$

where $b^{(i)}(k)$ is the galaxy bias as defined in Equation (1.65) and $p^{(i)}(\chi)$ is the distribution of lens galaxies along the line of sight within galaxy population, i. The galaxy bias, $b^{(i)}(k)$, is often assumed to be approximately scale-independent such that $b^{(i)}(k) \approx b^{(i)}$, particularly, in the linear regime, but other approaches to induce scale-dependence exist (see e.g. Pen 1998; Baldauf et al. 2010).

To measure the galaxy-galaxy lensing signal between a lens sample of galaxies and a source sample, one needs the following measurements: the galaxy position measurements of a lens sample as described in Section 1.3.1 including their redshifts, the shape measurements of the galaxies in the source sample as described in Section 1.3.2, and the associated redshifts of each source. Firstly, as before, we bin the lens galaxies into redshift bins, i, and the source galaxies into redshift bins, j. Then, we may define the estimator for the tangential shear as follows

$$\hat{\gamma}_+^{(ij)}(\theta) = \frac{\sum_{\alpha,\beta} w_{(\alpha\beta)}^{(ij)} \gamma_{+,(\alpha\beta)}^{(ij)}(\theta)}{\sum_{\alpha,\beta} w_{(\alpha\beta)}^{(ij)}}, \tag{1.105}$$

where $w_{(\alpha\beta)}^{(ij)}$ is the weight for the pair of lens, α, and source, β, in the lens redshift bin, i, and the source redshift bin, j, respectively, and γ_+ is the tangential shear for the associated galaxy pair. Additionally, it is common to the define observed excess surface density profile or density contrast, $\widehat{\Delta\Sigma}(\theta)$, as follows (Bartelmann, 1995),

$$\widehat{\Delta\Sigma}^{(ij)}(\theta) \equiv \widehat{\overline{\Sigma}}^{(ij)}(<\theta) - \widehat{\overline{\Sigma}}^{(ij)}(\theta) = \hat{\gamma}_+^{(ij)}(\theta)\Sigma_{\rm crit}^{(ij)}, \tag{1.106}$$

where $\widehat{\overline{\Sigma}}(<\theta)$ and $\widehat{\overline{\Sigma}}(\theta)$ are the azimuthally averaged projected surface density distributions within a disk and within a ring of angular distance θ, respectively, and $\Sigma_{\rm crit}$ is the critical surface density given by

$$\Sigma_{\rm crit} = \frac{1}{4\pi G}\frac{d_{\rm OS}}{d_{\rm OL}\, d_{\rm LS}}, \tag{1.107}$$

where $d_{\rm OS}$, $d_{\rm OL}$ and $d_{\rm LS}$ are the observer-source, observer-lens plane, and lens plane-source angular diameter distances, respectively. $\widehat{\Delta\Sigma}(\theta)$ as an estimator has the advantage that it includes a correction to the lensing boost which makes the response of the data vector with cosmology easier to observe.

From $\hat{\gamma}_+(\theta)$, we can get the observed position-shape angular power spectrum $\hat{C}_{{\rm n}\varepsilon}^{(ij)}$, performing a Hankel transform as follows

$$\hat{C}_{{\rm n}\varepsilon}^{(ij)} = -2\pi \int_0^\infty {\rm d}\theta\, \theta\, \hat{\gamma}_+^{(ij)}(\theta) J_2(\ell\theta). \tag{1.108}$$

To relate the observed position-shape angular power spectrum, $C_{{\rm n}\varepsilon}^{(ij)}(\ell)$, to the underlying galaxy-shear angular power spectrum, $C_{{\rm g}\kappa}^{(ij)}(\ell)$, we must take into consideration the main systematics which affect the lens clustering and the weak lensing of the sources. As is the case for the galaxy clustering signal described in Section 1.3.1, magnification of the foreground lenses is a considerable systematic. At the same time, as with the cosmic shear signal discussed in Section 1.3.2, intrinsic alignments of source galaxies which are physically close enough to foreground

lenses to have a considerable correlation between their shapes and positions, respectively. Hence, the observed signal can be understood as

$$\hat{C}_{\mathrm{n}\varepsilon}^{(ij)}(\ell) = C_{\mathrm{g}\kappa}^{(ij)}(\ell) + C_{\mathrm{gI}}^{(ij)}(\ell) + C_{\mathrm{mG}}^{(ij)}(\ell), \tag{1.109}$$

where $C_{\mathrm{gI}}^{(ij)}(\ell)$ is the correlation between the intrinsic alignments in the observed source galaxy shapes and the positions of the lenses, and $C_{\mathrm{mG}}^{(ij)}(\ell)$ is the correlation of the magnification of the lenses and the shear of the background sources. Both quantities are discussed in more detail in Chapter 2.

1.3.4 The State-of-the-Art in Cosmology and Future Outlook

In recent years and decades, the field of cosmology has made large strides in advancing our understanding of the Universe at its largest scales, theoretically and observationally. Yet, many open questions still remain.

In the last few years, two particular questions have come to the forefront of attention in observational cosmology: the H_0 tension and the σ_8 tension/discrepancy. The former refers to the observed disagreement between late-Universe and early-Universe measurements of the Hubble constant, H_0, as defined in Equation (1.15). In the late Universe, the most recent observations of Type Ia supernovae are consistent with $H_0 = (73.04 \pm 1.04)$ km s^{-1} Mpc^{-1} (Riess et al., 2022). In the early Universe, the most recent observations of the cosmic microwave background imply $H_0 = (67.4 \pm 0.5)$ km s^{-1} Mpc^{-1} (Planck Collaboration et al., 2020). These measurements are discrepant with a statistical significance in excess of 5σ which implies a tension which only has a one in 35 million chance of being due to random noise. Currently, it is unclear whether this tension is caused by unaccounted systematics or whether it is a sign of inconsistencies within the ΛCDM model. Akin to this tension, another discrepancy between late and early Universe probes has been observed in the estimates of the σ_8 parameter (as defined in Equation (1.63)).

The σ_8 parameter is measured by large-scale structure probes, as it is an analog for the amplitude of the primordial matter power spectrum, A_{s}. As described in Section 1.3, galaxy surveys can put constraints on σ_8 as seen at low redshifts and

small scales through probing the matter power spectrum (see Sections 1.2.3.1 and 1.2.4). As shown in Figure 1.5, observations of the cosmic microwave background can also probe the matter power spectrum, but at higher redshifts and larger scales. Figure 1.6 shows the most recent constraints on the cosmological parameters which govern large-scale structure: $\Omega_{\rm m}$ and σ_8. It also shows the inferred values of the S_8 parameter, given by Equation (1.95), because cosmic shear surveys are not sensitive to $\Omega_{\rm m}$ or σ_8 in their own right, but rather to S_8. The galaxy surveys shown in Figure 1.6 are the three currently ongoing cosmic shear surveys: the Kilo-Degree Survey (KiDS), the Dark Energy Survey (DES) and the Hyper Suprime-Cam (HSC) survey. The top panel compares the constraints from the CMB to the constrains from cosmic shear (see Section 1.3.2) alone, while the bottom panel compares the same CMB constraints to the posterior contours from 3x2pt analysis (a combination of cosmic shear, galaxy clustering and galaxy-galaxy lensing measurements). It becomes apparent from Figure 1.6, that all current cosmic shear surveys observe a value of $S_8 \approx 0.75$ which is consistently lower by a margin of $\sim 2\sigma$ than the values of measured from CMB observations, e.g. $S_8 = 0.832 \pm 0.013$ from Planck Collaboration et al. (2020). Although the discrepancy between each individual galaxy survey constraint and the CMB constraints has a 5% chance of being caused by random noise, the fact that KiDS, DES and HSC all roughly agree on the discrepancy despite using different methodologies and galaxy measurements might be an indication of a possible physical tension in S_8 and σ_8.

The veracity of this possible tension might be addressed in the near future by upcoming galaxy surveys. Ongoing and future CMB surveys, such as the ground-based ACT (Thornton et al., 2016; Madhavacheril et al., 2023; Qu et al., 2023), CMB-Stage 4 (Abazajian et al., 2016) and the Simons Observatory (Ade et al., 2019), and the space-based LiteBIRD (Matsumura et al., 2014), will further improve the cosmological constraints from the early Universe. In tandem, *Euclid* (Laureijs et al., 2011), *Rubin* (LSST Science Collaboration et al., 2009) and *Roman* (Spergel et al., 2015), will surmount the previous constraints on S_8, as these surveys will be wider, deeper and more precise. However, wider and deeper observations of cos-

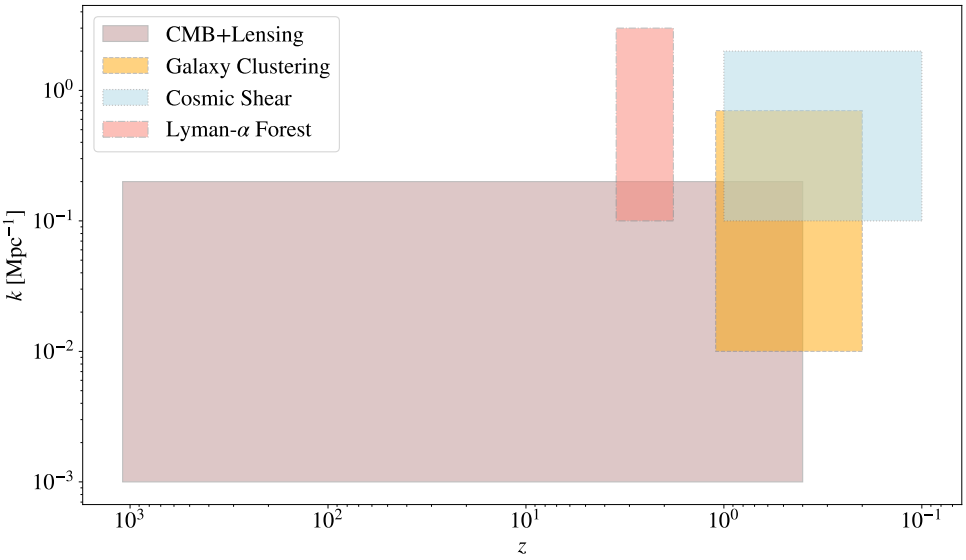

Figure 1.5: Overview of the scale/wavenumber, k, and redshift, z, dependence of the main observational probes of large-scale structure: the cosmic microwave background plus the gravitational lensing thereof (in brown with a solid outline), galaxy clustering (in orange with a dashed outline), cosmic shear (in blue with a dotted outline), and the Lyman-α forest (in red with a dot-dashed outline). Note that the edges of these intervals are not fixed, and they depend on the depth and precision of a given experiment. The limits of the CMB+Lensing observations are based on Planck Collaboration et al. (2020), the limits for galaxy clustering are based on eBOSS (Alam et al., 2021), the limits of cosmic shear are based on KiDS-1000, DES-Y3 and HSC-Y3 (Asgari et al., 2021; Amon et al., 2022; Li et al., 2023), and the limits for the Lyman-α forest measurements are also based on eBOSS (Alam et al., 2021). Figure made by the author.

mic shear, galaxy clustering and galaxy-galaxy lensing will make the modelling of the signal more challenging as non-linearities and observational systematics are no longer masked by random noise. This thesis aims to address this need by showcasing methods to model such systematics (see Chapters 2 and 3) and demonstrating the implementation of novel inference techniques which accurately propagate measurement uncertainties to cosmological parameter estimates even when the physical and systematic models can no longer be computed analytically (see Chapter 3).

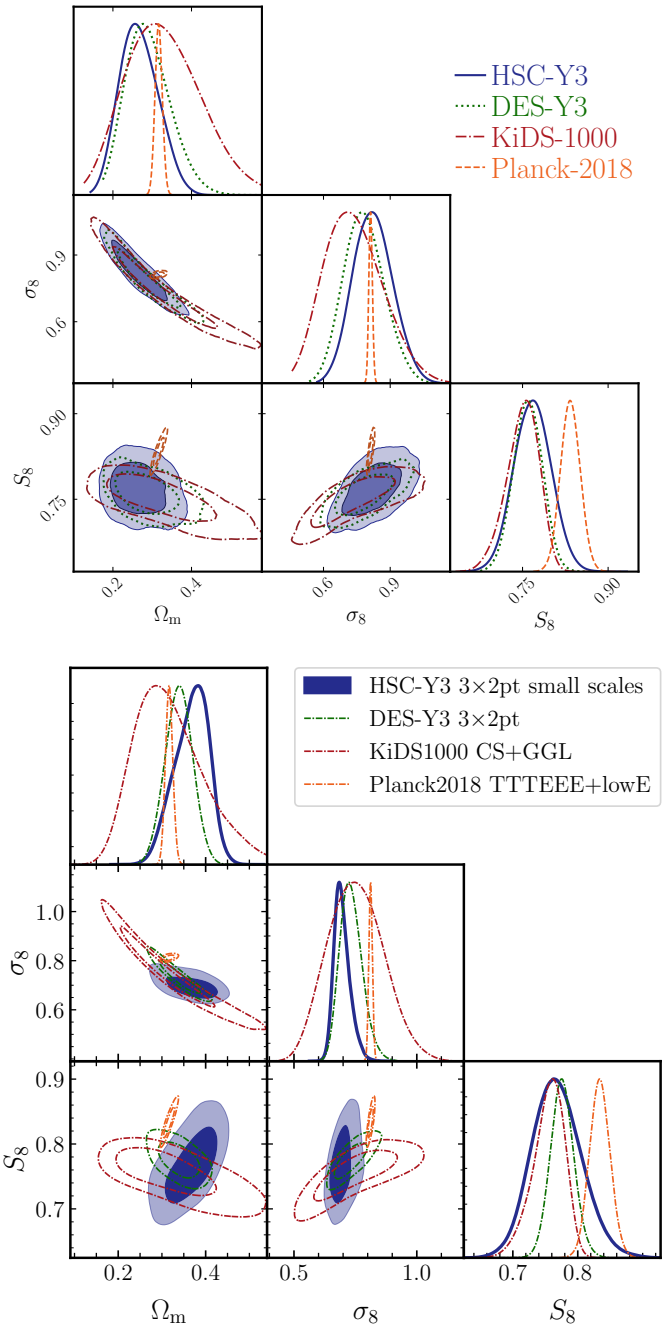

Figure 1.6: Posterior contours of the most recent cosmological constraints from galaxy surveys compared to constraints from the cosmic microwave background. The top figure only shows the cosmic shear results, while the bottom figure shows results from 3x2pt analysis (cosmic shear + galaxy clustering + galaxy-galaxy lensing). The Kilo-Degree Survey results (Asgari et al., 2021; Heymans et al., 2021) are shown with the red dot-dashed contours, the Dark Energy Survey results (Amon et al., 2022; Abbott et al., 2022) are given by the green dotted contours and the Hyper-Suprime Cam survey results (Li et al., 2023; Miyatake et al., 2023) are in blue solid contours. As a reference, the dashed orange contour shows the cosmological constraints from the cosmic microwave background as observed by the ESA Planck space telescope (Planck Collaboration et al., 2020). Figures from Li et al. (2023) and Miyatake et al. (2023).

1.4 Statistical Inference

In Sections 1.1 and 1.2, I have presented ΛCDM and its predictions for large-scale structure. From ΛCDM a set of free parameters, $\boldsymbol{\Theta}$, emerges which completely defines a ΛCDM cosmology: the Hubble constant, H_0, the energy density parameters of each of the Universe's ingredients, $\Omega_{\rm c}$, $\Omega_{\rm b}$, Ω_Λ, Ω_γ and Ω_ν, the amplitude of the primordial power spectrum, $A_{\rm s}$, and the scalar spectral index of the primordial power spectrum, $n_{\rm s}$. As described in Section 1.3, variations in these cosmological parameters cause measureable changes in the observables seen by galaxy surveys. Hence, when one makes measurements of such observables, one can infer a set of cosmological parameters which is most consistent with these observations. The rigorous study of this procedure while taking into consideration any uncertainties on the measurements is known as statistical inference.

1.4.1 Bayesian Inference

When conducting statistical inference, one can make a choice between two different philosophies: frequentist probability and Bayesian probability. The former assumes a notion of probability which is given by "the number of times the event occurs over the total number of trials, in the limit of an infinite series of equiprobable repetitions", while Bayesian statistics takes the view that probability is the "measure of the degree of belief about a proposition" (Trotta, 2008). Although both are valid in their own contexts, it can be shown that, in the presence of uncertainty, which is the case for cosmological observations of the Universe, Bayesian statistics is a generalisation of Boolean logic (Jaynes, 2003). As this thesis is concerned with uncertain observations from galaxy surveys which cannot be repeated for different sets of cosmological parameters as we only have access to one Universe, all cosmological inference is conducted in a Bayesian manner (see e.g. Jaynes 2003 for a detailed review).

The basis of Bayesian inference is Bayes' theorem which is given by

$$P(\boldsymbol{\Theta}|\boldsymbol{d},\mathrm{H}) = \frac{P(\boldsymbol{d}|\boldsymbol{\Theta},\mathrm{H})\,P(\boldsymbol{\Theta}|\mathrm{H})}{P(\boldsymbol{d}|\mathrm{H})}, \tag{1.110}$$

where $P(\boldsymbol{\Theta}|\boldsymbol{d},\mathrm{H})$ is the *posterior* probability as a function of the parameters, $\boldsymbol{\Theta}$, given some observations, $\boldsymbol{d}$, and a model/hypothesis, H; $P(\boldsymbol{d}|\boldsymbol{\Theta},\mathrm{H}) \equiv L(\boldsymbol{d}|\boldsymbol{\Theta},\mathrm{H})$ is known as the *likelihood* which is defined as the probability of observing a data vector, $\boldsymbol{d}$, given a set of parameters, $\boldsymbol{\Theta}$, within the hypothesis/model, H; $P(\boldsymbol{\Theta}|\mathrm{H})$, the *prior*, is the probability of parameters, $\boldsymbol{\Theta}$, given a model and any external information known prior to the observations; and $P(\boldsymbol{d}|\mathrm{H})$ is the *marginal evidence* which normalises the posterior probability.

As a consequence of Bayes' theorem, any observation yields a result which is a posterior probability distribution that is an update of our prior information based on the likelihood distribution obtained from the measured data. This often requires one to assume the form of the log-likelihood distribution, $\mathscr{L}(\boldsymbol{d}|\boldsymbol{\Theta})$, which is usually assumed to be Gaussian given by

$$\mathscr{L}(\boldsymbol{d}|\boldsymbol{\Theta}) \equiv \ln(L(\boldsymbol{d}|\boldsymbol{\Theta})) = -\frac{1}{2}\sum_{\alpha}^{|\boldsymbol{d}|}\sum_{\beta}^{|\boldsymbol{d}|}[d_\alpha - t_\alpha(\boldsymbol{\Theta})]\left(C_{\alpha\beta}\right)^{-1}[d_\beta - t_\beta(\boldsymbol{\Theta})], \quad (1.111)$$

where d_α is an element of the observed data vector, $\boldsymbol{d}$, $t_\alpha(\boldsymbol{\Theta})$ is the associated expected data vector for a given set of parameters $\boldsymbol{\Theta}$ in the model, and $C^{-1}_{\alpha\beta}$ is the inverse of the covariance matrix of the data which characterises the uncertainty on the data. Although the assumption of a Gaussian likelihood is common and often a good approximation to analytically intractable likelihoods, it is not always valid, even for two-point statistics in cosmic shear (Sellentin et al., 2018). In Section 1.4.3, I discuss in detail how the assumption of a Gaussian likelihood may be dropped and in Chapter 3 I present the first implementation of such an analysis on cosmic shear data.

1.4.2 Data Compression

When conducting statistical inference in cosmology, it is often the case that the measured data vector, $\boldsymbol{d}$, has a high dimensionality, even upon compression into summary statistics such as two-point correlation functions. Therefore, it can be computationally expensive to sample the likelihood if the data vector is too large.

In such cases, it is common to compress the data vector to an even lower dimensionality, while conserving as much information to constrain the cosmological model as possible. To measure the information, we can define the Fisher information matrix, F_{ij}, as

$$F_{ij}(\boldsymbol{d}_*, \boldsymbol{\Theta}_*) = -\left\langle \frac{\partial^2 \mathscr{L}(\boldsymbol{d}_*|\boldsymbol{\Theta}_*)}{\partial\theta_i \partial\theta_j} \right\rangle, \tag{1.112}$$

where $\theta_i, \theta_j \in \boldsymbol{\Theta}$ and $\boldsymbol{\Theta}_*$ is a fiducial set of cosmological parameters.

There exist many different compression schemes which reduce the dimensionality of a data vector while mostly conserving the Fisher information, $F_{ij}(\boldsymbol{d}_*, \boldsymbol{\Theta}_*)$. For example, principal component analysis (Pearson, 1901) where the inverse of the covariance matrix which constructs the likelihood is decomposed into its eigenvalues and eigenvectors, and only a fraction of the eigenvectors with the largest eigenvalues are not discarded, as they contain most of the Fisher information. A more sophisticated method for Gaussian likelihoods is the Massively Optimised Parameter Estimation and Data compression (Heavens et al., 2000b) which linearly compresses $\boldsymbol{d}$ down to the dimensionality of $\boldsymbol{\Theta}$, if the noise on $\boldsymbol{d}$ is independent of $\boldsymbol{\Theta}$. In this thesis, we use a generalisation of this scheme called score compression (Alsing & Wandelt, 2018) which only assumes that the likelihood is approximately Gaussian near the maximum (which is known to be the case for cosmic shear, Sellentin et al. 2018). To apply score compression, we Taylor-expand the log-likelihood, $\mathscr{L}$, up to the second order around the fiducial set of parameters, $\boldsymbol{\Theta}_*$, as follows

$$\mathscr{L}(\boldsymbol{\Theta}_* + \boldsymbol{\delta\Theta}) \approx \mathscr{L}(\boldsymbol{\Theta}_*) + \boldsymbol{\delta\Theta}^{\mathrm{T}} \boldsymbol{\nabla}\mathscr{L}(\boldsymbol{\Theta}_*) - \frac{1}{2}\boldsymbol{\delta\Theta}^{\mathrm{T}} \boldsymbol{\nabla}\boldsymbol{\nabla}^{\mathrm{T}} \mathscr{L}(\boldsymbol{\Theta}_*)\boldsymbol{\delta\Theta}, \tag{1.113}$$

where $\boldsymbol{\delta\Theta}$ is a small variation in the cosmological parameter vector near the fiducial value. The term $\boldsymbol{\nabla}\mathscr{L}(\boldsymbol{\Theta}_*)$ necessarily has the same dimensionality as $\boldsymbol{\Theta}$, so it can be viewed as a natural compression scheme which reduces the dimensionality of the data vector down to $|\boldsymbol{\Theta}|$. Therefore, it defines the score function as $\boldsymbol{s} \equiv \boldsymbol{\nabla}\mathscr{L}$ which

when evaluated at the fiducial cosmology gives the compressed data vector, $\boldsymbol{t}$, as

$$\boldsymbol{t} \equiv \boldsymbol{\nabla}\mathscr{L}_{*}. \tag{1.114}$$

1.4.3 Simulation-Based Inference

The assumption of a Gaussian likelihood as given in Equation (1.111) may not always hold. This may be due to non-Gaussian noise being inherent to a given summary statistic/measurement or due to the observations being biased by a systematic effect which induces non-Gaussian noise. To make sure that in such cases the assumption of a Gaussian likelihood is not biasing the inference as the non-Gaussian likelihood for some applications is unknown, it can be of interest to drop this assumption through the use of Simulation-Based Inference (SBI), also known as Likelihood-Free Inference or Implicit Likelihood Inference. SBI makes use of the fact that the posterior distribution is not only proportional to the likelihood, but also to the joint probability distribution of the data and the parameters of interest, $P(\boldsymbol{\Theta}, \boldsymbol{d}|\mathrm{H})$, such that

$$P(\boldsymbol{\Theta}|\boldsymbol{d}, \mathrm{H}) \propto P(\boldsymbol{\Theta}, \boldsymbol{d}|\mathrm{H})\, P(\boldsymbol{\Theta}|\mathrm{H}). \tag{1.115}$$

This implies that if one can define a mapping from parameters, $\boldsymbol{\Theta}$, to the associated data vector, $\boldsymbol{d}$, given a model, H, with a suite of forward-simulations (see Section 1.4.3.1), then one can numerically characterise the joint probability distribution, $P(\boldsymbol{\Theta}, \boldsymbol{d}|\mathrm{H})$, and directly evaluate the posterior distribution. Traditionally, this could be achieved through Approximate Bayesian Computation (Rubin, 1984; Pritchard et al., 1999; Beaumont, 2019) which rejects/accepts forward-simulations which contribute to the joint probability distribution according to their distance from the observed data vector. However, in this thesis, I opt for Density Estimation Likelihood-Free Inference (DELFI), as it has been shown to be more computationally effective, while also discarding less information (Alsing et al., 2019), which I discuss in detail in Section 1.4.3.2 and apply to cosmic shear data in Chapter 3.

1.4.3.1 Statistical Simulations

As alluded to in Section 1.2.4, it is necessary to create numerical simulations of large-scale structure under ΛCDM if one wishes to capture the maximal complexity of structure formation, particularly, in the non-linear regime. An accurate yet computationally expensive approach is the use of N-body and hydrodynamical simulations. N-body simulations simulate the evolution of many rigid and massive bodies evolving under the influence of gravity alone, while hydrodynamical simulations simulate the contents of the Universe as fluids, while also including the effects of electromagnetic interactions, star formation or the exchange of intergalactic gas (see e.g. Dolag et al. 2008; Nagamine 2018 for a review of such simulations). Such simulations require a prohibitively large amount of computational resources that it is unfeasible to obtain enough forward-simulations at different cosmological parameter values such that a prior volume in parameter space can be sampled extensively. This means that it is currently impractical to use N-body simulations or hydrodynamical simulations in order to conduct SBI from large-scale structure observables.

As a computationally effective but less accurate alternative, in this thesis, I develop a suite of statistical simulations of large-scale structure with the aim of forward-modelling galaxy survey observables, such as cosmic shear, galaxy clustering and galaxy-galaxy lensing, for a given set of cosmological parameters (see Section 3). Rather than simulating individual particles or fluids evolving within a ΛCDM Universe, statistical simulations directly simulate the statistical properties of the matter fields at different snapshots during the cosmic evolution of the Universe. The simplest form of such a statistical simulation of the matter overdensity, δ, is a Gaussian random field. These fields are exclusively defined by their mean and variance at every point in space, which can be randomly sampled in order to evaluate a specific instance of such a Gaussian random field, $\delta_{\rm G}$. However, Gaussian fields do not capture the full complexity of large-scale structure as by definition they have zero skewness or kurtosis or beyond. Realistic matter overdensity fields have higher-order variations as filamentary structure and voids tend to appear under

the influence of gravity and baryonic effects, even if the seeds that originated the overdensities were distributed in a Gaussian fashion. It has been found that a good approximation to such fields is a log-normally distributed random field, $\delta_{\log}$, which can be derived from a Gaussian random field. In fact, log-normal fields are even able to produce reasonable approximations to the three-point and four-point statistics measured from N-body simulations (Hall & Taylor, 2022; Piras et al., 2023). For this reason, log-normal random fields are a common approximation used for matter overdensity fields (Coles & Jones, 1991; Böhm et al., 2017; Abramo et al., 2016, 2022) as well as for convergence fields (Hilbert et al., 2011; Clerkin et al., 2017; Giocoli et al., 2017; Gatti et al., 2020).

To motivate the use of log-normal fields, it is helpful to note that primordial density fluctuations (see Section 1.2.1) occur due to random quantum fluctuations which are independent from each other. In the limit of a large number of random variations, due to the central limit theorem, the primordial density field is thought to be described by a Gaussian random field (Guth & Pi, 1982; Bardeen et al., 1983; Brandenberger, 1985; Barrow & Coles, 1990). As these fluctuations continue to evolve according to gravity and baryonic effects in a non-linear fashion, the matter density field deviates from Gaussianity. The simplest model for such fields arises when considering the central limit theorem for a random field of non-linear noise, where each random independent variation is multiplied with all others, rather than being linearly superposed (Coles & Jones, 1991). It can be shown that such non-linear noise produces a random field, which in its central limit, is described by a log-normal distribution (Aitchison & Brown, 1969). Besides this fact, log-normal distributions are motivated directly by perturbation theory (see Section 1.2.2). When considering the continuity equation for a small density perturbation as shown in Equation (1.43), it can be re-expressed as follows (Coles & Jones, 1991),

$$\frac{1}{\rho a^3}\frac{\mathrm{d}(\rho a^3)}{\mathrm{d}\tau} = -\nabla_{\boldsymbol{\Sigma}} \cdot \boldsymbol{u}, \tag{1.116}$$

where τ is the conformal time such that $a\,\mathrm{d}\tau = \mathrm{d}t$. Upon inspection of Equation (1.116), one finds that if the initial peculiar velocity field, $\boldsymbol{u}$, is Gaussian, $\nabla_{\boldsymbol{\Sigma}} \cdot \boldsymbol{u}$

is also Gaussian, and it will remain Gaussian as it linearly evolves with τ. In this case, given that the solution to Equation (1.116) gives that $\rho \propto \exp(-\nabla_{\boldsymbol{\Sigma}} \cdot \boldsymbol{u})$, the density field, ρ, is given by a log-normal field. This implies that, in the weak non-linear regime (see Section 1.2.4), a log-normal field describes the kinematics of large-scale structure in accordance with the continuity equation. Of course, such a description also has its limitations: as it cannot take into consideration three-dimensional topology of supergalactic gas clouds in the velocity field, Zel'dovich pancaking (Zel'dovich, 1970) cannot be reproduced by log-normal fields.

Additionally, it has also been observed empirically that the distribution of galaxies is approximately log-normal (Hubble, 1934; Coles & Jones, 1991; Wild et al., 2005), which makes log-normal fields a sufficiently good model for galaxy clustering. Although this could be driven by non-linear galaxy bias, it has also been found that log-normal distributions (with some generalisations) are a reasonable approximation for the weak lensing convergence fields in N-body simulations (Taruya et al., 2002; Hilbert et al., 2011; Das & Ostriker, 2006; Joachimi et al., 2011b; Takahashi et al., 2011; Hall & Taylor, 2022; Piras et al., 2023). This has been empirically confirmed as well in Clerkin et al. (2017) where it was found that log-normal fields accurately describe the weak lensing convergence between scales of 3 to 10 Mpc, which are the scales typically probed by weak lensing surveys.

For log-normal random fields to capture the statistics of a matter overdensity field at a given cosmology, we can set its power spectrum based on the matter power spectrum as described by ΛCDM in Section 1.2. For simplicity, it is convenient to project the matter field into a shell which is discrete along the line of sight and covers the whole sky in the angular directions, $\boldsymbol{\theta}$, such that

$$\langle \delta^{(i)}_{\log}(\boldsymbol{\theta}_1) \delta^{(j)}_{\log}(\boldsymbol{\theta}_2) \rangle = C^{(ij)}_{\delta\delta}(\theta = |\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2|), \tag{1.117}$$

where $C^{(ij)}_{\delta\delta}(\theta)$ is the two-dimensional two-point correlation function of the log-normal random field in real-space which is given by

$$C_{\delta\delta}^{(ij)}(\theta) = \sum_{\ell=0}^{\infty} \frac{\ell+1}{4\pi} C_{\delta\delta}^{(ij)}(\ell)\mathscr{P}_\ell(\cos\theta), \tag{1.118}$$

where $\mathscr{P}_\ell$ is the Legendre polynomial of ℓ^{th} order and $C_{\delta\delta}^{(ij)}(\ell)$ is the angular matter power spectrum given by the projection of the three-dimensional matter power spectrum into the discrete shell along the line of sight (see Chapter 3) for more details.

In practice, it is not simple to sample a log-normal random field from $C_{\delta\delta}^{(ij)}(\theta)$ directly, so it is more convenient to sample a Gaussian random field first, which is defined by

$$\langle \delta_{\rm G}^{(i)}(\boldsymbol{\theta}_1)\delta_{\rm G}^{(j)}(\boldsymbol{\theta}_2)\rangle = G_{\delta\delta}^{(ij)}(\theta = |\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2|), \tag{1.119}$$

where $G_{\delta\delta}^{(ij)}(\theta)$ is the two-dimensional power spectrum of the Gaussian random field in real-space. This can then be transformed into a log-normal random field through the following transformation (Coles & Jones, 1991; Kayo et al., 2001; Hilbert et al., 2011; Xavier et al., 2016; Tessore et al., 2023),

$$G_{\delta\delta}^{(ij)}(\theta) = \ln\left[1 + \frac{C_{\delta\delta}^{(ij)}(\theta)}{\alpha^{(i)}\alpha^{(j)}}\right], \tag{1.120}$$

where $\alpha^{(i)} = \langle \delta_{\rm log}^{(i)}(\boldsymbol{\theta})\rangle + \lambda$ and $\alpha^{(j)} = \langle \delta_{\rm log}^{(j)}(\boldsymbol{\theta})\rangle + \lambda$, and λ is the shift of the log-normal distribution, which is typically $\lambda = 1$ for matter overdensity fields to ensure that $\delta_{\rm log} = 0$ and $\delta_{\rm G} = 0$ happen at the same point $\boldsymbol{\theta}$. The Gaussian two-dimensional power spectrum can be expressed as an angular power spectrum through

$$G_{\delta\delta}^{(ij)}(\ell) = 2\pi \int_0^{2\pi} {\rm d}\theta\, G_{\delta\delta}^{(ij)}(\theta)\mathscr{P}_\ell(\cos\theta)\sin(\theta). \tag{1.121}$$

Once the Gaussian random field, $\delta_{\rm G}$, has been sampled from $G_{\delta\delta}^{(ij)}(\ell)$ in accordance with the input cosmology, it can then be transformed into a log-normal random field, $\delta_{\rm log}$, using the following relation

$$\delta_{\rm log}^{(i)}(\boldsymbol{\theta}) = e^{\delta_{\rm G}^{(i)}(\boldsymbol{\theta})} - \lambda. \tag{1.122}$$

This allows one to efficiently create statistical simulations of log-normal random matter fields which approximate the statistics of an N-body-simulated matter field reasonably well. Then, one can stack a series of such fields along the line of sight and populate the simulation with galaxies which trace the overdensities in the field. This ensures that the cosmological dependence is encoded into the galaxy clustering signal in the forward-simulations. Moreover, integrating the log-normal random matter field, $\delta_{\log}$, along the line of sight as given by Equation (1.88), we obtain the convergence field, κ, along the line of sight. From this, we can compute the shear and simulate the cosmic shear and galaxy-galaxy lensing signal in the galaxy population. A more detailed description of this is given in Chapter 3.

1.4.3.2 Neural Density Estimation

Once a mapping from cosmological parameters, $\boldsymbol{\Theta}$, to the compressed data vector, $\boldsymbol{t}$, has been established through a suite of forward-simulations, they can be used to learn the effective likelihood, $L(\boldsymbol{t}|\boldsymbol{\Theta})$, using Density Estimation Likelihood-Free Inference (DELFI; Alsing et al. 2019). DELFI uses neural networks to parametrise the effective likelihood distribution from the given parameter-data vector pairs. DELFI can use ensembles of multiple neural network architectures, such as Mixture Density Networks (MDNs; Bishop 1994), Gaussian Masked Autoencoders for Density Estimation (MADEs; Germain et al. 2015), and Masked Autoregressive Flows (MAFs; Papamakarios et al. 2017).

An MDN is a neural network in which the network weights, $\boldsymbol{w}$, parametrise a mixture density model of Gaussian distributions. Therefore, the learnt likelihood by a given MDN, i, is given by

$$P(\boldsymbol{t}|\boldsymbol{\Theta},\boldsymbol{w}) = \sum_{\alpha}^{N} A_{\alpha}(\boldsymbol{\Theta},\boldsymbol{w})\, G[\mu_{\alpha}(\boldsymbol{\Theta},\boldsymbol{w}), C_{\alpha}(\boldsymbol{\Theta},\boldsymbol{w})], \tag{1.123}$$

where N is the number of Gaussian, $A_{\alpha}(\boldsymbol{\Theta},\boldsymbol{w})$ is the amplitude of each Gaussian and $G[\mu_{\alpha}(\boldsymbol{\Theta},\boldsymbol{w}), C_{\alpha}(\boldsymbol{\Theta},\boldsymbol{w})]$ is a Gaussian distribution with mean $\mu_{\alpha}(\boldsymbol{\Theta},\boldsymbol{w})$ and covariance $C_{\alpha}(\boldsymbol{\Theta},\boldsymbol{w})$.

Instead, for the purposes of weak gravitational lensing observables, particu-

larly in high-dimensional parameter spaces, MADEs and MAFs perform better. MADEs harness the fact that any probability density distribution can be written as a product of one-dimensional conditionals through the chain rule. This implies that

$$P(\boldsymbol{t}|\boldsymbol{\Theta},\boldsymbol{w}) = \prod_{\alpha}^{|\boldsymbol{t}|} P(t_\alpha|\boldsymbol{t}_{1:\alpha-1},\boldsymbol{\Theta},\boldsymbol{w}). \tag{1.124}$$

From this, Gaussian MADEs choose to parametrise each $P(t_\alpha|\boldsymbol{t}_{1:\alpha-1},\boldsymbol{\Theta},\boldsymbol{w})$ as a Gaussian distribution whose mean and variance is learnt by a neural network which only depend on $\boldsymbol{t}_{1:\alpha-1},\boldsymbol{\Theta}$. Through this, MADEs learn the transform of the random variable, $\boldsymbol{t}$, to a unit Gaussian, $G(\boldsymbol{0},\boldsymbol{I})$, such that

$$u_\alpha = \frac{t_\alpha - \mu_\alpha(\boldsymbol{t}_{1:\alpha-1},\boldsymbol{\Theta},\boldsymbol{w})}{\sigma_\alpha(\boldsymbol{t}_{1:\alpha-1},\boldsymbol{\Theta},\boldsymbol{w})}, \tag{1.125}$$

where u_α is the transformed variable which the MADE learns for a given element of $\boldsymbol{t}$, α, $\mu_\alpha(\boldsymbol{t}_{1:\alpha-1},\boldsymbol{\Theta})$ is the learnt mean of the transformation, while $\sigma_\alpha(\boldsymbol{t}_{1:\alpha-1},\boldsymbol{\Theta})$ is the learnt variance. This means that Equation (1.124) reduces to

$$\begin{aligned} P(\boldsymbol{t}|\boldsymbol{\Theta},\boldsymbol{w}) &= G[\boldsymbol{u}(\boldsymbol{t},\boldsymbol{\Theta},\boldsymbol{w})|\boldsymbol{0},\boldsymbol{I}] \left|\frac{\partial \boldsymbol{u}(\boldsymbol{t},\boldsymbol{\Theta},\boldsymbol{w})}{\partial \boldsymbol{t}}\right| \\ &= G[\boldsymbol{u}(\boldsymbol{t},\boldsymbol{\Theta},\boldsymbol{w})|\boldsymbol{0},\boldsymbol{I}] \prod_{\alpha}^{|\boldsymbol{t}|} \sigma_\alpha(\boldsymbol{t},\boldsymbol{\Theta},\boldsymbol{w}). \end{aligned} \tag{1.126}$$

As found in Papamakarios et al. (2017) and Alsing et al. (2019), MADEs, on their own, are sensitive to changes in the order of the factors in Equation (1.124). Additionally, assuming only Gaussian conditional distributions may not be suitable for some applications. For this reason, for the purposes of this thesis, it is more useful to use MAFs which overcome these issues. Constructing a MAF involves initiating a stack of MADEs where the output transformation $\boldsymbol{u}$ of one autoencoder is taken as the input of the next MADE after being randomly reordered. Therefore, Equation (1.124) for a MAF with N_{MADE} MADEs becomes

$$P(\boldsymbol{t}|\boldsymbol{\Theta},\boldsymbol{w}) = G[\boldsymbol{u}_{\beta=N_{\mathrm{MADE}}}(\boldsymbol{t},\boldsymbol{\Theta},\boldsymbol{w})|\boldsymbol{0},\boldsymbol{I}] \prod_{\beta}^{N_{\mathrm{MADE}}} \prod_{\alpha}^{|\boldsymbol{t}|} \sigma_{\alpha\beta}(\boldsymbol{t},\boldsymbol{\Theta},\boldsymbol{w}), \tag{1.127}$$

where $\boldsymbol{u}_{\beta=N_{\mathrm{MADE}}}(\boldsymbol{t},\boldsymbol{\Theta},\boldsymbol{w})$ is the output of the last MADE in the MAF.

To ensure that the learnt effective likelihood, $P(\boldsymbol{t}|\boldsymbol{\Theta},\boldsymbol{w})$, converges to the true underlying likelihood distribution, DELFI optimises its neural networks based on the following loss function

$$-\ln[U(\boldsymbol{w})] = -\sum_{i=1}^{N_{\mathrm{sims}}} \ln[P(\boldsymbol{t}_i|\boldsymbol{\Theta}_i,\boldsymbol{w})], \tag{1.128}$$

where N_{sims} is the number of forward-simulations and the loss function, $-\ln[U(\boldsymbol{w})]$, is defined such that it is a Monte Carlo estimate of the Kullback-Leibler divergence between the learnt effective likelihood and the true effective likelihood (see Alsing et al. 2019 for a more detailed explanation). As is the case for the analysis shown in this thesis in Chapter 3, this process is usually repeated over an ensemble of many independent neural density estimator networks which each learn the effective likelihood separately. The final estimate of the effective likelihood is then given by the weighted sum of the learnt distributions from each neural density estimator.

Chapter 2

Magnification Bias in Galaxy Surveys with Complex Sample Selection Functions

Over the last few decades, weak gravitational lensing has become a powerful tool to directly measure the matter distribution of the late Universe, while allowing for the inference of the cosmological parameters which govern it. Surveys, such as the currently ongoing Kilo Degree Survey (KiDS, Kuijken et al. 2015), the Dark Energy Survey (DES, Flaugher & DES Collaboration 2013; Dark Energy Survey Collaboration et al. 2016), the Hyper Suprime-Cam Subaru Strategic Program (HSC SSP, Aihara et al. 2018), have become increasingly limited by systematics rather than statistics as ever-growing sample sizes reduce uncertainties. The impact of the systematics will become even more exaggerated for the next generation of surveys, e.g. *Euclid* (Laureijs et al., 2011), the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST, LSST Science Collaboration et al. 2009), and the Nancy Grace Roman Space Telescope (also known as WFIRST, Spergel et al. 2015). For this reason, recent efforts have focused on improving our physical understanding of often neglected phenomena which can influence cosmological parameter inference based on shear and clustering measurements. These effects include intrinsic galaxy alignments (Kiessling et al., 2015; Kirk et al., 2015; Troxel & Ishak, 2015) and magnification (Hildebrandt et al., 2009; Duncan et al., 2014; Hildebrandt, 2016; Unruh

et al., 2020; Thiele et al., 2020). In this chapter, we will focus on the magnification effects.

While the magnification due to gravitational lensing partially manifests itself as a change in the angular diameter of an object, it also changes the observed solid angle of a field with respect to the intrinsic solid angle. This can affect the observed galaxy counts and their fluxes, leading to magnification effects, which have been detected in the past by Chiu et al. (2016) and Garcia-Fernandez et al. (2018). It is important to note that this affects the counts of source galaxies and lens galaxies, such that the magnification due to large-scale structure can also change the shear-clustering cross-correlations (galaxy-galaxy lensing, GGL) and the clustering measurements (Hui et al., 2007; Ziour & Hui, 2008; Duncan et al., 2014; Unruh et al., 2020; Thiele et al., 2020). Therefore, if this effect is not accurately modelled in such analyses, a *magnification bias* can be induced. However, we also note that, in the literature and in this thesis, the term *magnification bias* is regularly used to refer magnification effects even when they are modelled.

We break down the magnification effect into two separate phenomena: *flux magnification* and *lensing dilution*. The first is caused by an increase/decrease in the flux observed from a source due to gravitational lensing which can push otherwise unobserved galaxies over the flux limit or push galaxies with magnitudes below the flux limit out of the observational window. At the same time, lensing dilution increases/decreases the number of observed sources within a certain area of the sky by (de-)magnifying the solid angle behind the gravitational lens. The magnification effect can be measured directly from changes in the apparent size and magnitude of lensed galaxies (Schmidt et al., 2012) or by comparing the observed galaxy effective radii to the intrinsic radii derived from their surface brightness and stellar velocity dispersion (Huff & Graves, 2014). Nonetheless, it is most commonly measured through the bias in the observed number density of sources (Scranton et al., 2005). Since this bias directly contributes to the clustering and GGL signal, we will rely on this approach in our analysis.

The constraining power of weak lensing samples is constantly growing (Troxel

& Ishak, 2015; Hikage et al., 2019; Asgari et al., 2021) by including additional measurements (Abbott et al., 2019b,a, 2022) and through joint analyses between different surveys like, for example, in the recent joint analysis of KiDS-1000 with BOSS (methodology described in Joachimi et al. 2021 and the results are shown in Heymans et al. 2021 and in Tröster et al. 2021). In all these analyses, the understanding of the systematics is becoming a priority. One potential systematic could appear from unaccounted magnification biases in the clustering signal of a non-flux-limited spectroscopic surveys such as BOSS (Dawson et al., 2013) or DESI (DESI Collaboration et al., 2016) or color-selected photometric samples such as DES REDMAGIC (Rozo et al., 2016) or luminous red galaxy (LRG) samples (Vakili et al., 2020; Fortuna et al., 2021). Thus also biasing the GGL correlations with shear signal from weak lensing surveys.

This work aims to provide a method for estimation of the magnification bias for surveys which have complex sample selection functions which are not purely flux/magnitude-limited. We use the standard framework for estimating the magnification bias from observables in flux-limited surveys as a basis for the parametrisation of a semi-empirical model for non-flux-limited surveys. This model is then tested by comparing the estimates for the magnification bias in BOSS observations (Dawson et al., 2013) to the estimates from MICE2 cosmological simulations. We then use our results to forecast some of the potential biases which could be induced in a joint analysis of KiDS-1000 or HSC Wide with BOSS and a *Euclid*-like survey with a DESI-like survey.

This chapter is structured in the following manner. In Section 2.1, the theoretical background is described. In Section 2.2, we provide an outline and presentation of our methods and simulations. The magnification bias estimates from a BOSS-like galaxy population are presented in Section 2.3. The forecasts for current and future joint analyses are found in Section 2.4. Lastly, we conclude the chapter and provide an outlook in Section 2.5. Appendix A repeats the analysis shown in Section 2.3 for a magnitude limited galaxy sample.

2.1 Theoretical Background

2.1.1 Magnification Bias for Flux-Limited Surveys

As described in the review by Bartelmann & Schneider (2001), a lensed population of galaxies with a cumulative galaxy count N at redshift z, given a flux limit of S, can be described in terms of the unlensed population, N_0, as

$$N(>S,z) = \frac{1}{\mu(z)} N_0\left(> \frac{S}{\mu(z)}, z \right), \tag{2.1}$$

where $\mu(z)$ is the magnification for a redshift z as defined in Equation (1.87) in Section 1.3.2. Here, the $1/\mu(z)$ factor accounts for the dilution of galaxies due to magnification. The unlensed population has been observationally shown to be similar to a power law in flux (in particular, for faint galaxies) given by

$$N_0(>S,z) = AS^{-\alpha} p_0(z;S), \tag{2.2}$$

where A and α parametrise the power law and $p_0(z;S)$ is the redshift probability distribution of the galaxies. To contextualise such a flux distribution, let us consider the simplest case: in Euclidean space, a sample of uniformly distributed galaxies with density $n_{\rm gal}$ which are all standard candles with the same luminosity, $L = 4\pi r^2 S$, where r is the Euclidean distance between a galaxy and the observer. For a given flux limit, this gives

$$N_0(>S) = n_{\rm gal}\frac{4}{3}\pi r_{\rm lim}^3 = n_{\rm gal}\frac{4}{3}\pi\left(\frac{4\pi S}{L}\right)^{-\frac{3}{2}}, \tag{2.3}$$

where $r_{\rm lim}$ is the distance at which the observed flux equals to the flux limit. From Equation (2.3), we find that for this idealised case $\alpha = 1.5$, so deviations in α from this value can be regarded as deviations in the underlying galaxy population from a uniformly distributed sample of standard candles.

Taking the ratio of these two populations, assuming that we can approximate the $\mu(z)$ with the magnification μ of a fiducial source at infinity (which should hold mainly at low redshifts, Bartelmann & Schneider, 2001) and integrating over

redshift, we get the following expression:

$$\frac{N(>S)}{N_0(>S)} = \mu^{\alpha-1}. \tag{2.4}$$

If $\alpha \approx 1$, we can see from Equation (2.4) that the magnification bias would vanish (with slight deviations from this depending on the redshift range). The magnification can be related directly to the local surface density κ in the weak lensing limit ($|\kappa| \ll 1$, $|\gamma| \ll 1$) with $\mu \approx 1+2\kappa$ (Broadhurst & Lehar, 1995). Therefore, one can relate κ to the relative difference between the magnified and the unmagnified galaxy populations and the exponent α of the flux power spectrum with

$$\frac{N(>S)-N_0(>S)}{N_0(>S)} \approx 2(\alpha_\kappa - 1)\kappa, \tag{2.5}$$

where α_κ is the same as the α in Equation (2.4) in the weak lensing limit. When analysing samples with a complex selection function, Equation (2.5) does not necessarily apply anymore. Nonetheless, we use the parameter α_κ as an analogue to estimate the magnitude of the magnification bias in a given galaxy sample.

2.1.2 Estimating the Magnification Bias in Flux-Limited Surveys

By considering Equations (2.1), (2.2), and the definition of magnitude as a function of flux, one can derive that $\alpha_{\rm obs}$ can be determined from the differential galaxy count $n(m)$ over a given band magnitude range from m to $m+dm$ as follows (Binggeli et al., 1988; Bartelmann & Schneider, 2001; Hildebrandt et al., 2009),

$$\alpha_{\rm obs}(m) = 2.5\frac{{\rm d}\log_{10}[n(m)]}{{\rm d}m}. \tag{2.6}$$

One could get the same estimates of $\alpha_{\rm obs}(m)$ (at least, for a flux-limited sample) by replacing $n(m)$ in Equation (2.6) with the cumulative galaxy count distribution. However, here we choose to derive $\alpha_{\rm obs}(m)$ from the differential distribution, $n(m)$, instead, because we find that it gives more robust estimates when deviating from the flux-limited case. Also, note that sometimes the differential galaxy count distri-

bution is given over flux, S, instead of magnitude, m. Then, $\alpha_{\rm obs}(S)+1$ is given by $-{\rm dlog}_{10}[n(S)]/{\rm dlog}_{10}(S)$.

This $\alpha_{\rm obs}$ near the faint end of the galaxy population is considered as an *effective* luminosity function slope if it is consistent with the α_κ value given by Equation (2.5). Therefore, by estimating the luminosity function slope, α_κ, through the observed α, one can estimate the systematic effects that may be introduced to galaxy number counts through the magnification bias, and therefore the systematics affecting the clustering and GGL signals derived from this observable.

2.1.3 Signal Modelling

In accordance with the framework outlined in Section 2 of Joachimi et al. (2021) as the methodology for the inference of cosmological parameters from KiDS-1000, we opt to quantify the influence of the magnification bias on cosmology through its contribution to the GGL angular power spectra. These angular power spectra are line-of-sight projections of the three-dimensional matter power spectrum. As shown in Section 1.3.3, we express the observable GGL angular power spectrum correlating galaxy positions, n, and galaxy shapes, ε, as a linear functional of derived statistics as

$$C_{{\rm n}\varepsilon}^{(ij)}(\ell) = C_{\rm gG}^{(ij)}(\ell) + C_{\rm gI}^{(ij)}(\ell) + C_{\rm mG}^{(ij)}(\ell), \tag{2.7}$$

where i is the index for lens galaxy redshift bins, j is the index of the source galaxy samples, gG stands for the cross-correlation between the lens galaxy distribution and the source gravitational shear, gI stands for the intrisinc alignment of source galaxies physically close to foreground lenses and mG stands for the correlation between gravitational shear and the lensing-induced magnification bias in the lens sample. $C_{\rm ga}^{(ij)}(\ell)$ for $a \in \{{\rm G},{\rm I}\}$ are defined as Limber-approximated line-of-sight projections of the three-dimensional cross-power spectrum between the galaxy and matter distribution, $P_{\rm gm}$, given by (Kaiser, 1992; LoVerde & Afshordi, 2008)

$$C_{\rm ga}^{(ij)}(\ell) = \int_0^{\chi_{\rm hor}} {\rm d}\chi \frac{n_{\rm L}^{(i)}(\chi) W_{\rm a}^{(j)}(\chi)}{f_K^2(\chi)} P_{\rm gm}\left(\frac{\ell+1/2}{f_{\rm K}(\chi)}, \chi\right), \tag{2.8}$$

where χ is the comoving distance, $\chi_{\rm hor}$ is the comoving distance to the horizon, $n_{\rm L}^{(i)}$ is the comoving distance distribution of the lens sample i and $f_{\rm K}$ is the comoving angular diameter distance. $W_{\rm G}^{(j)}$ is the weak lensing kernel and is given by

$$W_{\rm G}^{(j)}(\chi) = \frac{3H_0^2\Omega_{\rm m}}{2}\frac{f_{\rm K}(\chi)}{a(\chi)}\int_{\chi}^{\chi_{\rm hor}} {\rm d}\chi' n_{\rm S}^{(j)}(\chi')\frac{f_{\rm K}(\chi'-\chi)}{f_{\rm K}(\chi')}, \tag{2.9}$$

where H_0 is the Hubble constant, $\Omega_{\rm m}$ is the matter density parameter, c is the speed of light, a is the scale factor and $n_{\rm S}^{(j)}$ is the comoving distance distribution of the source sample j. $W_{\rm I}^{(j)}$ is the intrinsic alignment (IA) kernel. Here, we choose an IA kernel in accordance with the non-linear alignment model (NLA, Bridle & King 2007) given by

$$W_{\rm I}^{(j)}(\chi) = -A_{\rm IA}\frac{C_1\rho_{\rm cr}\Omega_{\rm m}}{D(a[\chi])}n_{\rm S}^{(j)}(\chi), \tag{2.10}$$

where $A_{\rm IA}$ is the IA amplitude, $z_{\rm pivot}$ is an arbitrary pivot which is set to 0.3 in line with previous IA analyses (Joachimi et al., 2011a), C_1 denotes a normalisation constant, $\rho_{\rm cr}$ is the critical density, D is the linear growth factor normalized to unity at the present day. We normalise the IA kernel by setting $C_1\rho_{\rm cr} \approx 0.0134$, i.e. $C_1 = 5\times10^{-14}(h^2M_\odot/Mpc^{-3})^{-2}$, in accordance with the value from Hirata & Seljak 2004 and Bridle & King 2007 which is set using the galaxy ellipticity measurements from SuperCOSMOS (Hambly et al., 2001; Brown et al., 2002).

The magnification term in Equation (2.7) is modelled as

$$C_{\rm mG}^{(ij)}(\ell) = 2(\alpha_{\rm obs}^{(i)} - 1)C_{\rm GG}^{(ij)}(\ell), \tag{2.11}$$

where i again indexes lens galaxy samples, j indexes source samples, mG stands for the lensing-induced magnification bias in the lens sample and GG stands for shear-shear correlation signal. $C_{\rm GG}^{(ij)}(\ell)$ is defined as the cosmic shear angular power spectrum purely from gravitational lensing effects, i.e. without any intrinsic alignment signals, and is given by

$$C_{\mathrm{GG}}^{(ij)}(\ell) = \int_0^{\chi_{\mathrm{hor}}} \mathrm{d}\chi \frac{W_{\mathrm{G}}^{(i)}(\chi) W_{\mathrm{G}}^{(j)}(\chi)}{f_K^2(\chi)} P_\delta \left(\frac{\ell + 1/2}{f_{\mathrm{K}}(\chi)}, z(\chi) \right), \tag{2.12}$$

where $P_{\mathrm{m,nl}}$ is the non-linear matter power spectrum. This power spectrum is computed with a non-perturbative model using HMCODE (Mead et al., 2015, 2016) integrated within CAMB (Code for Anisotropies in the Microwave Background; Lewis et al. 2000; Lewis & Bridle 2002; Howlett et al. 2012). HMCODE incorporates baryonic feedback in its halo modelling approach. We solely parametrise the baryonic feedback model using one free parameter, A_{bary}, in line with Hildebrandt et al. (2017). The non-linear matter power spectrum P_δ is also used to compute the cross-power spectrum between the galaxy and matter distribution P_{gm} used in Equation (2.8) as in the analysis shown in Joachimi et al. (2021).

2.2 Methodology

The method outlined in this work aims to provide an accurate estimate of the *effective* luminosity function slope, α, of a galaxy sample with a complex sample selection. This estimate can be used to quantify the magnification bias in clustering and GGL lensing analyses. To achieve this, we rely on realistic weak lensing simulations to calibrate the α_{obs} estimate from observables, based on Equation (2.6), such that it agrees with the value of α_κ derived from unobservable quantities using Equation (2.5). The procedure gives a magnitude range that yields the most optimal α_{obs} value. This value is used to estimate α_{obs} from observations. If the simulations are accurate, α_{obs} should agree with the underlying α_κ even though it cannot be directly measured. The code used for the analysis presented in this chapter can be found in the MAGBET[1] GitHub repository.

2.2.1 BOSS DR12 Data

We develop our method using lens samples derived from the Sloan Digital Sky Survey (SDSS)-III BOSS (Eisenstein et al., 2011; Dawson et al., 2013). BOSS is a spectroscopic survey with a complex sample selection function which is commonly

[1] https://github.com/mwiet/MAGBET

used in cosmological analyses of galaxy clustering and GGL (Alam et al., 2017; Sánchez et al., 2017; Beutler et al., 2017; Speagle et al., 2019; Tröster et al., 2020; Heymans et al., 2021). For more details about the nature of the galaxy selection process, see Alam et al. (2015). A lens galaxy sample selected in such a way could be introducing a substantial magnification bias in any analysis, while its complexity does not allow to measure it with current means. For the BOSS sample, the bias becomes even more important to model, because it is commonly used in GGL analysis with the source galaxy samples of weak lensing surveys whose footprint significantly overlaps with the BOSS footprint.

For this work, we use the photometric data from the final data release of BOSS, DR12 (Alam et al., 2015) with the same target selection as in Sánchez et al. (2017). This sample combines the BOSS LOWZ and CMASS galaxy samples to produce a catalogue which covers approximately 9,300 deg^2 (Reid et al., 2016). Its normalised redshift distribution can be seen in Figure 2.1. The sample is then split into two redshift ranges: "zlow" ($0.2 < z \leq 0.5$) and "zhigh" ($0.5 < z \leq 0.75$). From this photometric data, we use SDSS composite model (cmodel) band magnitudes which are defined in Stoughton et al. (2002).

2.2.2 MICE2 Simulations

For the analysis discussed in Section 2.3 and in Appendix A, we rely on datasets of simulated galaxies, selected from the MICE2 galaxy mock catalogue (Fosalba et al., 2015a,b; Crocce et al., 2015; Hoffmann et al., 2015). This catalogue is based on the MICE dark matter-only simulation, generated from 7×10^{10} particles in a box with a side length of 3 Gpc and assuming a ΛCDM cosmological model with $\Omega_{\rm m} = 0.25$, $\Omega_\Lambda = 0.75$, $\Omega_{\rm b} = 0.044$ and $h = 0.7$. A light cone, spanning $\sim 5{,}000$ deg^2, is constructed from this simulation box and populated with galaxies up to a redshift of $z = 1.4$ using a hybrid Halo Occupation Distribution (HOD) and Halo Abundance Matching (HAM) technique. Additionally, MICE2 embeds gravitational lensing by providing estimates of the shear components, convergence as well as true and lensed position for each galaxy. MICE2 derives weak lensing properties by constructing all-sky shells in steps of 35 Myr of lookback time (Fosalba et al., 2015b). These

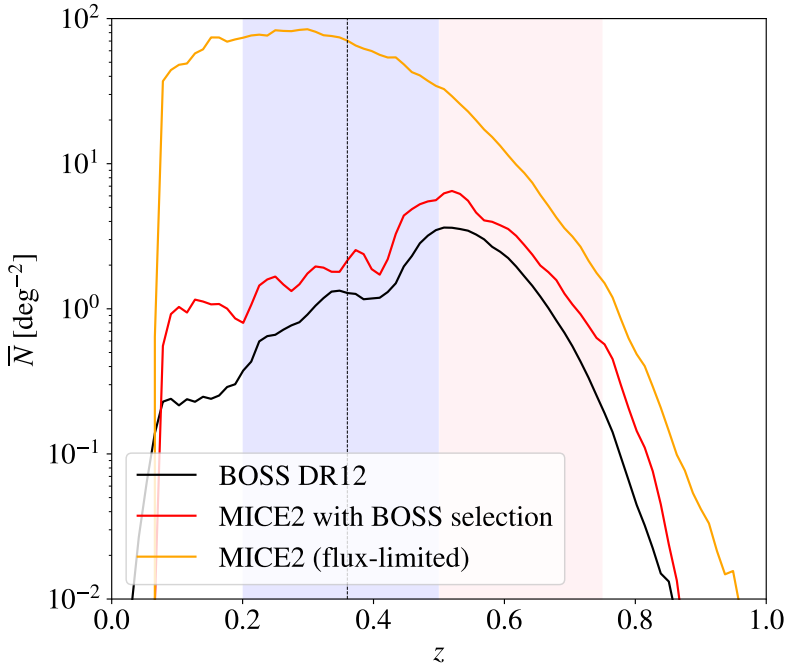

Figure 2.1: Galaxy counts per unit area on the sky, $\overline{N}$, for 100 redshift bins within $0 < z \leq 1$. The SDSS DR12 BOSS sample is shown in black, the MICE2 sample with the BOSS selection function in red and the flux-limited MICE2 sample in orange. The blue area marks the domain between $z = 0.2$ and $z \leq 0.5$ which defines the zlow bin, while the red area marks the domain of the zhigh bin ($0.5 < z \leq 0.75$). The dashed black horizontal line indicates the boundary between the LOWZ and the CMASS samples within the BOSS DR12 sample at $z \sim 0.36$.

are then projected into HEALPIX maps (Górski et al., 2005) from which the convergence is computed using the Born-approximation (Fosalba et al., 2015b). Therefore, galaxies within the same HEALPIX pixel inherit the same lensing properties which are, due to this limitation, accurate down to scales of 1 arcmin. We compute the magnified galaxy magnitudes according to Equation (2.13) which uses the weak lensing assumption $|\kappa| \ll 1$ by approximating $\mu \approx 1 + 2\kappa$.

We start from this MICE2 input catalogue and apply an evolutionary correction to the provided SDSS $g'r'i'$-band magnitudes and calculate an additional set of magnitudes

$$m^{mag} = m^{evo} - 2.5 \log_{10}(1 + 2\kappa)\,, \tag{2.13}$$

that factor in magnification, where m^{evo} are the evolution corrected MICE2 mag-

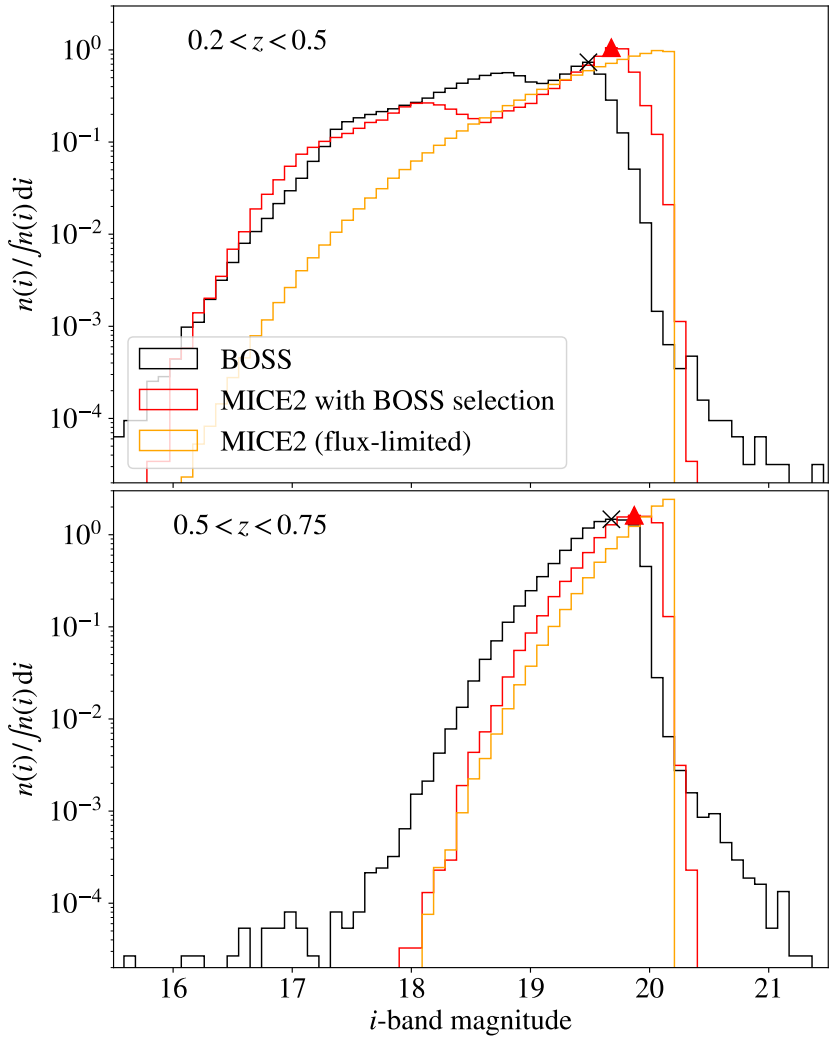

Figure 2.2: Normalised differential galaxy count distribution, $n(i)$, with respect to the i-band magnitude. The BOSS sample is shown in black, the MICE2 sample in red, while the flux-limited MICE2 sample is shown in yellow. In the top plot, we see the population of galaxies with $0.2 < z \leq 0.5$ and at the bottom, the population of galaxies with $0.5 < z \leq 0.75$. The black cross indicates the effective magnitude limit determined for the BOSS sample by finding the faintest prominent peak in $n(i)$. The red triangle indicates the same for the MICE2 mock sample.

nitudes and κ the convergence (see van den Busch et al. 2020 for details). This allows us later to separate the effects of lensing dilution and magnification in the mock data. Unruh et al. 2020 recently showed that the weak lensing assumption ($|\kappa| \ll 1, |\gamma| \ll 1$) used to derive Equation (2.5) and Equation (2.13) might lead to biases when simulating magnified galaxy samples. Since 99.9% of the galaxies in the MICE2 simulations have $|\kappa| < 0.09$, the assumption should still hold. However, it should be investigated in the future whether this is really the case.

Finally, we select two samples from this base catalogue, one with an arbitrary magnitude limit in the SDSS *i*-band at $m_i^{\mathrm{mag}} \leq 20.2$ (applied in Appendix A) and one that resembles the SDSS BOSS survey, using a target selection similar to Eisenstein et al. (2011) (applied in Section 2.3). The details of this BOSS mock sample selection are summarised in van den Busch et al. (2020).

The *i*-band number counts of these two samples and the original BOSS data is shown in Figure 2.2. In Figure 2.2, it becomes apparent how the BOSS selection function differs from a flux-limited sample. The cut-off of the galaxy population at the magnitude limit is not as pronounced, while the $n(i)$ no longer increases monotonically, especially at low redshifts. The galaxy counts per unit area as a function of redshift of the three samples is shown in Figure 2.1. Here we see how with the BOSS selection function applied, the redshift distribution is altered in a highly non-linear manner causing it to be multi-peaked with a main peak at $z \sim 0.5$. The magnitude-limited sample, on the other hand, follows a roughly single-peaked distribution dominated by low redshift galaxies ($z \sim 0.3$).

Having knowledge of the underlying matter distribution allows us to compare estimates of the scale of flux magnification through $\alpha_{\mathrm{obs}}(m)$ from observables as given by Equation (2.6) with the α_κ estimate as given by Equation (2.5). When analysing the MICE2 mock observations, we only consider the SDSS model *i*-band magnitude, due to a lack of available SDSS cmodel magnitudes from the simulations.

As a sanity check of our methods outlined in Section 2.2.3, we conduct an estimate of the magnification bias induced by a flux/magnitude-limited sample selec-

tion function on a galaxy survey over an eighth of the sky in Appendix A. For this, we use the MICE2 simulations to obtain the positions and magnitudes of galaxies before and after magnification, while knowing the true underlying matter density. We set the magnitude limit in the i-band to a magnitude of 20.2 (similar to the magnitude limit in the i-band of the BOSS survey). We find that the calibrated $\alpha_{\rm obs}$ values accurately recover α_κ near the faint limit. At the same time, the $\alpha_{\rm obs}$ estimates are robust over large changes in the calibration range chosen, showing that the power law approximation holds within $\sim 1\sigma$ over $\Delta i \sim 1$ and within $\sim 2\sigma$ over the whole magnitude range for both *zlow* and *zhigh*.

To conduct the analysis for the case where the target selection function is not flux or magnitude limited, we select a $\sim 5,000$ deg^2 area from the MICE2 simulations and apply the aforementioned sample selection function to it. The i-band magnitude distribution of the BOSS and MICE2 galaxies within each of the redshift bins is shown in Figure 2.2. Here, we see that, although the overall shape of the population is similar between the BOSS and the MICE2 galaxies, the MICE2 objects are consistently shifted towards the fainter end of the distribution. This is at least partially caused by the fact that the BOSS magnitudes are i-band cmodel magnitudes and the MICE2 magnitudes are SDSS model i-band magnitudes. In addition, the MICE2 simulations with a BOSS-like selection function do not seem to capture the population of galaxies at the extremes of the magnitude distribution. Both of these biases might also be due to some assumptions in the galaxy formation and evolution models used in the MICE2 simulations. In addition, the fiducial cosmology assumed for the simulations might not agree with the cosmological parameters preferred by the BOSS data. However, the method of calibrating the $\alpha_{\rm obs}$ estimates from the observations with the simulations is not sensitive to a constant shift in the distribution nor is it sensitive to the extremes of the magnitude distribution by construction.

2.2.3 Calibration Procedure on Simulations

To calibrate the $\alpha_{\rm obs}$ obtained from observations, we first have to determine an accurate estimate of the underlying luminosity function slope, α_κ, in the MICE2

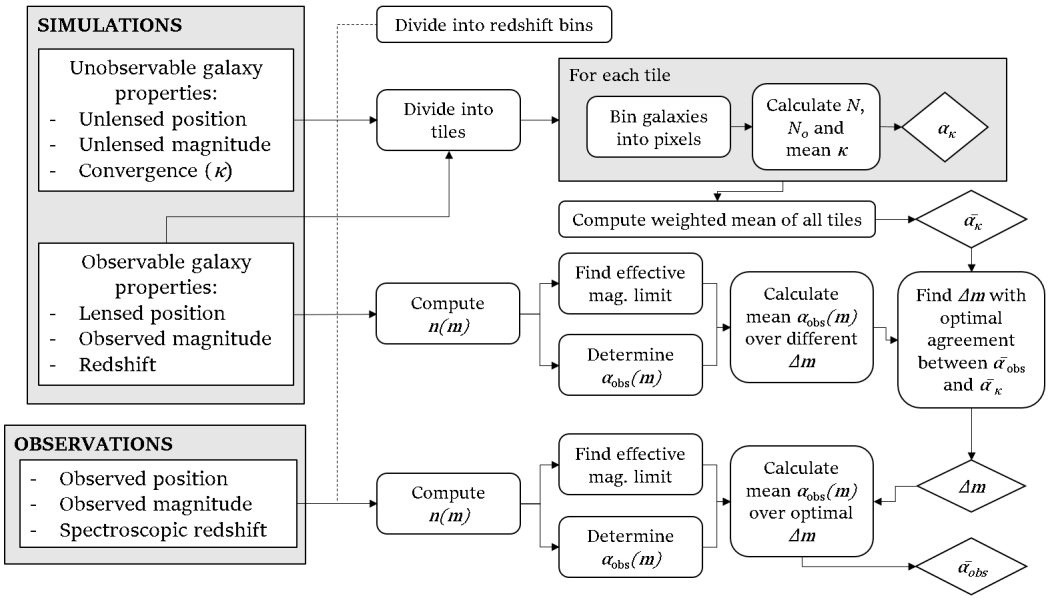

Figure 2.3: Flow chart outlining the method used to estimate the magnification bias of galaxy samples with an arbitrary sample selection. N stands for the count of lensed galaxies, N_0 refers to the counts of unlensed galaxies, κ to the convergence, α_κ to the luminosity function slope determined from the known κ, $n(m)$ is the differential galaxy count distribution over magnitude, m, $\alpha_{\rm obs}$ is the luminosity function slope as determined from $n(m)$.

simulations as given by Equation (2.5). As outlined in Figure 2.3, we first spatially bin the lensed and unlensed galaxy positions using HEALPIX at a resolution of nside = 64 (Górski et al., 2005). Within each bin/pixel, we evaluate lensed and unlensed cumulative galaxy counts, N and N_0 respectively, as well as the average convergence, κ. We then perform a least squares linear fit of the relative difference between lensed and unlensed galaxy counts over the convergence, κ, to estimate α_κ (as shown in Figure 2.4). This is a consequence of the linearity between these two quantities which emerges in the weak lensing limit as given by Equation (2.5).

In order to obtain better estimates of the uncertainties of α, the HEALPIX pixels are grouped into tiles (HEALPIX pixels with a resolution of nside = 4) for which we repeat the analysis independently each time. The weighted mean of these values obtained from each tile gives the final estimate for α_κ, $\overline{\alpha_\kappa}$, while the standard deviation between these values is used to estimate the uncertainty as given by

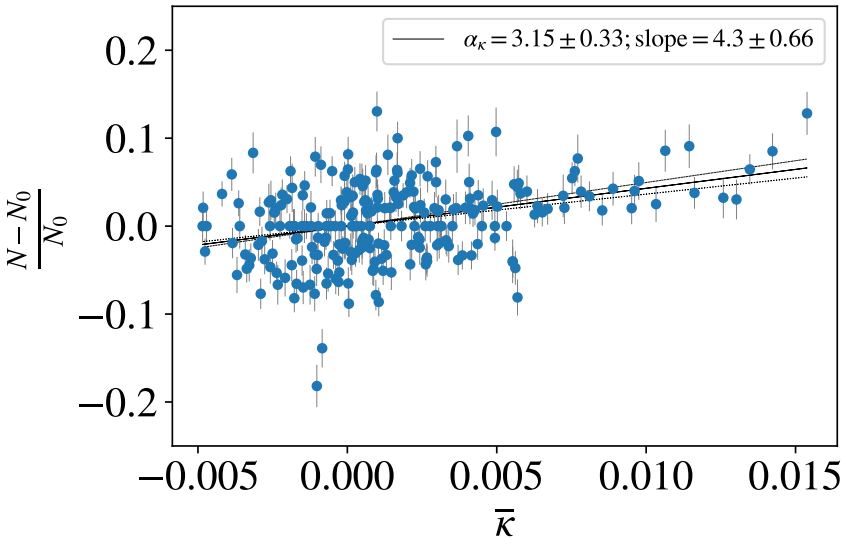

Figure 2.4: Plot of the relative difference in galaxy counts per pixel over the mean convergence ($\overline{\kappa}$) in each pixel for a HEALPIX pixelation with nside = 64 and in the zhigh redshift bin ($0.5 < z \leq 0.75$). The graph only shows pixels within 1 of 28 tiles. The relative difference between the lensed and unlensed galaxy counts in each pixel are shown as blue points. The black line is fitted to the blue data points with Equation (2.5) to give the α_κ value shown in the legend.

$$\sigma_\alpha^2 \equiv \frac{\sum_{i=1}^{M} \frac{(\alpha_i - \overline{\alpha})^2}{\sigma_i^2}}{(M-1)\sum_{i=1}^{M} \frac{1}{\sigma_i^2}}, \tag{2.14}$$

where α_i are the α estimates from each tile or bin, σ_i is their associated uncertainty, $\overline{\alpha}$ is the weighted mean of the α estimates and M is the number of tiles over which the analysis is repeated. When $\sigma_i = \sigma$, Equation (2.14) reduces to the equation for the error of the mean, i.e. $\sigma_\alpha \equiv \sigma_{sd}/\sqrt{M}$ where $\sigma_{sd} \equiv (\sum_{i=1}^{M}(\alpha_i - \overline{\alpha})^2/(M-1))^{1/2}$.

As an alternative, one might think that it would be enough to assume that the uncertainty on the galaxy counts is given by a noise, which considers the correlation between the lensed and unlensed galaxy counts (which is shown in the errorbars of the data points in Figure 2.4). We find, however, that this approach leads to underestimates of the uncertainties. Sampling α_κ over many different areas in the sky gives a more conservative estimate of the uncertainty, while also accounting for the local fluctuations in the BOSS sample.

A possible cause for concern when comparing the magnified and unmagnified galaxy populations can be the edge cases where, for a given bin or pixel, the unmagnified galaxy number count $N_0 = 0$, while the magnified number counts $N = 1$ or vice versa. These cases cause divergences in the relative difference and unrealistic uncertainties, since they introduce null denominators. For this reason, they are excluded in the analysis. In any case, the frequency of these occurrences is usually found to be negligible for the HEALPIX resolutions and redshift bins used in this work. Dividing the 5,000 deg^2 MICE2 simulations into two redshift bins at a HEALPIX nside = 64, there are none of these cases. While considering 19 redshift bins at the same HEALPIX resolution, only $\sim 0.7\%$ of the pixels have to be discarded.

2.2.4 Determining Magnification Bias from Observations

After having determined the luminosity function slope, $\overline{\alpha}_\kappa$, from the simulations as described in Section 2.2.3, we estimate the optimal magnitude range, Δm, to calibrate the estimate of $\alpha_{\rm obs}$ from mock observations using α_κ.

To do this, we first choose a magnitude band, m, that has been used to select (at least, partially) the galaxy sample of interest. Another magnitude band will carry less information about flux magnification. Then, we determine the discrete differential galaxy count distribution, $n(m)$, over the chosen magnitude, m, for a given redshift range. Subsequently, we find the magnitude at which the faintest most dominant peak in $n(m)$ occurs. This value is considered to be the effective magnitude limit of the galaxy sample. From $n(m)$, we compute $\alpha_{\rm obs}(m)$ using Equation (2.6). Thereafter, we calculate the weighted mean of $\alpha_{\rm obs}(m)$, $\overline{\alpha}_{\rm obs}$, over all possible magnitude ranges, Δm, below the effective magnitude limit determined before.

In order to find the optimal Δm which will be used for the calibration of $\overline{\alpha}_{\rm obs}$ from the actual observations, we find the value of $\overline{\alpha}_{\rm obs}(\Delta m)$ which is in best statistical agreement with the value of $\overline{\alpha}_\kappa$ determined previously for the same galaxy sample and redshift range. Therefore, the optimal $\overline{\alpha}_{\rm obs}(\Delta m)$ value is the one which minimises the number of standard deviations it deviates from $\overline{\alpha}_\kappa$, i.e. $|\overline{\alpha}_{\rm obs}(\Delta m) - \overline{\alpha}_\kappa| / \sigma_{\alpha_\kappa}$.

The reason behind choosing a magnitude range, Δm, relative to the effective magnitude limit of the differential galaxy count distribution, $n(m)$, for calibration is to account for one of the simplest forms of disagreement between the observed $n(m)$ and the $n(m)$ from mock observations. This disagreement being a constant shift in the domain of $n(m)$. For instance, such a shift exists between the $n(m)$ from the BOSS and MICE2 samples which has been discussed in Section 2.2.2 and shown in Figure 2.2 already. If we were to evaluate $\overline{\alpha}_{\rm obs}^{\rm MICE2}$ and $\overline{\alpha}_{\rm obs}^{\rm BOSS}$ over the same magnitude range, while disregarding the difference between their $n(m)$ distributions, the $\overline{\alpha}_{\rm obs}$ estimates will be biased. This happens because we would be probing regimes of $n(m)$ from the observed galaxy sample beyond or far below its magnitude limit when calculating $\overline{\alpha}_{\rm obs}$. Other higher-order biases in the $n(m)$ from mock observations may exist which would require more complex parametrisations of the calibration procedure. Nevertheless, in such cases, it might be more efficient and physically motivated to adjust the models used to produce the mock galaxy samples such that the agreement in $n(m)$ improves up to a point where it can be mostly parametrised by a constant shift in the magnitude.

In any case, once the optimal Δm to reconcile $\overline{\alpha}_{\rm obs}$ and $\overline{\alpha}_{\kappa}$ from the mocks has been determined, it may be used to calibrate $\overline{\alpha}_{\rm obs}$ from the observations. As summarised in the lower third of Figure 2.3, we first compute $n(m)$ for the given redshift range. We again find the faintest most dominant peak in $n(m)$ and set it as the effective magnitude limit and evaluate $\alpha_{\rm obs}(m)$ from $n(m)$. Lastly, we calculate the weighted mean of $\alpha_{\rm obs}(m)$ over the optimal magnitude range below the effective magnitude limit, Δm, determined before from the simulations over the same redshift range. Thus, we produce the final $\overline{\alpha}_{\rm obs}$ estimate for that sample.

2.3 Applications to BOSS Lenses

We proceed to apply the method described in Sections 2.2.3 and 2.2.4 to the BOSS lens galaxy sample introduced in Section 2.2.1. The magnitude bands selected for this are cmodel magnitudes, since they are better indicators of the overall flux emitted by a galaxy. The specific magnitude band chosen is based on which band was

used to select the dominant population within a sample. In other words, when working with LOWZ-dominated galaxy samples ($z < 0.36$), we use the r-band and when working with CMASS-dominated samples ($z > 0.36$), we use the i-band (Eisenstein et al., 2011). To allow for accurate forecasting of the KiDS-1000+BOSS analysis (Heymans et al., 2021), we choose the same convention for the redshift bins: $0.2 < z \leq 0.5$ and $0.5 < z \leq 0.75$. Consequently, both bins are dominated by CMASS galaxies, so we opt to use i-band magnitudes for the analysis of both samples.

As demonstrated in Appendix A, for the flux-limited case, we can accurately and robustly estimate the magnitude of the magnification bias by determining the effective luminosity function slope α through the weighted mean of $\alpha_{\rm obs}$ near the magnitude limit. In this section, we discuss whether the same can be said when applying a complex sample selection function which does not have a clear flux/magnitude limit such as in the case of the BOSS survey.

Firstly, we directly estimate α_κ from the MICE2 simulations following the approach outlined in Section 2.2.3. An example of this is shown in Figure 2.4, where we see the α_κ estimate within a single ~ 200 deg^2 tile containing 256 pixels within the zhigh bin. This procedure is repeated for each tile and redshift bin. Then, we find the weighted mean between the α_κ from each tile to determine the $\overline{\alpha}_\kappa$ for each redshift bin and its uncertainty given by Equation (2.14). This gives $\overline{\alpha}_\kappa^{\rm zlow} = 2.43 \pm 0.09$ and $\overline{\alpha}_\kappa^{\rm zhigh} = 3.26 \pm 0.07$.

Next, applying the procedure discussed in Section 2.2.4 and using the differential galaxy count distributions for each redshift bin shown in Figure 2.2, we can estimate $\overline{\alpha}_{\rm obs}$; once for the simulated BOSS-MICE2 observations, and once for the actual BOSS observations. In Figure 2.5, for zlow, we find that the estimate is optimal near the faint end of the count distribution, which is expected, since the assumed flux power law should be most accurate in the faint limit. However, this does not appear to be the case for the high redshift sample, zhigh. For this range, the estimate is optimal when considering the whole magnitude range up to the turn-off magnitude. This might be due to incompleteness in the sample and/or the complex

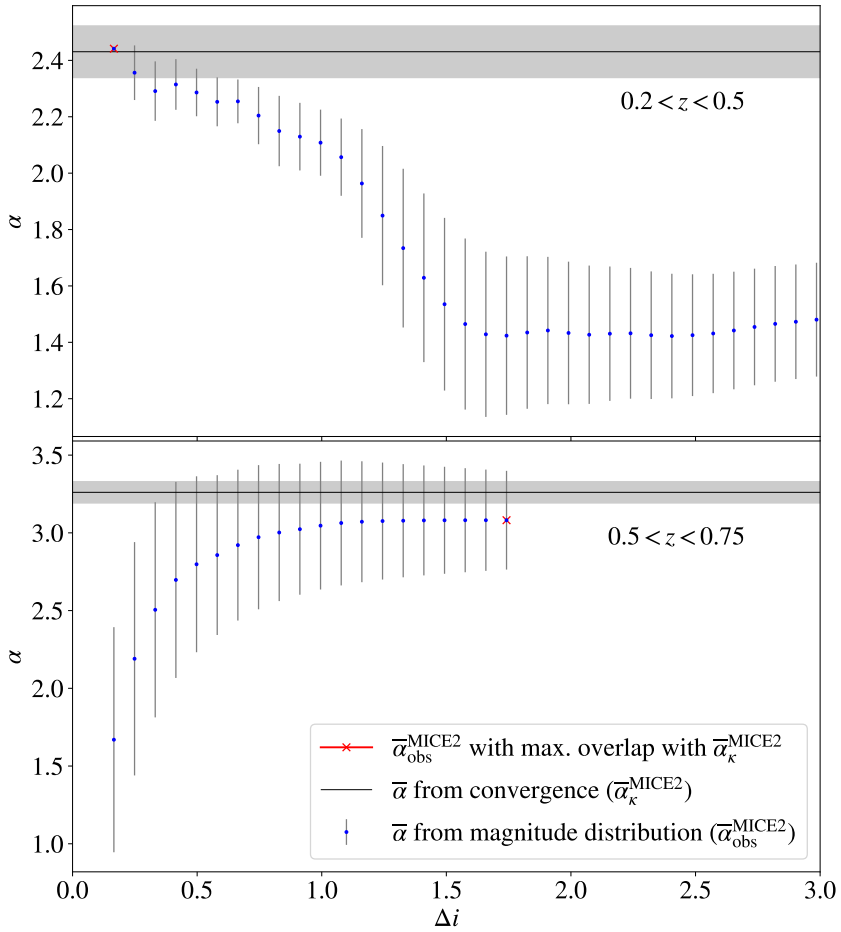

Figure 2.5: $\overline{\alpha}_{\rm obs}$ estimates from MICE2 simulations with the BOSS selection function over different i-band magnitude ranges below the turn-off magnitude, Δi, considered to calculate the weighted average. Two redshift ranges are considered: zlow with $0.2 < z \leq 0.5$ (top) and zhigh with $0.5 < z \leq 0.75$ (bottom). The red cross marks the α estimate which overlaps the most with the $\overline{\alpha}_{\kappa}$ truth from the simulations (black vertical line).

selection, which flattens the observed number counts (Hildebrandt, 2016).

Taking the magnitude range from the optimal $\overline{\alpha}_{\rm obs}^{\rm MICE2}$ estimate to calibrate $\overline{\alpha}_{\rm obs}^{\rm BOSS}$ gives the estimates shown in Figure 2.6. For the MICE2 mocks, we find that $\overline{\alpha}_{\kappa}^{\rm zlow} = 2.43 \pm 0.09$, while $\overline{\alpha}_{\rm obs}^{\rm zlow} = 2.442 \pm 0.002$. In addition, $\overline{\alpha}_{\kappa}^{\rm zhigh} = 3.26 \pm 0.07$, while $\overline{\alpha}_{\rm obs}^{\rm zhigh} = 3.08 \pm 0.32$ which indicates that the α estimates obtained from observations using Equation (2.6) are a good indicator of the scale of the magnification bias even when there is a complex sample selection function when they are properly calibrated. For this reason, we may consider the $\overline{\alpha}_{\rm obs}$ estimates given in Table 2.1 from the actual BOSS observations as unbiased indicators of the scale of the mag-

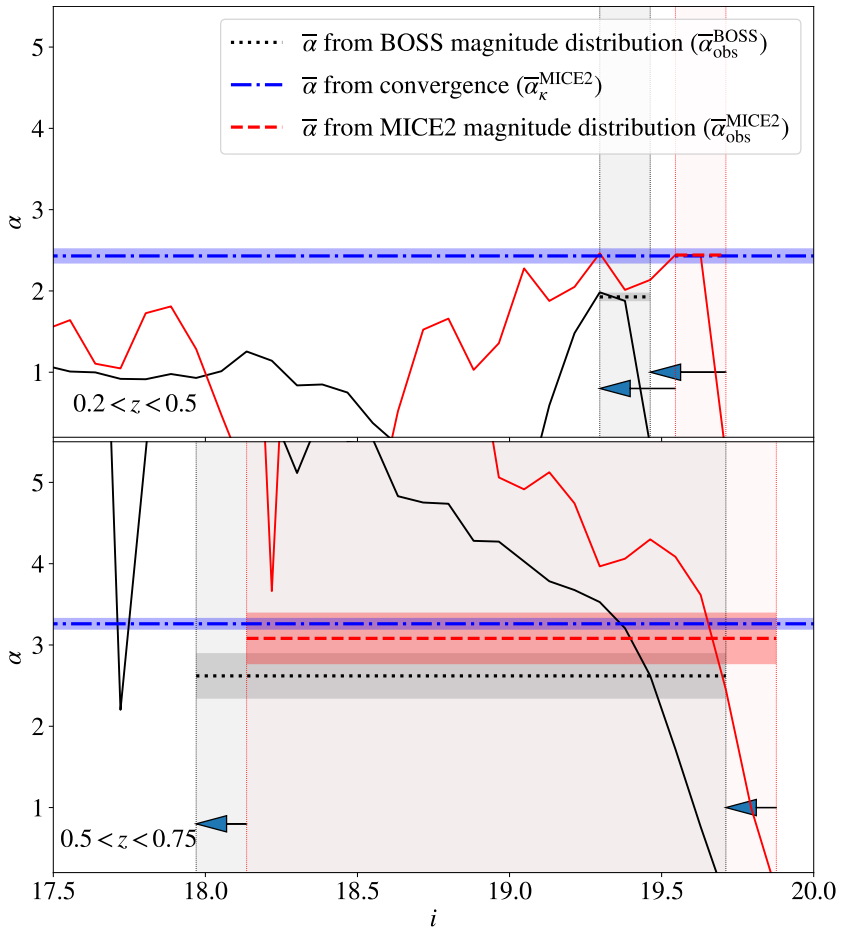

Figure 2.6: The slope of the BOSS luminosity function, α, as a function of the i-band magnitude (i) for two redshift bins: $0.2 < z \leq 0.5$ (top) and $0.5 < z \leq 0.75$ (bottom). The red line shows $\alpha_{\rm obs}(i)$ as given by Equation (2.6) calculated from the MICE2 mocks, while the black line shown $\alpha_{\rm obs}(i)$ as determined from the BOSS DR12 photometric data. The grey vertical lines mark the upper and the lower bounds of the magnitude range used to find $\overline{\alpha}_{\rm obs}^{\rm BOSS}$, while the red vertical lines mark the upper and lower bounds of the highlighted magnitude range used to determine $\overline{\alpha}_{\rm obs}^{\rm MICE2}$. The arrows indicate the constant magnitude shift applied to reconcile the differential galaxy count distribution, $n(m)$, from observations with the $n(m)$ from mocks. The dotted black horizontal line marks the $\overline{\alpha}_{\rm obs}$ estimate from BOSS galaxies, the dashed red horizontal line marks the $\overline{\alpha}_{\rm obs}$ estimate from MICE2 mock galaxies and the blue dot-dashed horizontal line marks the effective $\overline{\alpha}_{\kappa}^{\rm MICE2}$ determined from the weak lensing convergence with Equation (2.5) and used to calibrate $\overline{\alpha}_{\rm obs}^{\rm MICE2}$.

nification bias. Note that the value for zlow, slightly deviates from the value of $\overline{\alpha}_{\rm obs}^{\rm BOSS} = 1.80 \pm 0.15$ quoted in Joachimi et al. (2021), since there have been minor adjustments in the way peaks in $n(m)$ are detected. This leads to a 16% change in the amplitude of the mag. bias contribution, which has no effect on the KiDS-1000 analysis as the GGL contributions are marginal.

When comparing the $\alpha_{\rm obs}(i)$ curves for each bin in Figure 2.6, one might notice that the turn-off near the effective magnitude limit is not as steep for zhigh as for zlow. This is due to the complex BOSS selection function which deviates particularly strongly from a simple flux limit at high redshifts. Here is where the semi-empirical calibration of the magnitude range considered in order to determine the effective luminosity function slope $\overline{\alpha}_{\rm obs}$ is especially relevant. As shown in Figure 2.5, we find that for zhigh we get a more accurate α estimate when considering the entire magnitude range Δi available below the effective magnitude limit which is in stark contrast with the results found for a flux-limited sample (see Figure A.2). The opposite is the case for zlow. As shown in Figure 2.6, the double peak in the zlow bin combined with a clearer 'flux limit' near the peak magnitude means that the power law model for the luminosity function holds best within a small magnitude range near the peak. In other words, the Δi intervals which provide the best agreement between $\overline{\alpha}_{\rm obs}$ and $\overline{\alpha}_{\kappa}$ are also the magnitude intervals over which $n(m)$ resembles a power law the most. Therefore, our method actively avoids basing its estimates on a magnitude domain where the power law approximation in Equation (2.2) does not hold.

We note that in Figure 2.2 the simulated and the observed differential count distributions do not quite match. The $n(i)$ from MICE2 mock observations is shifted by a $\Delta i \approx 0.2$ to the faint end with respect to the BOSS $n(i)$. This might be due to some limitations in the galaxy model of the MICE2 simulations. The fact that the $n(m)$ from the mocks and observations do not match perfectly seems to be driving the discrepancy between $\overline{\alpha}_{\rm obs}^{\rm MICE2}$ and $\overline{\alpha}_{\rm obs}^{\rm BOSS}$ shown in Figure 2.6. However, since our calibration is based on a magnitude range of a fixed width relative to the effective magnitude limit for each sample, the estimates are not sensitive to this apparent

Table 2.1: Table showing the effective luminosity function slopes derived from Figure 2.6 for each redshift bin of the BOSS galaxy sample.

Bin	Redshift range	Luminosity function slope ($\overline{\alpha}_{\rm obs}^{\rm BOSS}$)
zlow	$0.2 < z \leq 0.5$	1.93 ± 0.05
zhigh	$0.5 < z \leq 0.75$	2.62 ± 0.28

shift in the domain of $n(i)$. The only thing which can bias our estimates are any disagreements in higher-order derivatives of $n(m)$ near the effective magnitude limit between observations and simulations. However, the uncertainties of $\overline{\alpha}_{\rm obs}$ from Equation (2.14) are defined such that they consider the variations of $\overline{\alpha}_{\rm obs}$ within the calibration magnitude range.

In addition, to see how α evolves over redshift within zlow and zhigh, we repeat this analysis of the BOSS sample again for a different choice of redshift bins producing the α estimates shown in Figure 2.7. Here, the edges of the 15 redshift bins are given by $\{0.2, 0.225, 0.25, 0.3, 0.4, 0.5, 0.525, 0.55, 0.575, 0.6, 0.625, 0.65, 0.675, 0.7, 0.725, 0.75\}$. Since the redshift bins between $z = 0.2$ and $z = 0.4$ are dominated by LOWZ galaxies, we choose a bin width of 0.1 instead of 0.025 between $z = 0.3$ and $z = 0.5$. This is done to mitigate the sharp gradient changes in $n(m)$ in the BOSS sample at redshifts near $z = 0.36$, i.e. at the boundary between the LOWZ and CMASS samples as shown in Figure 2.1.

Figure 2.7 shows how the effective luminosity function slope $\overline{\alpha}_{\kappa}$ in the MICE2 sample varies smoothly. Nonetheless, $\overline{\alpha}_{\rm obs}$ for MICE2 and for BOSS varies more strongly with redshift, due to their sensitivity of small variations in $n(m)$. Also, $\overline{\alpha}_{\kappa}^{\rm MICE2}$ is consistent with $\overline{\alpha}_{\rm obs}^{\rm MICE2}$ over most of the redshift range. However, for a few redshift bins, $\overline{\alpha}_{\rm obs}^{\rm MICE2}$ is in a $\sim 1\sigma$ to $\sim 2\sigma$ tension with $\overline{\alpha}_{\kappa}$ despite being calibrated to optimally overlap. Taking $\overline{\alpha}_{\kappa}$ as the underlying truth, we consider $\overline{\alpha}_{\rm obs}^{\rm MICE2}$ and $\overline{\alpha}_{\rm obs}^{\rm BOSS}$ to be biased in these cases. This seems to be driven by small discrepancies between the faint-end of $n(m)$ from MICE2 and the faint-end of $n(m)$ from BOSS. These are then exacerbated, since a small change in the sample size can lead to radical changes in the gradient of the magnitude distribution $n(m)$ of these galaxies, causing substantial biases in the $\overline{\alpha}_{\rm obs}$ estimates, as discussed in Hildebrandt (2016). Nonetheless, these discrepancies become insignificant as we increase the

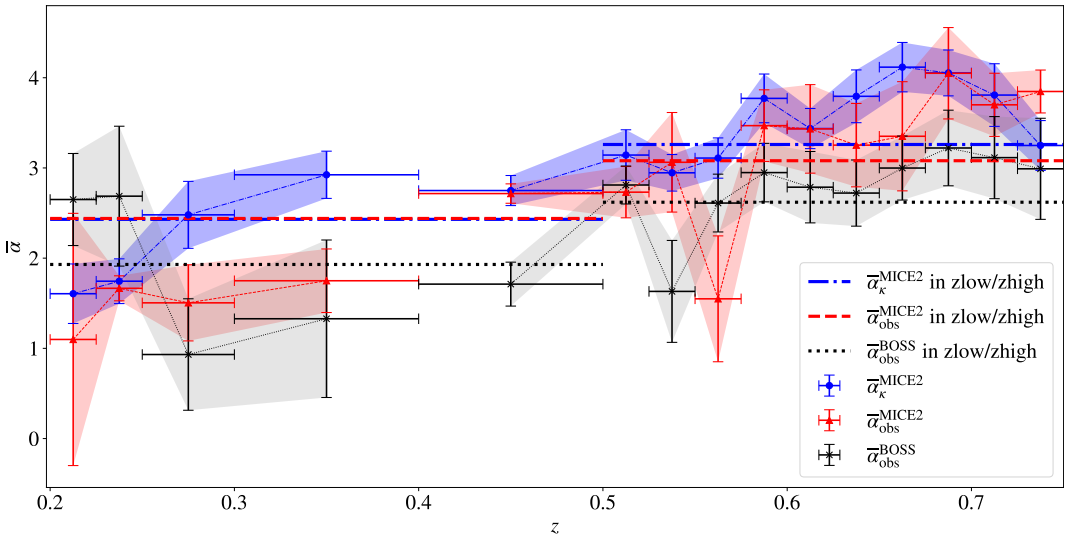

Figure 2.7: Plot of different α estimates for 16 different redshift (z) bins within $0.2 < z \leq 0.75$. The black crosses mark the $\overline{\alpha}_{\mathrm{obs}}^{\mathrm{BOSS}}$ estimates from observations within each bin, the red triangles mark $\overline{\alpha}_{\mathrm{obs}}^{\mathrm{MICE2}}$ estimates from mock observations and the blue circles mark the true effective $\overline{\alpha}_{\kappa}^{\mathrm{MICE2}}$ determined from the weak lensing convergence with Equation (2.5) and used to calibrate $\overline{\alpha}_{\mathrm{obs}}^{\mathrm{MICE2}}$.The values of $\overline{\alpha}_{\mathrm{obs}}^{\mathrm{MICE2}}$ and $\overline{\alpha}_{\mathrm{obs}}^{\mathrm{BOSS}}$ for the bins with $z < 0.4$ have been derived from the differential galaxy count distribution with respect to the r-band magnitude, $n(r)$, while the values for the bins with $z > 0.4$ have been derived from $n(i)$. The horizontal lines show the α estimates from simulations obtained for the zlow bin ($0.2 < z \leq 0.5$) and the zhigh bin ($0.5 < z \leq 0.75$).

sample size by widening the redshift bin width to the one used in the main analysis (i.e. $0.2 < z \leq 0.5$ and $0.5 < z \leq 0.75$). We also note that the $\overline{\alpha}_{\mathrm{obs}}^{\mathrm{BOSS}}$ estimates for the $0.2 < z \leq 0.225$ and $0.225 < z \leq 0.25$ bins may be biased. This is the case, since the profile of $n(m)$ as obtained from the MICE2 simulations deviates from the $n(m)$ observed in BOSS more strongly than over the remaining redshift range. Hence, the calibration range determined through our method does not necessarily apply anymore (as already mentioned in Section 2.2.4) and the estimates may be inaccurate. To avoid this, we highlight the necessity for accurate cosmological simulations over the whole redshift domain.

2.4 Magnification Bias in Weak Lensing Measurements

Having produced estimates for the effective luminosity function slope ($\overline{\alpha}_{\rm obs}$) for the BOSS DR12 galaxy sample, we now proceed to make forecasts of the importance of magnification bias in the GGL signals. The forecasts are produced from cross correlating source galaxies from weak lensing surveys with the BOSS lens samples considered in Section 2.3. First, we produce forecasts for the GGL signals for a KiDS-1000+BOSS DR12 analysis as described in Joachimi et al. (2021). Secondly, we produce similar forecasts for a GGL analysis of HSC Wide+BOSS DR12 similar to Speagle et al. (2019), while using the source bins described in Hikage et al. (2019). Lastly, we produce GGL forecasts for a potential *Euclid*-like+DESI-like analysis using the galaxy sample properties defined in the *Euclid* collaboration forecast choices (Euclid Collaboration et al., 2020), The properties of all of the aforementioned galaxy samples are given in Table 2.2 and their redshift distributions, $P(z)$, are given in Figure 2.8.

Throughout the forecasts, we assume a Planck 2018 TT,TE,EE+lowE flat ΛCDM cosmology (Planck Collaboration et al., 2020) with $\omega_{\rm b} = 0.02236$, $\omega_{\rm c} = 0.1202$, $h = 0.6727$, $n_{\rm s} = 0.9649$, $\ln(10^{10}A_{\rm s}) = 3.045$, $\Omega_{\kappa} = 0$, $w = -1$, and $\sum m_{\nu} = 0.06$ eV c^{-2}. To model the cross-power spectrum between galaxy and matter distribution ($P_{\rm gm}$, e.g. Section 2.1.3), we split the power spectrum into linear and a non-linear part as outlined in Joachimi et al. (2021) based on Sánchez et al. (2017) and set $b_1 = [2.1, 2.3]$, $b_2 = [0.2, 0.5]$, and $\gamma_3 = [0.9, 0.1]$ where the first value of each vector corresponds to the first lens bin (zlow) and the second values to the second lens bin (zhigh). b_1 is linear galaxy bias parameter, b_2 is the second-order galaxy bias parameter, and γ_3 is the non-local galaxy bias term. These values follow the rounded best-fit values from the cosmic shear and GGL analysis of KV450+BOSS (Tröster et al., 2020). We use the halo and intrinsic alignment models described in Section 2.1 and set $A_{\rm bary} = 3.13$ (upper limit of the KiDS-1000 prior) and $A_{\rm IA} = 0.8$ (best estimate from Tröster et al. 2020).

Table 2.2: Properties of the galaxy samples used to produce the galaxy-galaxy lensing forecasts.

Bin	z range	$\bar{z}$	$z_{\rm med}$	$n_{\rm gal}$	$\sigma_{\varepsilon,i}$
zlow	$0.2 < z_{\rm spec} \leq 0.5$	0.38	0.37	0.014	-
zhigh	$0.5 < z_{\rm spec} \leq 0.75$	0.60	0.55	0.016	-
KiDS1	$0.1 < z_{\rm phot} \leq 0.3$	0.26	0.21	0.62	0.27
KiDS2	$0.3 < z_{\rm phot} \leq 0.5$	0.40	0.36	1.18	0.26
KiDS3	$0.5 < z_{\rm phot} \leq 0.7$	0.56	0.54	1.85	0.27
KiDS4	$0.7 < z_{\rm phot} \leq 0.9$	0.79	0.75	1.26	0.25
KiDS5	$0.9 < z_{\rm phot} \leq 1.2$	0.98	0.93	1.31	0.27
HSC1	$0.3 < z_{\rm phot} \leq 0.6$	0.61	0.45	5.5	0.28
HSC2	$0.6 < z_{\rm phot} \leq 0.9$	0.78	0.72	5.5	0.28
HSC3	$0.9 < z_{\rm phot} \leq 1.2$	1.09	1.01	4.2	0.29
HSC4	$1.2 < z_{\rm phot} \leq 1.5$	1.37	1.30	2.4	0.29
Euclid1	$0.001 < z_{\rm phot} \leq 0.418$	0.33	0.21	3.0	0.21
Euclid2	$0.418 < z_{\rm phot} \leq 0.560$	0.51	0.49	3.0	0.21
Euclid3	$0.560 < z_{\rm phot} \leq 0.678$	0.63	0.62	3.0	0.21
Euclid4	$0.678 < z_{\rm phot} \leq 0.789$	0.75	0.73	3.0	0.21
Euclid5	$0.789 < z_{\rm phot} \leq 0.900$	0.85	0.84	3.0	0.21
Euclid6	$0.900 < z_{\rm phot} \leq 1.019$	0.96	0.96	3.0	0.21
Euclid7	$1.019 < z_{\rm phot} \leq 1.155$	1.09	1.09	3.0	0.21
Euclid8	$1.155 < z_{\rm phot} \leq 1.324$	1.23	1.24	3.0	0.21
Euclid9	$1.324 < z_{\rm phot} \leq 1.576$	1.42	1.45	3.0	0.21
Euclid10	$1.576 < z_{\rm phot} \leq 2.500$	1.85	2.04	3.0	0.21

Notes. $\bar{z}$ stands for the mean redshift in each tomographic bin, $z_{\rm med}$ for the median redshift, $n_{\rm gal}$ for the galaxy number density in arcmin^{-2} following the definition from Heymans et al. (2013) and $\sigma_{\varepsilon,i}$ for the dispersion per ellipticity component. zlow and zhigh are the lens bins based on the BOSS DR12 galaxy clustering data. The KiDS source bins have been defined in accordance with the methodology for the KiDS-1000 GGL analysis as given in Joachimi et al. (2021) and Heymans et al. (2021) based on the redshift calibration described in Hildebrandt et al. (2021) and Wright et al. (2020a). The properties of the HSC source bins are based on the information provided in Table 1 of the HSC Y1 cosmic shear analysis (Hikage et al., 2019) and the source $P(z)$ distributions are based on the DEmP photometric redshifts. The tomographic bins for the Euclid forecasts are in accordance with the *Euclid* collaboration forecast choices (Euclid Collaboration et al., 2020). The *Euclid* $P(z)$ distributions are determined using the fitting formula from Joachimi & Bridle (2010) assuming equi-populated binning with an overall median redshift of 0.8.

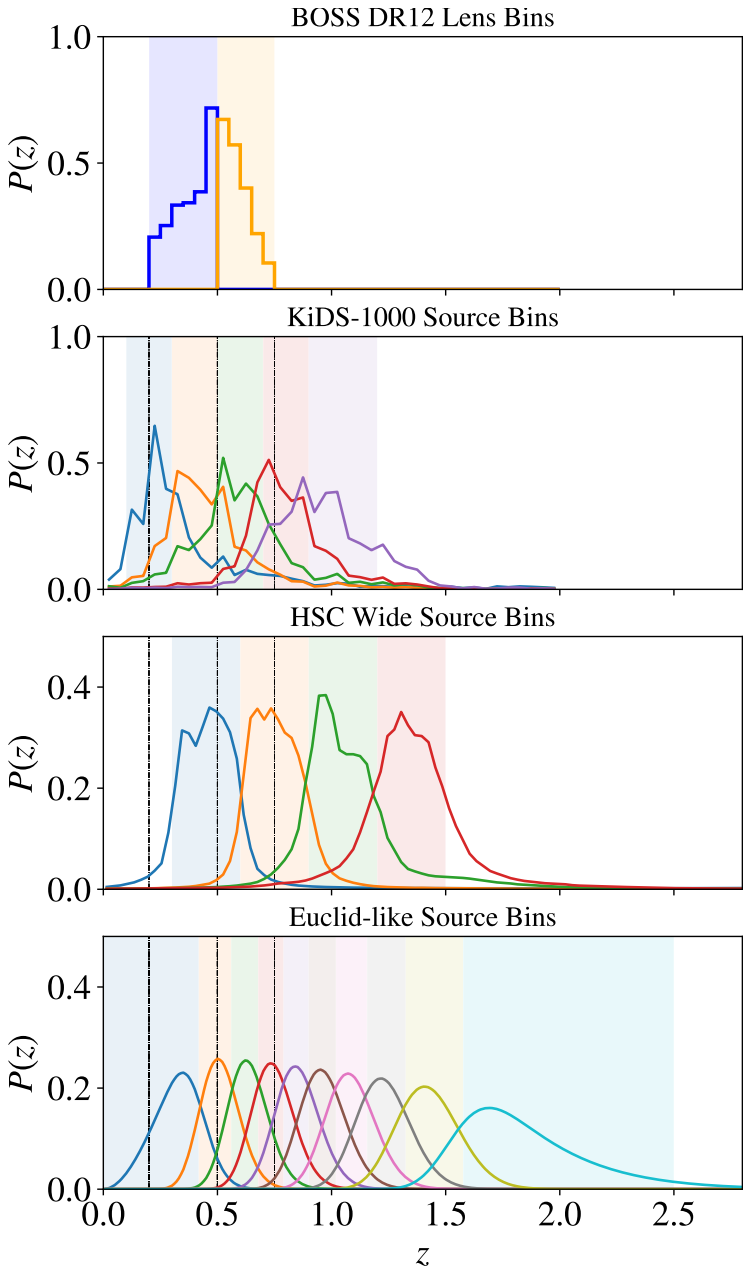

Figure 2.8: Redshift distributions $P(z)$ for the lens and source galaxy samples used in the forecasts for the galaxy-galaxy lensing signal in a KiDS-1000+BOSS, HSC Wide+BOSS and *Euclid*-like+DESI-like analysis. The properties of these redshift distributions are given in Table 2.2.The black vertical dot-dashed lines show the limits of the BOSS lens bins for comparison with the source bins.

2.4.1 KiDS-1000 + BOSS Forecasts

Following the approach outlined in Section 2.1.3, we propagate the $\overline{\alpha}_{\rm obs}$ measurements for zlow and zhigh shown in Table 2.1 into angular power spectrum prediction for the galaxy-galaxy lensing signal. We then determine the ratio between the angular power spectrum correlating gravitational shear with the lensing-induced magnification bias in the lens sample, $C^{(\rm ij)}_{\rm mG}(\ell)$, and the angular power spectrum correlating the lens galaxy distribution and the source gravitational shear, $C^{(\rm ij)}_{\rm gG}(\ell)$, as shown in Figure 2.9.

In order to put these contributions into perspective, we also estimate the statistical uncertainty in the GGL signal assuming shot and shape noise only (see for example Joachimi & Bridle, 2010). We calculate this for 6 logarithmically spaced ℓ bins per dex, while assuming the footprint area of the full KiDS survey, $A = 1,350$ deg^2. In Figure 2.9, we then compare the relative magnification-shear signal to the relative GGL uncertainty for each ℓ bin. The magnification-shear correlation found between these bins constitutes a few-per cent contribution to the galaxy-galaxy lensing signal correlated with the zlow bin. To compare that to the shape and shot noise, $\sigma_{\rm gG}$, we define the cumulative signal-to-noise ratio, SNR, within a range of angular scale, $\ell_{\rm min} < \ell < \ell_{\rm max}$, as follows

$$\begin{aligned}\mathrm{SNR}(\ell_{\rm min} < \ell < \ell_{\rm max}) &= \left(\frac{1}{K}\sum_{\rm i}^{K}\mathrm{SNR}_{\rm i}^2\right)^{1/2} \\ &= \left(\frac{1}{K}\sum_{\rm i}^{K}\frac{C^2_{\rm mG}(\ell_{\rm min,i} < \ell < \ell_{\rm max,i})}{\sigma^2_{\rm gG}(\ell_{\rm min,i} < \ell < \ell_{\rm max,i})}\right)^{1/2}, \quad (2.15)\end{aligned}$$

where K is the number of ℓ bins, i labels each ℓ bin, and $\ell_{\rm min,i}$ and $\ell_{\rm max,i}$ mark the lower and upper limits of each bin, respectively. For the correlations with the zlow bin, this implies a cumulative signal-to-noise ratio for $100 < \ell < 4600$ between 0.1 and 0.3. This contribution becomes larger for the high-redshift source bin (zhigh), from $\sim 5\%$ to $\sim 20\%$ of the GGL signal, while the shot and shape

noise is of a similar scale. Hence, the cumulative SNR($100 < \ell < 4600$) = 0.2 for the correlation between the zhigh and the first KiDS redshift bin, while the cumulative SNR($100 < \ell < 4600$) = 1.1 between the zhigh and the fifth KiDS bin. At the same time, these α values lead to a maximal contribution of the magnification bias to the clustering signal of $\sim 0.6\%$ (Joachimi et al., 2021). Even though we are assuming the area of the full 1,350 deg^2 KiDS footprint, these contributions to the GGL signal by magnification are large enough to prompt the consideration through modelling in the analysis of this systematic in the KiDS-1000+BOSS analysis outlined in Joachimi et al. (2021). Nonetheless, since the analysis shown here already provides an accurate estimate for the magnitude of the magnification bias, the contribution to the GGL signal in each bin can simply be fixed and added to the overall GGL angular power spectrum without the need to add any more free parameters in the astrophysical models.

We note the oscillations at low ℓ for some of the zhigh correlations in Figure 2.9. These originate from fluctuations from a power law of $< 1\%$ in $C_{\mathrm{gG}}(\ell)$ for $100 < \ell < 500$. They can be attributed to BAO as their amplitude decreases with ω_{b}. In Figure 2.9 and subsequent forecasts discussed Sections 2.4.2 and 2.4.3, the fluctuations in $C_{\mathrm{mG}}(\ell)/C_{\mathrm{gG}}(\ell)$ at low ℓ appear to be increased to amplitudes $> 1\%$ from the mean. This is caused by the non-BAO signal in $C_{\mathrm{gG}}(\ell)$ being approximately proportional to $C_{\mathrm{mG}}(\ell)$ at low ℓ. Hence, after taking their ratio, the only signal that does not approximately cancel is the BAO signal in $C_{\mathrm{gG}}(\ell)$. In any case, the variations in $C_{\mathrm{gG}}(\ell)$ are well below the uncertainties over that range (which are typically $\gg 2\%$), so they would be undetectable for now.

2.4.2 HSC Wide + BOSS Forecasts

We repeat the analysis for Section 2.4.1, considering the HSC Wide source bins. Figure 2.10 shows the ratio between $C_{\mathrm{mG}}^{(\mathrm{i,j})}(\ell)$ and $C_{\mathrm{gG}}^{(\mathrm{i,j})}(\ell)$ together with the relative uncertainty in the GGL signal for each ℓ bin assuming a full footprint area of 1,400 deg^2 (Aihara et al., 2018) as well as the galaxy sample properties shown in Table 2.2. Similar to KiDS, we find that the magnification-shear signal only contributes about $\sim 2\%$ to the GGL signal correlated with the zlow lens bin (giving a cumulative SNR

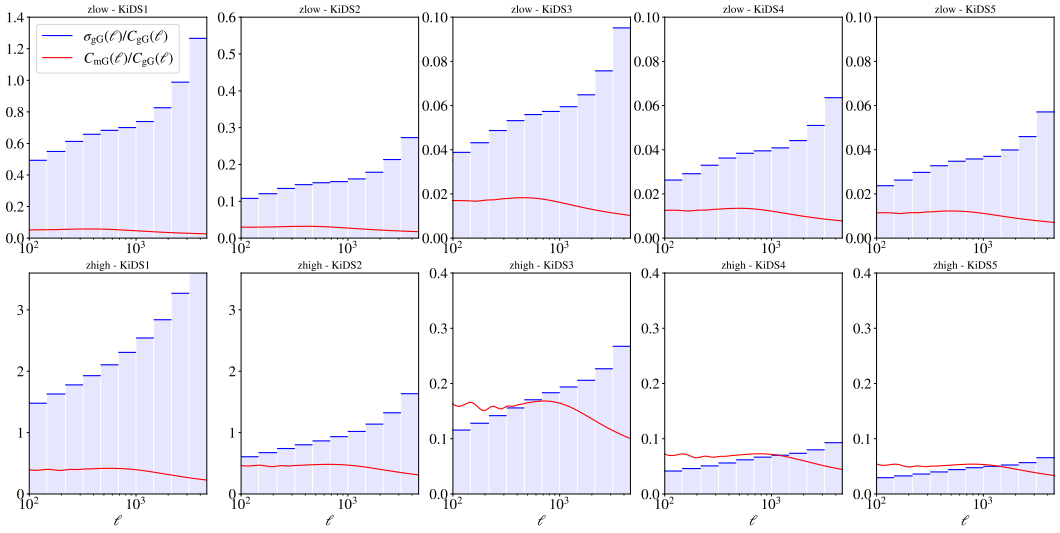

Figure 2.9: Magnification bias contribution $C_{\rm mG}(\ell)$ relative to the galaxy-galaxy lensing signal $C_{\rm gG}(\ell)$ over the angular scale ℓ (in red) for the crosscorrelations between the BOSS DR12 lens bins and the KiDS-1000 source bins assuming $\overline{\alpha}^{\rm zlow}_{\rm obs} = 1.93$ and $\overline{\alpha}^{\rm zhigh}_{\rm obs} = 2.62$. In blue, we show the expected relative uncertainty from shot and shape noise in the GGL signal, $\sigma_{\rm gG}(\ell)/C_{\rm gG}(\ell)$, within each ℓ bin (6 logarithmically spaced ℓ bins per dex). The uncertainties are calculated for a KiDS footprint with an area of 1,350 deg^2. The properties of the galaxy samples are given in Table 2.2.

within $100 < \ell < 4600$ between 0.4 and 0.5). In correlations with the zhigh lens bin, the contribution of the magnification-shear signal is larger and between $\sim 5\%$ and $\sim 20\%$ which is considerable above the shape and shot noise (with a cumulative SNR within $100 < \ell < 4600$ between 1.3 and 2.0). It is significant enough to give grounds for the consideration of this systematic during future GGL analyses which cross correlate the HSC Wide sample with the BOSS DR12 or a similarly selected lens sample.

2.4.3 *Euclid*-like Survey + DESI-like Survey Forecasts

We produce forecasts for a GGL analysis with Stage-IV (Albrecht et al., 2006), assuming lens and source samples akin to DESI (DESI Collaboration et al., 2016) and *Euclid* (Laureijs et al., 2011), respectively. We repeat the analysis shown in Section 2.4.1 and 2.4.2 for the *Euclid*-like source bins described in Table 2.2 and in Figure 2.8. We consider a footprint overlap between our source and lens sample of 6,000 deg^2, which is roughly the expected overlap between *Euclid* and DESI (Levi

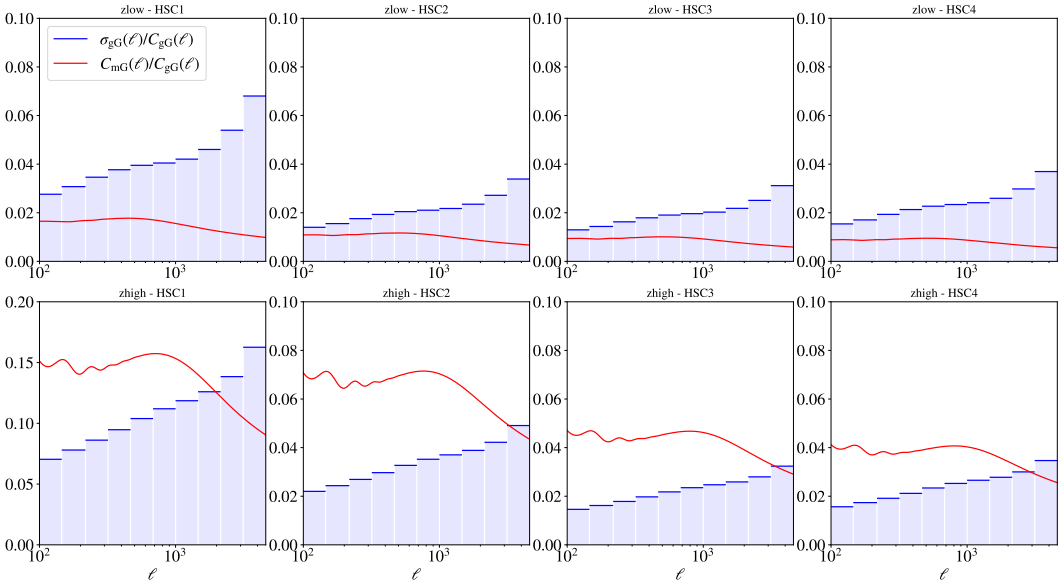

Figure 2.10: Same as Figure 2.9, but for HSC source bins as defined in Table 2.2 and Figure 2.8.

et al., 2013; DESI Collaboration et al., 2016). Therefore, the fictitious BOSS/DESI-like galaxy sample we are considering here has all the properties of the BOSS lens sample, but has the planned DESI footprint. Although DESI will probe higher redshifts and fainter galaxies than BOSS, it will be similar to BOSS in that it will not be a purely flux-limited survey. Targets in DESI are selected using a combination of different band magnitudes depending on the galaxy type and redshift range which is being observed (for more details see DESI Collaboration et al. 2016). For this reason, the magnification bias in the DESI sample cannot be modelled analytically either, warranting an analysis similar to the one discussed here. The *Euclid*-like source sample used in this work is designed to be split into the same redshift bins as suggested by *Euclid* collaboration forecast choices (Euclid Collaboration et al., 2020). In addition, within each bin, the median redshift is chosen to be in agreement with the one expected for the *Euclid* sources.

Considering 6 logarithmically spaced ℓ bins per dex (as in the previous sections), we obtain the magnification-shear signal forecasts shown in Figure 2.11. We see that the magnification-shear signal constitutes a considerable systematic when correlating with the zlow bin, since the observed cumulative SNR on scales within

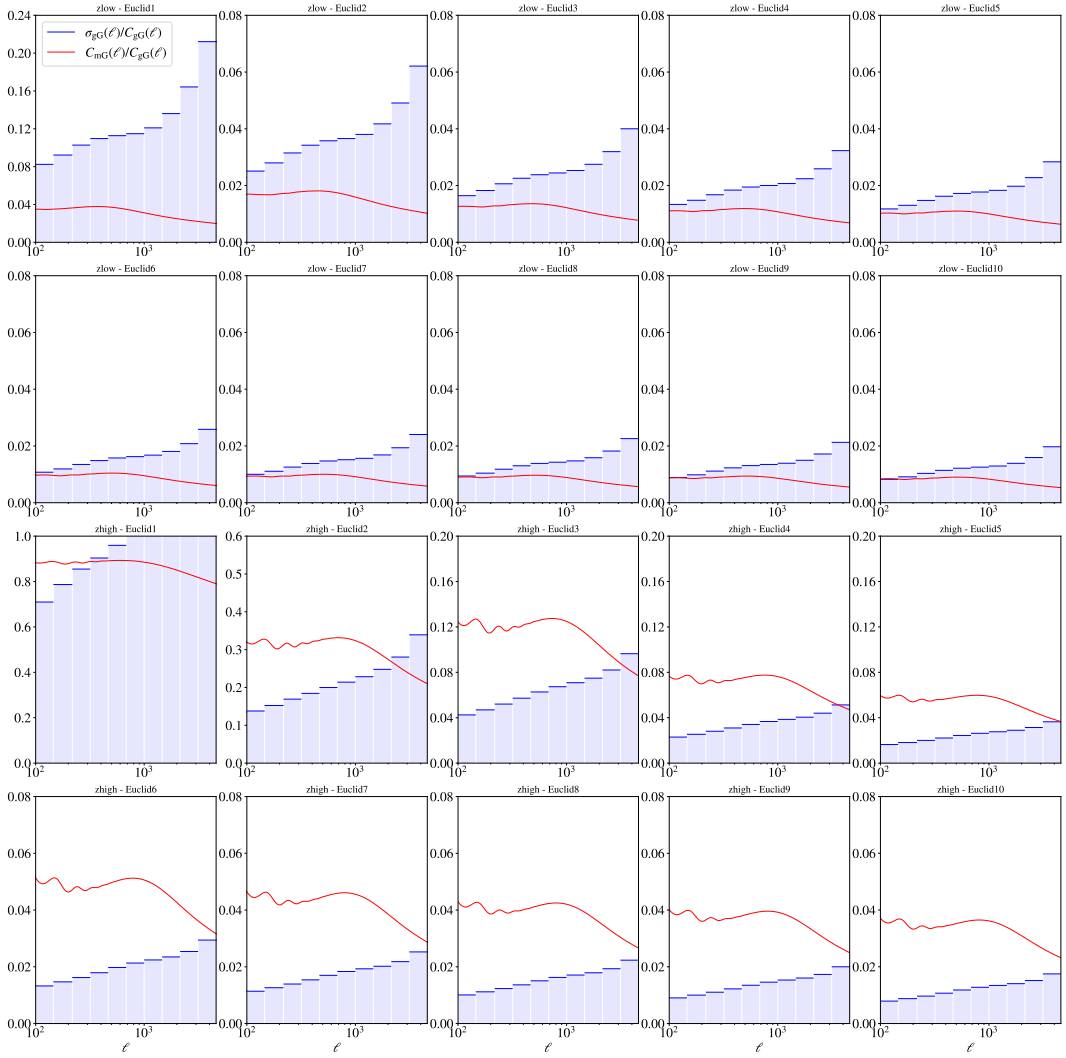

Figure 2.11: Same as Figure 2.9, but for *Euclid*-like source bins as defined in Table 2.2 and Figure 2.8.

$100 < \ell < 4600$ is between 0.3 and 0.7. The magnification bias signal becomes strong enough for correlations with zhigh, it would be a detectable signal (with the cumulative SNR within $100 < \ell < 4600$ ranging from 1.5 when correlating zhigh and Euclid1 to 2.8 when correlating zhigh and Euclid10). This might require any future GGL analysis of *Euclid*+BOSS or *Euclid*+DESI data to allow for the α parameters to freely vary as a nuisance parameter in order to properly account for this systematic. The method outlined in this work could be used to set informative priors on the α values within each lens bin.

2.5 Conclusions

In this chapter, we have introduced a novel method to estimate the effective luminosity function slope, α, of galaxy samples which have been defined with a complex selection function that is not simply flux/magnitude-limited. The method calibrates the α estimates from observables with accurate cosmological simulations with the same sample selection. This expands upon previous work where the flux magnification was only measured for flux-limited cases or found to be inaccurate in non-flux-limited cases (Hildebrandt, 2016).

The new method determines the underlying slope of the luminosity function of the simulated galaxy sample (α_κ) from unobservable properties such as the convergence, κ, and the unlensed galaxy position. It then finds the magnitude range relative to the magnitude limit over which the resulting $\alpha_{\rm obs}$ as calculated from the observable differential galaxy count distribution, $n(m)$, best agrees with α_κ. Finally, the same relative magnitude range is used to determine $\alpha_{\rm obs}$ from the observed galaxy sample.

A few things should be considered when employing this method. We find that the magnitude ranges up to the effective magnitude limit that are determined to be optimal from the simulations in order to calibrate $\alpha_{\rm obs}$ are only valid for a given redshift range, a given sample selection function and a given galaxy sample for which weak lensing simulations are available. Thus, it is important to note that this method cannot be generalised trivially, as it requires the availability of accurate cosmological simulations to assure consistency between the two independent α estimates, $\overline{\alpha}_{\rm obs}$ and $\overline{\alpha}_\kappa$. Nonetheless, when simulations are available, it provides a robust estimate of the scale of the magnification bias for non-flux-limited surveys such as BOSS.

Applying our calibration method to the BOSS DR12 sample split into two redshift bins, we find that $\overline{\alpha}_{\rm obs} = 1.93 \pm 0.05$ for $0.2 < z \leq 0.5$ and $\overline{\alpha}_{\rm obs} = 2.62 \pm 0.28$ for $0.5 < z \leq 0.75$ leading to a contribution to the galaxy-galaxy lensing signal of up to $\sim 2\%$ for KiDS-1000 and HSC Wide sources correlated with the $0.2 < z \leq 0.5$ lens bin. Although the contribution can go up to $\sim 20\%$ when correlating

KiDS-1000 and HSC Wide sources with the $0.5 < z \leq 0.75$ BOSS lens bin, the magnification-shear signal can go above the noise with a cumulative SNR going up to 1.1 and 2.0 for KiDS-1000 and HSC Wide, respectively. Hence, both for KiDS-1000 and HSC Wide, the magnification-shear signal appears to be dominant enough to warrant the modelling of this systematic in future GGL analyses involving BOSS lenses, as was already done in the recent KiDS-1000 analysis (Joachimi et al., 2021; Heymans et al., 2021). This necessity becomes even more evident in the forecasts for a GGL analysis of *Euclid*-like sources with DESI-like lenses. In this case, the magnification-shear signal is either a considerable systematic when correlating with the zlow bin (with a cumulative SNR of around 0.5), or it even becomes a detectable signal when correlated the source bins with zhigh giving cumulative SNRs around 2 which can go up to 2.8. This might require any future GGL analysis incorporating *Euclid* and any highly selected lens sample (e.g. DESI or BOSS) to allow for the effective luminosity function slope (α) of each lens sample to vary freely within the model using informative priors based on an analysis similar to the one conducted in this work. These results are in line with Duncan et al. (2014) as well as the recent findings from Mahony et al. (2022) where it was determined that the inclusion of the magnification bias in the modelling for surveys such as the next generation of surveys is necessary to accurately infer cosmological parameters.

We expect similar conclusions for other surveys. It might be desirable to estimate the magnification bias using the methodology outlined in this work in clustering and GGL analyses based on DES REDMAGIC lens galaxies such as the ones described in Clampitt et al. (2017), Elvin-Poole et al. (2018) and Prat et al. (2018), since it also follows a complex selection function (Rozo et al., 2016). The SNR should be comparable to HSC and KiDS, so the magnification bias will not have to be included as a free parameter. On the other hand, for surveys such as LSST (LSST Science Collaboration et al., 2009) and the Nancy Grace Roman Space Telescope (formerly known as WFIRST, Spergel et al. 2015), it may become necessary to make the α of the lens galaxy samples a nuisance parameter in any clustering or GGL analysis, as we suggest for a *Euclid*+DESI-like analysis.

Chapter 3

Simulation-Based Inference of KiDS-1000 Cosmic Shear from Statistical Forward-Simulations

Cosmic shear, the weak gravitational lensing effect on distant galaxies due to matter in the foreground, is a powerful tool to study the distribution of matter in the Universe and to probe its large-scale structure. By measuring the distortions in the shapes of galaxy images caused by the gravitational influence of intervening matter, we can infer a combination of the matter density and the amplitude of the matter power spectrum within the framework of a "Dark Energy Cold Dark Matter", ΛCDM, cosmology. Recent cosmic shear analyses of the data taken in stage-III galaxy surveys, such as the Kilo-Degree Survey[1] (KiDS; Kuijken et al. 2019; Asgari et al. 2021; Heymans et al. 2021; van den Busch et al. 2022), the Hyper Suprime-Cam survey[2] (HSC; Sugiyama et al. 2022; Aihara et al. 2022), and the Dark Energy Survey[3] (DES; Gatti et al. 2021; Amon et al. 2022; Secco et al. 2022), have constrained these cosmological parameters with unprecedented precision. Upcoming stage-IV galaxy surveys, such as *Euclid*[4] (Laureijs et al., 2011), *Rubin*[5] (LSST Science Collaboration et al., 2009) or *Roman*[6] (Spergel et al., 2015), will

[1] https://kids.strw.leidenuniv.nl/
[2] https://hsc.mtk.nao.ac.jp/ssp/
[3] https://www.darkenergysurvey.org/
[4] https://www.euclid-ec.org/
[5] https://www.lsst.org/
[6] https://roman.gsfc.nasa.gov/

further improve upon these constraints, as these surveys will be wider, deeper and more precise. However, many analysis challenges still remain.

Current stage-III galaxy survey measurements of the root-mean square of the matter overdensity field at 8 Mpc h_0^{-1}, σ_8, or its analogue $S_8 \equiv \sigma_8(\Omega_\mathrm{m}/0.3)^{0.5}$, agree well with each other (Hikage et al., 2019; Asgari et al., 2021; van den Busch et al., 2022; Amon et al., 2022) despite using largely independent methodologies. However, the constraints from observations of the late Universe (Asgari et al., 2021; Heymans et al., 2021; van den Busch et al., 2022; Amon et al., 2022; Abbott et al., 2022) disagree by up to $\sim 3.4\sigma$ with the σ_8 value consistent with the early-Universe observations from the Cosmic Microwave Background (CMB; Planck Collaboration et al. 2020). This discrepancy might be pointing at new physical phenomena, but it could also be caused by unconsidered systematic effects when modelling the cosmic shear or CMB signal and noise. Such systematic effects may be physical, e.g. baryonic feedback or the intrinsic alignments of galaxies (Kilbinger, 2015; Mandelbaum, 2018; Amon & Efstathiou, 2022; Li et al., 2023; Miyatake et al., 2023), while they may also be observational, e.g. shear mesurement bias, depth variability, point-spread function variation, etc.

We aim to shed some light onto this by, for the first time, conducting a full cosmic shear analysis with the same complexity as current stage-III analyses using simulation-based inference (SBI). SBI, also known as Likelihood-Free Inference or Implicit Likelihood Inference, is a Bayesian inference method which does not require an explicit formulation for the form of the likelihood function of the data given the parameters of interest. Instead, the likelihood is implicitly calculated by evaluating the joint probability of data and parameters of interest from forward simulations which map the parameters to the corresponding mock data vectors. This comes with multiple advantages with respect to other standard approaches which require an explicit form for the likelihood. Firstly, the likelihood is allowed to take an arbitrary form, so one can avoid the typical assumption of a Gaussian likelihood or avoid having to define a complex analytical expression for the likelihood. Second, for some models and measurements, it may not even be possible to define an analytical likeli-

hood or it would be too resource-intensive. In such cases, as long as the observables can be simulated, an effective/implicit likelihood may be found using SBI. Lastly, in many cases, it is already necessary to create forward simulations to validate the signal and noise modelling used within a standard inference pipeline. When running a Markov-Chain Monte Carlo (MCMC) to sample the posterior distributions from a Gaussian likelihood, one either has to compute a numerical covariance matrix from forward simulations or create an analytical model for the covariance which ought to be validated with a numerical one. In both cases, for a data vector of size $|\boldsymbol{d}|$, one would require a number of simulations $> |\boldsymbol{d}|$ to get an accurate covariance matrix. This is already the approximate amount of forward simulations needed for an SBI analysis (Alsing et al., 2018; Lin et al., 2023). Therefore, SBI allows one to do full Bayesian uncertainty propagation from data to parameters for any model which can be simulated, without substantial additional computational costs when compared to standard approaches at comparable accuracy.

In cosmic shear, we know that the commonly used two-point statistics to quantify the shear-shear correlations of galaxies on the sky have an approximately Gaussian likelihood, but this assumption has its limitations (Eifler et al., 2009; Schneider & Hartlap, 2009; Sellentin & Heavens, 2018; Sellentin et al., 2018; Taylor et al., 2019; Upham et al., 2021). Additionally, when observational and physical systematics are included, the likelihood may become even less Gaussian (Jeffrey et al., 2021). In such cases, like the one presented in this chapter, SBI allows one to capture potential non-Gaussanities and propagate them to the posterior distributions. SBI can also compute effective likelihoods for summary statistics for which analytical methods may be intractable. For example, the likelihoods of higher-order cosmic shear statistics can be complex to evaluate, but SBI can consider the full complexity of the signal and relevant systematics in the likelihood (Fluri et al., 2022). This could help current and upcoming galaxy surveys to make sure that as much cosmological information as possible is rigorously extracted from the observations. In this work, we limit the scope to only two-point measurements of cosmic shear, and leave higher-order statistics or field-level statistics within SBI as an av-

enue for future work. Of course, other summary statistics of cosmic shear such as higher-order correlators, full shear fields or entire shape catalogues could still benefit from the SBI approach laid out here, as it can implicitly characterise their likelihoods which are known to have significant non-Gaussianities.

A potential source of non-Gaussianities in the likelihood of two-point statistics is the anisotropy in the observational depth of a galaxy survey across the sky (Guzik & Bernstein, 2005; Shirasaki et al., 2019). The images at different pointings taken by a galaxy survey, such as KiDS, may be subject to different environmental conditions depending on the time of observation (e.g. atmospheric conditions, thermal expansion of the telescope, etc.). When the images of each pointing are subsequently combined, it may also be the case that due to overlapping areas between pointings, some galaxies are observed more often than others which can cause variations in the signal-to-noise ratio which are spatially correlated. In effect, all these variations lead to fluctuations in the effective galaxy density which are observed across the footprint, while also biasing the shape measurements made. This has been determined to be a percentage-level effect on the cosmic shear signal measured by KiDS (Heydenreich et al., 2020; Baleato Lizancos & White, 2023), and it has been found to be an effect $< 20\%$ on the standard deviation of the KiDS-1000 cosmic shear signal (Joachimi et al., 2021). Although for KiDS this is within the random uncertainties, it will become important for upcoming galaxy shear surveys such as *Euclid*, *Rubin* or *Roman*. Additionally, the effect of spatial variability on the non-Gaussianity of the likelihood has not been explored in previous work. To address these needs, in this work, our forward model includes the effects of the observational depth variations in KiDS, so that we may propagate the effects of spatial variations in the depth all the way to the uncertainty on inferred cosmological parameters.

Although other approaches to SBI exist (e.g. Approximate Bayesian Computation, ABC; Rubin 1984; Pritchard et al. 1999; Lin & Kilbinger 2015; Lin et al. 2016; Beaumont 2019), we choose to make use of Density Estimation Likelihood-Free Inference (DELFI), as it offers better performance and therefore scales better

when increasing the dimensionality of the parameter space of interest (Leclercq, 2018; Alsing et al., 2018). Simultaneously, DELFI discards less information from the forward simulations than ABC. ABC only keeps the forward simulations which have a certain level of agreement with the desired data vector. In contrast, DELFI learns a probability density distribution of the data vectors as a function of the model parameters based on all of the forward simulations obtained across the prior volume. To achieve this, DELFI employs ensembles of neural density estimators in order to learn the joint probability distribution of the data and the parameters from the forward simulations. Recently, this method has become increasingly popular, among other things, because it requires about an order of magnitude fewer evaluations of the likelihood to constrain the posterior than a standard MCMC analysis (Papamakarios & Murray, 2016; Alsing et al., 2018; Taylor et al., 2019; Gerardi et al., 2021; Leclercq & Heavens, 2021; Jeffrey et al., 2021; Lemos et al., 2021; Legin et al., 2021; Mishra-Sharma & Cranmer, 2022; Mancini et al., 2022; Lin et al., 2023; Chen et al., 2022; Hu et al., 2022; Lemos et al., 2023b).

In a previous paper (Lin et al. 2023; L22 hereafter), we introduced a new inference pipeline for estimating the cosmological parameters from the cosmic shear data of KiDS-1000 (Kuijken et al., 2019) using DELFI. In L22, the SBI pipeline for KiDS is shown to be robust, accurate and efficient, even when constraining a 12-dimensional posterior distribution. However, in that work, the simulated vectors are based on random samples from the covariance matrix used in the fiducial KiDS-1000 analysis (Joachimi et al., 2021; Asgari et al., 2021; van den Busch et al., 2022), so the likelihood is Gaussian by construction to allow for direct comparison with the standard formalism. In this work, we present a novel suite of physically motivated and realistic forward simulations of the cosmic shear observables as seen by KiDS, so the neural density estimators may learn any non-Gaussianities in the likelihood induced by our model plus any complex relations which may arise between the data and the model parameters. We then use these simulations to create an inference pipeline consistent in its modelling choices with previous analyses, while also considering some additional systematic effects in our uncertainty model. The

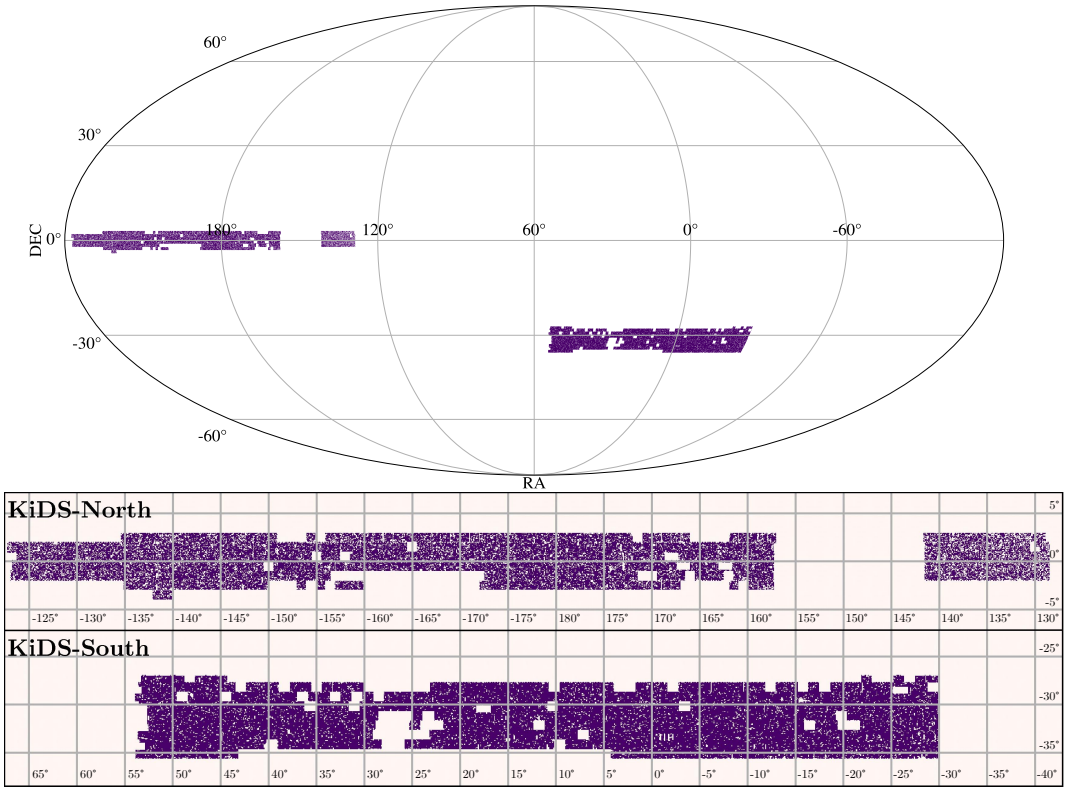

Figure 3.1: Spatial map of the KiDS-1000 footprint. The top panel shows a Mollweide projection of the full KiDS-1000 footprint, while the two panels at the bottom show zoomed-in Cartesian projections of KiDS-North and KiDS-South fields, respectively.

obtained posteriors are then scrutinised for accuracy and robustness.

This chapter has the following structure. Section 3.1 presents the Kilo-Degree Survey's data used for this analysis. Section 3.2 gives a detailed description the forward simulations developed for the SBI analysis. Section 3.3 describes the setup of the simulation-based inference pipeline using DELFI based on L22. Section 3.4 presents the choices made in the forward simulations, based on previous KiDS-1000 analyses while including anisotropies in the observational depth and a more complex shear bias model. We also validate the inference pipeline by inferring the cosmological parameters from a mock data vector, and by conducting a coverage test of the learned posterior. Finally, Section 3.5 shows the posterior contours and the estimates of the cosmological parameters obtained from mock KiDS-1000 DR4 data. We then provide some final remarks and conclusions in Section 3.6, while Appendix B shows additional details on the cosmic shear signal modelling.

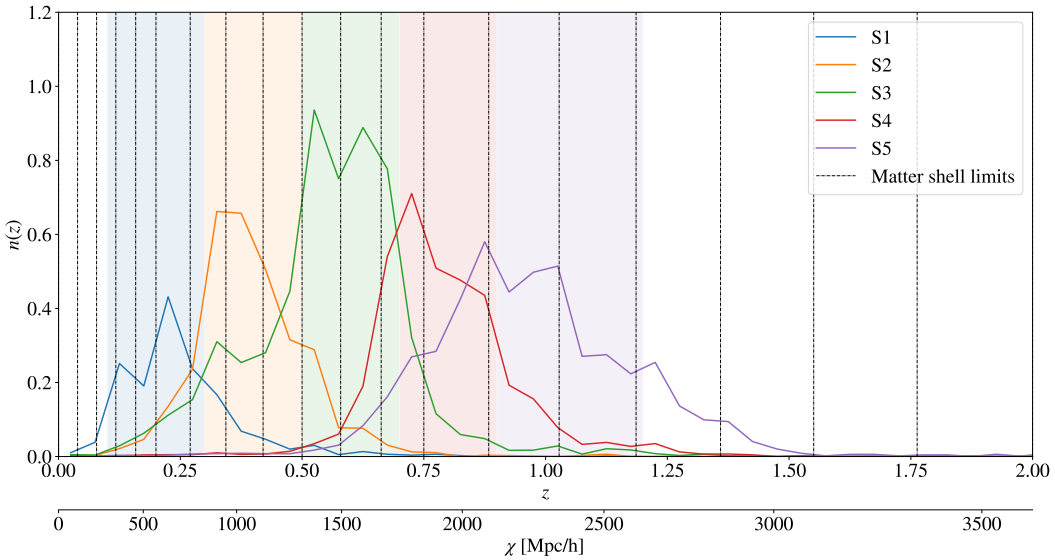

Figure 3.2: Plot of the redshift distributions of the five KiDS-1000 tomographic bins. The shaded areas show to limits of each tomographic bin, while the solid lines show the $n(z)$ of the source galaxies in each tomographic bin as a function of both redshift, z, and comoving distance, χ (the latter is derived assuming a Planck 2018 cosmology; Planck Collaboration et al. 2020). The black dashed lines show the limits of the spherical matter shells in our forward simulations.

3.1 KiDS-1000 Data

The Kilo-Degree Survey (KiDS) is a large public galaxy survey conducted by the European Southern Observatory using the OmegaCAM CCD mosaic camera (Kuijken, 2011) which is attached to the 2.6 m VLT Survey Telescope (VST). The survey covers approximately 1,350 deg^2 between two distinct fields known as KiDS North (which straddles along the DEC = 0º) and KiDS South (which straddles along the DEC = -30º). This area is the same as the one covered by the VISTA Kilo-degree INfrared Galaxy survey (VIKING; Edge et al. 2013), which means that both surveys together observe every object with a total of nine photometric bands: $ugriZYJHK_s$. The analysis presented in this work makes use of DR4 (Kuijken et al., 2019), also known as KiDS-1000, which covers approximately 1,000 deg^2. The effective area covered by KiDS-1000 shape measurements shown in Figure 3.1 covers 773.3 deg^2 (Joachimi et al., 2021). This area is calculated directly from the mosaic mask defined a the native OmegaCAM pixel scale of 0.213 arcsec.

KiDS images are processed using ASTRO-WISE (McFarland et al., 2013) and

the methodology outlined in Wright et al. (2019) to join the VST and VIKING photometric measurements. To obtain the shape measurements, THELI (Erben et al., 2005) is used to process the r-band images, while the gravitational shear estimates are then obtained from LENSFIT (Miller et al., 2007, 2013; Fenech Conti et al., 2017). Depending on the shape noise variance and the ellipticity measurement noise variance, LENSFIT assigns a weight to the shear measurement of each galaxy, w_i, which scales the shear signal such that the signal-to-noise ratio is optimal. Although KiDS-SBI currently does not model LENSFIT weights (all galaxies are weighted equally by default), the weights are carried through the measurement pipeline when measuring the cosmic shear summary statistics from the real KiDS DR4 data. If the LENSFIT weight is mischaracterised, this may lead to shear bias. For this purpose, the shape measurements in KiDS DR4 are calibrated to determine any shear biases due selection biases, noise, weight bias, point-spread function residuals or otherwise are calibrated (Giblin et al., 2021). The resulting estimate of the linear multiplicative shear bias per tomographic bin is then used to unbias the galaxy shear measurements in KiDS-1000. All details on the reduction of the images and the shape calibration are given in Kuijken et al. (2019); Wright et al. (2019); Giblin et al. (2021).

The photometric redshifts are estimated from the KiDS 9-band photometry using Bayesian template-fitting as incorporated within the BPZ code (Benítez, 2000; Wright et al., 2019, 2020a). The photometric redshifts are calibrated with spectroscopic redshift measurements made of galaxies in the KiDS sample (Hildebrandt et al., 2021) based on a subsample of 99% completeness. The mapping from photometric to spectroscopic redshift, $P(z|z_{\rm ph})$, is learnt through a self-organising map (Wright et al., 2020a). The maximum posterior redshift is subsequently used to split the source galaxy catalogue into five tomographic bins at the following boundaries: $\{0.1, 0.3, 0.5, 0.7, 0.9, 1.2\}$. These then give the redshift distributions shown in Figure 3.2. As the cosmological constraints are sensitive to shifts in the redshift distributions, the KiDS-1000 analysis also allows the shift in the mean of each tomographic bin to vary freely in the cosmological inference as a nuisance parameter

(Asgari et al., 2021; Heymans et al., 2021; van den Busch et al., 2022). This does not only avoid biases in the cosmological parameter estimates, it also provides an additional test on the redshift calibration.

3.2 Forward Simulations

The forward simulations for cosmic shear analyses described in this chapter are fully implemented and available in the KiDS-SBI[7] module. This code is built within the Cosmological Survey Inference System (CosmoSIS; Zuntz et al. 2015) and it is based upon the KiDS Cosmology Analysis Pipeline (KCAP; Joachimi et al. 2021; Asgari et al. 2021; Heymans et al. 2021). In addition, the KiDS-SBI pipeline is generalisable to other cosmic shear analyses.

The overall outline of the forward simulations is shown in Figure 3.3. Firstly, the simulation is fed a set of cosmological and astrophysical parameters, $\mathbf{\Theta}$, from a sampler of our choice. All of these parameters are passed into CAMB (Lewis et al., 2000; Lewis & Challinor, 2002; Howlett et al., 2012) which calculates the three-dimensional matter power spectrum, $P_{\delta,\mathrm{nl}}$, at a given cosmology (see also Section 3.2.1 for more details). We incorporate baryonic feedback using HMCode (Mead et al., 2016). Next, the simulation splits the light-cone along the line-of-sight into concentric spherical shells centred at the observer, based on the approach in Xavier et al. (2016) and Tessore et al. (2023). Such an approach is computationally efficient as all observables are simulated with a coarse resolution along the line-of-sight. Within each shell, we sample log-normal random matter fields to efficiently model the two-point statistics with high accuracy, while also incorporating some higher-order fluctuations. As an added benefit, this simplifies the implementation of effects such as redshift evolution as well as survey characteristics, as they usually also follow a spherical geometry. On the other hand, the concentric shells also allow to split the Universe into shells, which are of a scale large enough such that a log-normal field becomes an accurate description of the matter over-densities within it (see Sects. 3.2.2 for details).

[7]Kilo-Degree Survey – Simulation-Based Inference; `https://github.com/mwiet/kids_sbi`

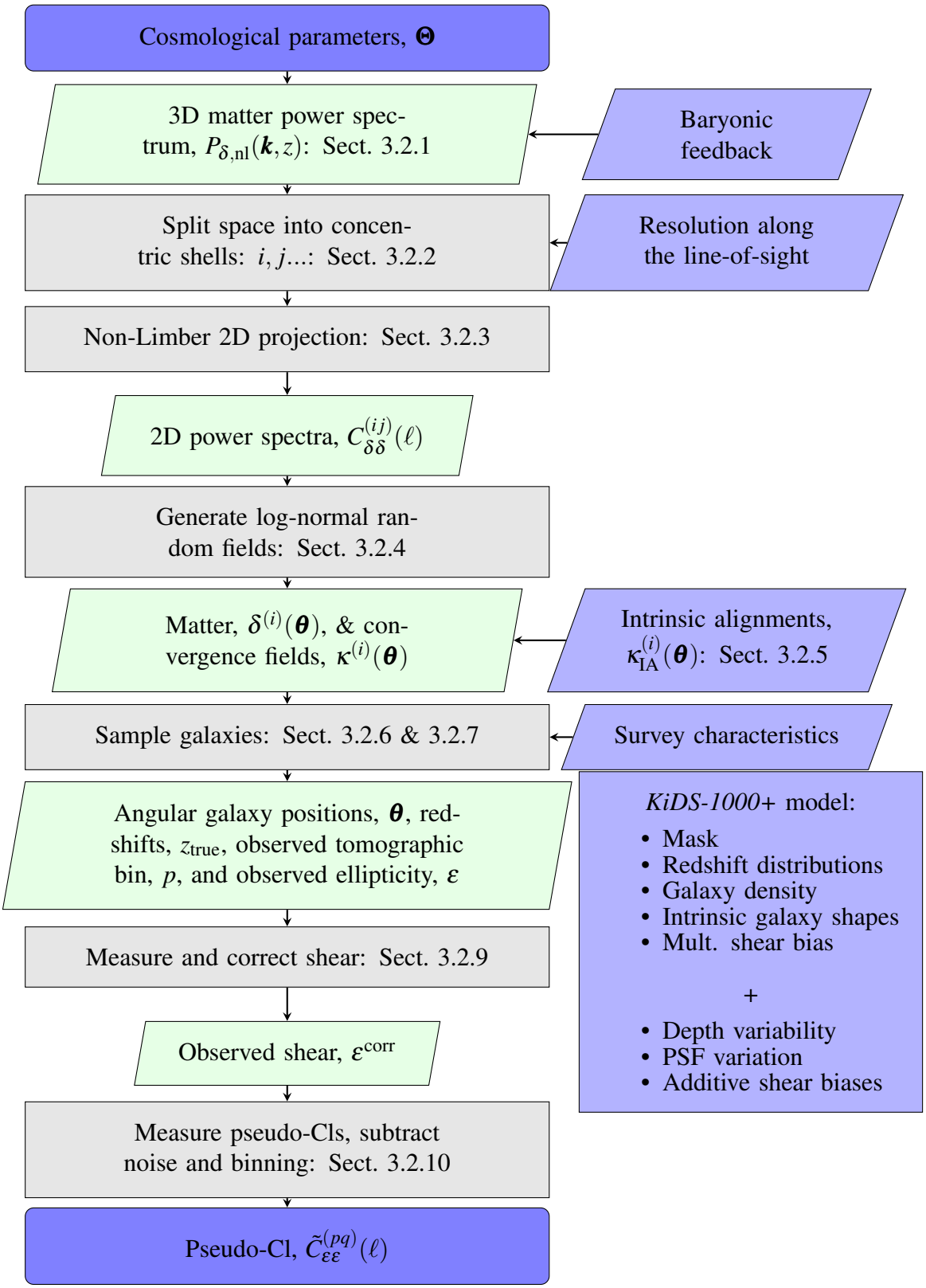

Figure 3.3: Flowchart describing the steps in a single forward simulation of cosmic shear observables from cosmological parameters. The dark blue rounded boxes represent the inputs and outputs which are given to the simulation-based inference pipeline. The green slanted boxes represent relevant quantities which are calculated during the simulation. The grey rectangular boxes show steps in the calculations, while the blue slanted boxes show any (systematic) effects which are included. All variables are defined within the respective sections quoted in the diagram.

Once we know the geometry of our matter shells, we can project the three-dimensional matter power spectrum using the redshift distributions of these shells in order to obtain their two-dimensional angular power spectra, $C_{\delta\delta}$. The latter describes the correlations between the matter over-densities in each shell. Since we also require accurate cross-correlations between different shells to compute the log-normal fields, the Limber approximation is not applicable here, as the cross-correlations are zero in that approximation with the shells being non-overlapping along the line-of-sight. To avoid this, we project the three-dimensional matter power spectrum with a non-Limber integral. This is usually an expensive computation, so we show here a novel implementation of the Levin method of integration (Levin, 1996) which allows for the efficient integration of the Bessel functions found in non-Limber integrals (more details are given in Section 3.2.3).

Knowing how the matter overdensity fields within the concentric shells correlate with themselves and each other, we can then construct log-normal matter fields which are consistent with these correlations and their associated cosmology. To do this, we use the Generator for Large Scale Structure (GLASS[8], Tessore et al. 2023), which efficiently generates a log-normal random matter field, δ, within each shell. Then, it accurately computes the respective convergence field, κ, (see Section 3.2.4) while also incorporating an effective convergence field due to intrinsic galaxy alignments, $\kappa_{\rm IA}$ (see Section 3.2.5).

Upon the construction of the matter, δ, and convergence fields, κ, we use them to Poisson-sample galaxies that reside within the matter fields as well as their properties, i.e. intrinsic ellipticities and shear. During this step, with the aid of the SALMO module (Joachimi et al., 2021), we also consider survey characteristics, such as the survey footprint, its spatial variability in depth, in its redshift distributions, as well as its spatial variability in the shape and shot noise as described in Section 3.2.6. Thus, we obtain a full galaxy catalogue with shape information consistent with the input cosmology as well as with the previously mentioned survey properties.

[8]Generator for Large Scale Structure; `https://github.com/glass-dev/glass`

The last steps of the forward simulation pipeline involve the post-processing and compression of the simulated data catalogue. In principle, any data vector could be obtained here, such as shear two-point correlation functions (Asgari et al., 2021; van den Busch et al., 2022), the shear fields (Porqueres et al., 2022) or even the full catalogues. However, in order to reduce the dimensionality of the data vector efficiently, while still retaining most information about the cosmology, we choose observed angular power spectra, a.k.a. pseudo-Cls, as our data vector of choice. This allows us to compare our SBI analysis to the results from the standard analysis of KiDS-1000 data using pseudo-Cls in Loureiro et al. (2021). The KiDS-SBI pipeline is, in principle, set up to make cosmological inferences from cosmic shear based on any statistic which could be derived from a galaxy catalogue, but the underlying statistical random fields are currently only capable of achieving percentage-level accuracy for two-point statistics. The statistics and its post-processing are defined in Section 3.2.10.

The forward simulations are designed to model weak gravitational lensing observations on the level of galaxy catalogues. The galaxy populations are sampled such that they trace the underlying log-normal random matter fields, while the galaxy shapes are lensed in accordance with the lensing potential of the matter fields along the line-of-sight. As the simulations do not model images, any image-level systematic effects, such as shear biases or variable depth, are included as spatially varying probability density functions.

3.2.1 Cosmology Dependence: 3D Matter Power Spectrum

The basis of the forward simulations is the 3-dimensional matter power spectrum, $P_{\delta,\mathrm{nl}}(k,z;\mathbf{\Theta})$. In the linear regime, the equal-time three-dimensional matter power spectrum's cosmology dependence comes from the following relation

$$P_{\delta,\mathrm{l}}(k,z;\mathbf{\Theta}) = T^2(k;\mathbf{\Theta})\,D^2(z;\mathbf{\Theta})\,P_{\mathscr{R}}(k;\mathbf{\Theta}), \tag{3.1}$$

where k is a wavenumber, z is redshift, $\mathbf{\Theta}$ is the set of cosmological parameters, $P_{\mathscr{R}}(k;\mathbf{\Theta})$ is the primordial density fluctuation power spectrum, D is the

growth factor of structure, T is the transfer function between the matter overdensity field, $\boldsymbol{\delta}(\boldsymbol{k},z)$, and the primordial curvature fluctuation field, $\mathscr{R}(\boldsymbol{k};\boldsymbol{\Theta})$, such that $\boldsymbol{\delta}(\boldsymbol{k},z) = T(\boldsymbol{k};\boldsymbol{\Theta})D(z;\boldsymbol{\Theta})\mathscr{R}(\boldsymbol{k};\boldsymbol{\Theta})$. To compute this, we make use of CAMB[9] (Lewis et al., 2000; Lewis & Challinor, 2002; Howlett et al., 2012). To stay in line with the main KiDS-1000 analysis (Joachimi et al., 2021; Asgari et al., 2021; van den Busch et al., 2022), we assume a normal neutrino hierarchy, while also assuming a fixed sum of neutrino masses of $\sum m_\nu = 0.06$ eV/c^2.

We compute the non-linear matter power spectrum via a non-perturbative model, HMCODE-2016 (Mead et al., 2015, 2016). Although updated iterations of this model exist (Mead et al., 2021), we limit ourselves to HMCODE-2016, so our model remains comparable to the KiDS-1000 cosmic shear analysis. HMCODE uses a halo model approach to incorporate the effects of baryonic feedback on the matter distribution. The main driving factor in this is the baryonic matter expelled by Active Galactic Nuclei (AGN). This suppresses the power at small scales, i.e. large k, as a function of the amplitude of the halo mass-concentration relation, $A_{\rm bary}$, and the halo bloating parameter, η_0. We also fix the relation between these two parameters to $\eta_0 = 0.98 - 0.12A_{\rm bary}$ in accordance with Joudaki et al. (2018) and Joachimi et al. (2021), and treat $A_{\rm bary}$ as the only free parameter related to baryonic feedback.

We note that this treatment may lead to systematic biases at small scales, as the galaxies will be sampled from a matter power spectrum which is already modified by baryonic feedback. It would be more physically accurate to sample the galaxies from a matter power spectrum without baryonic non-linearities, and subsequently adding a perturbation to the matter density contrast fields due to baryonic feedback (similar to the approach used for the intrinsic alignments described in Section 3.2.5). Nevertheless, we do not expect this to cause a large discrepancy for the scales which we are probing in the cosmic shear signal ($\ell < 1500$). Field-level implementations of baryonic feedback already exist (Schneider & Teyssier, 2015; Schneider et al., 2019, 2022), and they would constitute interesting avenues for future extensions to

[9]Code for Anisotropies in the Microwave Background; `https://camb.info`

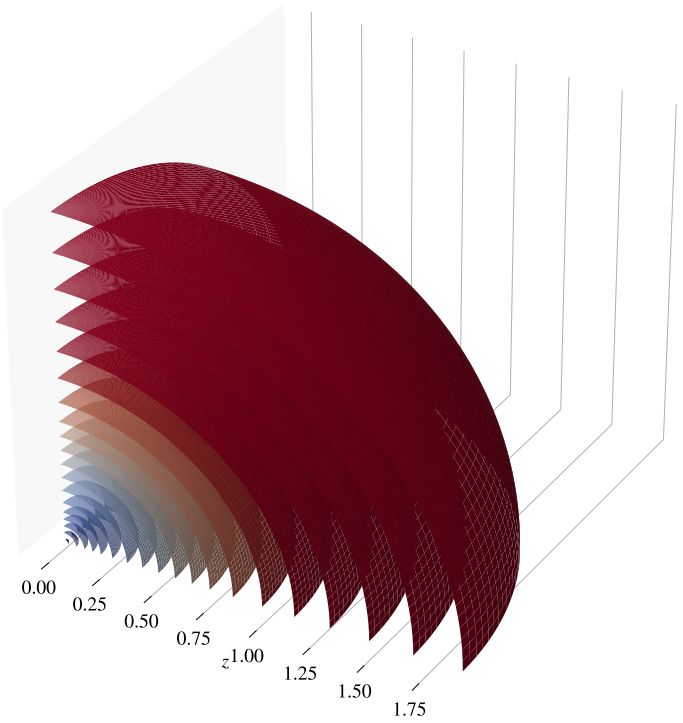

Figure 3.4: Three-dimensional diagram showing an octant of 19 concentric shells as they are simulated within KiDS-SBI and GLASS. The radius of each shell is given by its mean redshift.

KiDS-SBI.

3.2.2 Working on the Sphere

Astronomical observations are typically projected onto a two-dimensional surface of a sphere with the observer at its centre. Since this is the case for large surveys such as KiDS, most observational biases and some systematics depend on the spherical coordinates on the sky. For this reason, it is natural to express our model using the same geometry.

The KiDS-SBI forward simulations use GLASS (Tessore et al., 2023)

to model the underlying large-scale structure. Within GLASS, we choose to model large-scale structure through log-normal random matter fields within non-overlapping concentric shells centred at the observer. Therefore, a given shell is a comoving volume which spans the full sky and has a finite width along the line-of-sight from one redshift, z_i, to another, z_{i+1}, such that $i \in \{1,2,3,...,N_{\text{shells}}\}$. We define the matter weight function $W^{(i)}(z)$ describing the distribution along the line-of-sight of the comoving volume spanned by a given shell as (Tessore et al., 2023),

$$W^{(i)}(z;\boldsymbol{\Theta}) = \begin{cases} f_k^2(z;\boldsymbol{\Theta})/E(z;\boldsymbol{\Theta}), & \text{if } z_i \leq z < z_{i+1}, \\ 0, & \text{otherwise}, \end{cases} \tag{3.2}$$

where $E(z;\boldsymbol{\Theta})$ is the dimensionless Hubble function and $f_k(z;\boldsymbol{\Theta})$ is the transverse comoving distance for a given redshift and cosmology, which is defined as

$$f_k(z;\boldsymbol{\Theta}) = \begin{cases} \frac{c}{H_0\sqrt{\Omega_k}} \sinh\left(\frac{\sqrt{\Omega_k}H_0}{c}\chi(z;\boldsymbol{\Theta})\right), & \text{if } \Omega_k > 0, \\ \chi(z;\boldsymbol{\Theta}), & \text{if } \Omega_k = 0, \\ \frac{c}{H_0\sqrt{|\Omega_k|}} \sin\left(\frac{\sqrt{|\Omega_k|}H_0}{c}\chi(z;\boldsymbol{\Theta})\right), & \text{if } \Omega_k < 0, \end{cases} \tag{3.3}$$

where H_0 is the Hubble constant, Ω_k is the curvature density parameter and c is the speed of light in a vacuum.

A given shell has to cover at least approximately 100 Mpc of comoving distance along the line-of-sight for log-normal random fields to be an accurate enough representation of the distribution of large-scale structure (Xavier et al., 2016; Hall & Taylor, 2022; Tessore et al., 2023). Otherwise, discretisation effects would smooth away large-scale structure (Tessore et al., 2023). Additionally, one needs to be mindful of computational resources, because the computation of angular power spectra for the correlation between the matter fields within each shell scales as $N_{\text{shells}}(N_{\text{shells}}+1)/2$. Simultaneously, the simulations' run-time for a given set of cosmological parameters scales linearly with N_{shells} and the memory requirements by $2N_{\text{shells}}$ (as we need to store a matter field map and a convergence map per shell). To balance all of these factors, we aim to have as many shells along the line-of-

sight as computationally feasible, while also choosing widths which come as close as possible to the 100 Mpc limit.

When choosing the number of shells and their widths in redshift, it is important to consider that the accuracy of the weak gravitational lensing signal is more sensitive to the resolution a low redshifts than at high redshifts. The lensing signal in a given matter shells depends on the weighted sum of the matter fields of the shells with lower redshifts. Since shells at low redshifts are along the line-of-sight of more shells than high-redshift shells, prioritising thinner shells at low redshifts helps to reduce discretisation effects. We find that the following set of redshift limits for 19 shells are sufficient for the resolution needed for KiDS-1000, while still being efficient: $\{0, 0.04, 0.08, 0.12, 0.16, 0.2, 0.27, 0.34, 0.42, 0.5, 0.58, 0.66, 0.75, 0.88, 1.03, 1.19, 1.36, 1.55, 1.76, 2\}$ (see Figure 3.2).

3.2.3 Non-Limber Projection

Although the three-dimensional matter power spectrum characterises all cosmological dependence of the matter fields in our simulation, it does not take into consideration the geometry of the different volumes considered in Section 3.2.2. To determine the correlations between the matter fields of comoving spherical shells, we define the two-dimensional angular matter power spectra for a given cosmology, $\mathbf{\Theta}$, and a given shell combination (ij) as follows

$$C^{(ij)}_{\delta\delta,\ell}(\mathbf{\Theta}) = \langle \tilde{\delta}^{(i)}_{\ell m}(\mathbf{\Theta})\, \tilde{\delta}^{(j)*}_{\ell m}(\mathbf{\Theta}) \rangle, \tag{3.4}$$

where $\tilde{\delta}^{(i)}_{\ell m}(\mathbf{\Theta})$ represents the harmonic coefficients defined through a spherical harmonic transform of the projection of the spin-0 matter field, $\delta^{(i)}$, as given by (Zaldarriaga & Seljak, 1997; Reinecke, 2011),

$$\tilde{\delta}^{(i)}_{\ell m}(\mathbf{\Theta}) = \int \mathrm{d}^2\boldsymbol{\theta}\, \delta^{(i)}(\boldsymbol{\theta};\mathbf{\Theta})\, {}_0Y^*_{\ell m}(\boldsymbol{\theta}), \tag{3.5}$$

where ${}_0Y_{\ell m}(\boldsymbol{\theta})$ are the spherical harmonics for a spin value of 0 as a function of the two-dimensional spatial sky position $\boldsymbol{\theta}$. The projected matter field, $\delta^{(i)}(\boldsymbol{\theta};\mathbf{\Theta})$, can

in turn be defined with respect to the underlying three-dimensional matter overdensity field, $\delta(\boldsymbol{\theta}, z; \boldsymbol{\Theta})$, as

$$\delta^{(i)}(\boldsymbol{\theta}; \boldsymbol{\Theta}) = \sum_{\ell m} \tilde{\delta}^{(i)}_{\ell m}(\boldsymbol{\Theta})\, {}_0Y_{\ell m}(\boldsymbol{\theta}) = \int dz W^{(i)}(z; \boldsymbol{\Theta})\, \delta(\boldsymbol{\theta}, z; \boldsymbol{\Theta}), \tag{3.6}$$

where $W^{(i)}(z; \boldsymbol{\Theta})$ is the weight function of the given shell as defined in Equation (3.2).

The three-dimensional matter overdensity field's cosmology dependence can be quantified as follows

$$\langle \delta(\boldsymbol{k}, z; \boldsymbol{\Theta})\, \delta^*(\boldsymbol{k}', z'; \boldsymbol{\Theta}) \rangle = (2\pi)^3 \delta_{\mathrm{D}}(\boldsymbol{k} - \boldsymbol{k}')\, P_{\delta,\mathrm{nl}}(k, z, z'; \boldsymbol{\Theta}), \tag{3.7}$$

where $\delta(\boldsymbol{k}, z; \boldsymbol{\Theta})$ is the two-dimensional Fourier transform of the two-dimensional matter overdensity field at z, $\delta(\boldsymbol{\theta}, z, ; \boldsymbol{\Theta})$ and $\delta_{\mathrm{D}}(\boldsymbol{k} - \boldsymbol{k}')$ is the Dirac delta function.

If we combine the relations in Equations (3.4) to (3.7) and, for the sake of simplicity, assume that $\Omega_k = 0$, i.e. $f_k(z; \boldsymbol{\Theta}) = \chi(z; \boldsymbol{\Theta})$, we can characterise the cosmology dependence of $C^{(ij)}_{\delta\delta}(\ell; \boldsymbol{\Theta})$ directly with respect to the three-dimensional matter power spectrum, $P_{\delta,\mathrm{nl}}(k, z, z'; \boldsymbol{\Theta})$, using the following relation

$$C^{(ij)}_{\delta\delta}(\ell; \boldsymbol{\Theta}) = \frac{2}{\pi} \int d\chi\, W^{(i)}(z[\chi]; \boldsymbol{\Theta}) \int d\chi'\, W^{(j)}(z'[\chi]; \boldsymbol{\Theta}) \int dk\, k^2\, P_{\delta,\mathrm{nl}}(k, z[\chi], z'[\chi]; \boldsymbol{\Theta})\, j_\ell(k\chi)\, j_\ell(k\chi'), \tag{3.8}$$

where $\chi' \equiv z[\chi']$, $j_\ell(k\chi)$ are spherical Bessel functions of rank ℓ, and $\ell \in \mathbb{Z}^{0+}$. We can simplify Equation (3.8) by making the geometric approximation, i.e. taking the geometric mean of two equal-time power spectra (Castro et al., 2005; Kitching & Heavens, 2017; Kilbinger et al., 2017), such that

$$P_{\delta,\mathrm{nl}}(k, z[\chi], z'[\chi']; \boldsymbol{\Theta}) = \sqrt{P_{\delta,\mathrm{nl}}(k, z[\chi]; \boldsymbol{\Theta})\, P_{\delta,\mathrm{nl}}(k, z'[\chi']; \boldsymbol{\Theta})}, \tag{3.9}$$

which has been shown to be an accurate approximation both in the linear and non-linear regimes (Kitching & Heavens, 2017).

At this point, it is common to also make the so-called Limber approximation (Limber, 1953; Kaiser, 1992) which consists in a Taylor expansion around the approximate maximum of the Bessel function at $k\chi = \ell + 1/2$, and only considering the first-order term. This approximation is reasonably accurate in the autocorrelations, $C_{\delta\delta}^{(ii)}(\ell;\boldsymbol{\Theta})$, but depending on the thickness of the shells, the Limber approximation can lead to substantial biases, especially, at large scales. However, when making the Limber approximation, all off-diagonal angular power spectra, $C_{\delta\delta}^{(ij)}$, where $i \neq j$, will be zero as long as their associated weight functions, $W^{(i)}(z[\chi];\boldsymbol{\Theta})$, are non-overlapping, as is the case for the weights constructed in Equation (3.2). This can be seen by looking at the full Taylor expansion of Equation (3.8) as given by (LoVerde & Afshordi, 2008)

$$
\begin{aligned}
C_{\delta\delta}^{(ij)}(\ell) = \int \frac{d\chi}{\chi} W^{(i)}(z[\chi])\, W^{(j)}(z[\chi])\, P_{\delta,\mathrm{nl}}\left(\frac{\ell+1/2}{\chi}, z[\chi]\right) \\
\left\{1 - \frac{1}{(\ell+1/2)^2}\left[\frac{\chi^2}{2}\left(\frac{W^{(i)''}(z[\chi])}{W^{(i)}(z[\chi])} + \frac{W^{(j)''}(z[\chi])}{W^{(j)}(z[\chi])}\right)\right.\right. \\
\left.\left. + \frac{\chi^3}{6}\left(\frac{W^{(i)'''}(z[\chi])}{W^{(i)}(z[\chi])} + \frac{W^{(j)'''}(z[\chi])}{W^{(j)}(z[\chi])}\right)\right] + \mathscr{O}\left([\ell+1/2]^{-4}\right)\right\}, \quad (3.10)
\end{aligned}
$$

where each apostrophe, $'$, denotes a partial derivative with respect to χ, i.e. $\partial/\partial\chi$, while we also omit the explicit cosmological dependence of the power spectrum and the weight functions here for the sake of clarity. The non-Limber integral therefore also contributes to the autocorrelations, where $|i-j| = 0$, particularly for small ℓ. If not considered this can induce a higher than percentage-level bias in the angular power spectra of the shells which could significantly bias cosmological constraints from KiDS-1000. At the same time, we find that for the cross-correlations of adjacent shells, such that $|i-j| = 1$, the angular power can be of the order of $10^{-1}C_{\delta\delta}^{(ii)}$ or less. In this case, the simulated $\delta^{(i)}(\boldsymbol{\theta};\Theta)$ would have a percentage-level bias, while being void of any large-scale structure along the line-of-sight beyond the scales of individual shells. To avoid this, we choose to calculate the full non-Limber projection as given by Equation (3.8). The correlations between shell pairs which are

not immediate neighbours, i.e. $|i-j|>1$, contribute less than 0.1% to the overall angular power of a shell, so their effect is undetectable for KiDS.

Calculating the full non-Limber integral is a computationally expensive endeavour as it involves three numerical integrals over highly oscillatory Bessel functions. To address this, we present the LEVIN module[10]. This is a novel Python module written in C++ which implements the Levin integration method for oscillatory functions (Levin, 1996) in the context of non-Limber integrals for weak lensing and galaxy clustering.

The module allows one to divide the domain of the angular power spectra into three regions: a non-Limber domain, a second-order extended Limber domain and a Limber domain. In this work, we choose to split the calculation of $C_{\delta\delta}^{(ij)}(\ell)$ into two domains: non-Limber ($1<\ell\leq\ell_{\mathrm{max,nL}}$) and second-order extended Limber ($\ell_{\mathrm{max,nL}}<\ell\leq 30,000$). Over the former domain, LEVIN numerically integrates the expression shown in Equation (3.8) for the given matter power spectrum and the weights. Over the latter domain, we perform the numerical integration of Equation (3.10) up to the second-order term in $(\ell+1/2)^{-1}$. This reduces the computational resources needed for the integration without any substantial loss of accuracy as long as $\ell_{\mathrm{max,nL}}$ is sufficiently large, since the residuals scale with $(\ell+1/2)^{-4}$ (LoVerde & Afshordi, 2008). With the shells shown in Figure 3.2 weighted with Equation (3.2), we obtain angular power spectra, $C_{\delta\delta}^{(ij)}(\ell;\boldsymbol{\Theta})$, shown in Figure 3.5.

With this configuration, we find that we can obtain computationally efficient and accurate (consistent with CAMB within 0.1%) angular power spectra which describe the correlations between the matter fields within each spherical shell of the simulation. This is in line with testing conducted on the LEVIN module in previous applications (Zieser & Merkel, 2016; Spurio Mancini et al., 2018a,b; Baleato Lizancos & White, 2023).

3.2.4 Log-normal Matter Field Simulations

With the input angular power spectra for a given cosmology computed, the next step in the forward simulations involves sampling random matter fields consistent with

[10] LEVIN; `https://github.com/rreischke/nonLimber_matter_shells`

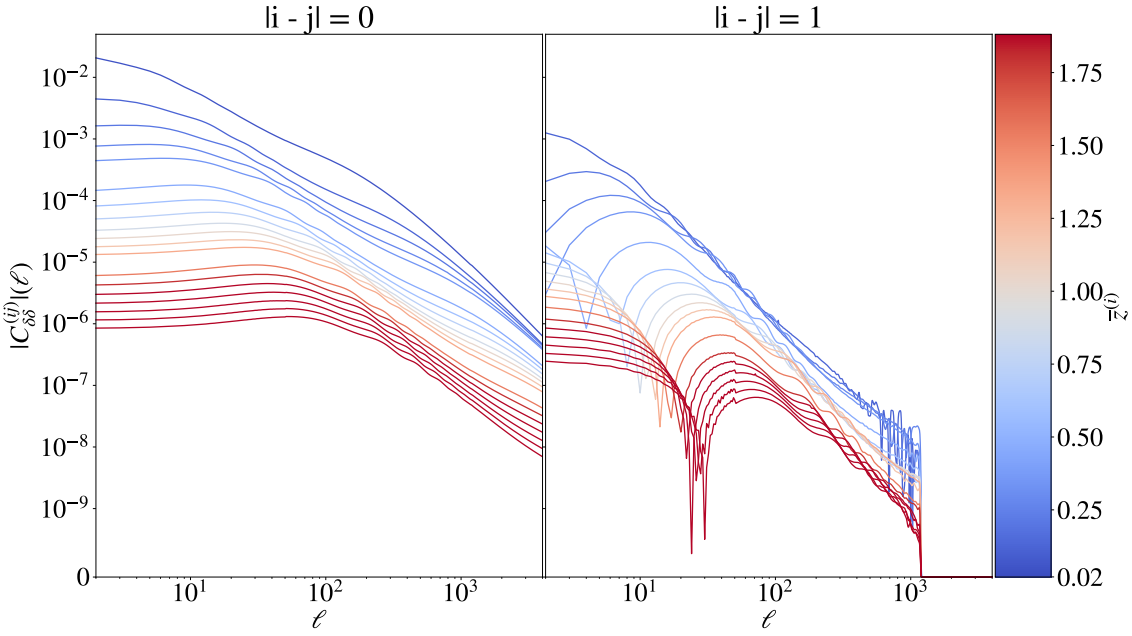

Figure 3.5: Plot of the two-dimensional angular matter power spectra, $C_{\delta\delta}^{(ij)}(\ell)$, projected with LEVIN describing the correlations between the large-scale structure within a set of 19 concentric shells spaced along the line-of-sight as shown in Figure 3.2. The left panel shows the autocorrelations of all shells, i.e. $|i-j|=0$. The right panel shows the correlations of each shell with its nearest neighbour, i.e. $|i-j|=1$ ($|i-j|>1$ are not shown as they do not have a large effect on matter fields within each shell, see Tessore et al. 2023). The colour of each line is given by the mean redshift of the i^{th} bin, $\bar{z}^{(i)}$. The underlying linear three-dimensional matter power spectrum is based on flat ΛCDM assuming $\Omega_c = 0.05$, $\Omega_b = 0.28$, $\sigma_8 = 0.79$, $S_8 = 0.84$ and $H_0 = 67$ km s^{-1} Mpc^{-1} calculated using CAMB (Lewis et al., 2000; Lewis & Challinor, 2002; Howlett et al., 2012), while the non-linear contribution is calculated using HMCODE-2016 (Mead et al., 2015, 2016) assuming $A_{\text{bary}} = 3.1$. The non-Limber projection done by LEVIN assumes $\ell_{\text{max,nL}} = 1200$ which causes the $C_{\delta\delta}^{(ij)}$ with $|i-j|>0$ go to a value of zero for $\ell > \ell_{\text{max,nL}}$.

these correlations. For this, we use GLASS(Tessore et al. 2023). This framework allows one to create efficient and accurate random field simulations of large-scale structure matter fields as well as the associated weak gravitational lensing signals. Within the GLASS suite, many different choices can be made, while it also allows one to sample the angular positions, redshifts, shapes and shears of galaxies. Nevertheless, in this analysis, we only choose to use it to simulate the underlying matter overdensity fields, $\delta^{(i)}(\boldsymbol{\theta};\boldsymbol{\Theta})$, as well as the associated convergence fields for each shell, $\kappa^{(i)}(\boldsymbol{\theta};\boldsymbol{\Theta})$. These are then integrated along the line-of-sight to give

the convergence field for a source within a given shell, $\kappa^{(i)}$.

Once the angular correlation functions, $C_{\delta\delta}^{(ij)}(\ell;\mathbf{\Theta})$, are defined, they can be transformed into spatial coordinates, $C_{\delta\delta}^{(ij)}(\theta;\mathbf{\Theta})$,, so one can now apply spatial transformations to the field. In general, we can transform any field, $X(\boldsymbol{\theta})$, into a transformed field, $Y(\boldsymbol{\theta}) = f[X(\boldsymbol{\theta})]$. Since we are interested in sampling cosmological fields describing large-scale structure, it is important to ensure that they are consistent with the cosmological principle. Therefore, any field, δ, must be homogeneous, i.e. invariant under rotational transformations. One such type of field is a Gaussian field which by construction is homogeneous and therefore has the same mean and variance at all points in space. In addition, Gaussian fields have the advantage that they are easy to sample and they can be sampled recursively for each shell for computational efficiency (see Tessore et al. 2023 for a more detailed description).

However, Gaussian fields do not capture the full complexity of large-scale structure as by definition they have zero skewness or kurtosis or beyond, i.e. they are constructed such that they can be fully characterised with only a mean and a variance. However, realistic matter overdensity fields have higher-order variations as filamentary structure and voids tend to appear under the influence of gravity and baryonic effects, even if the seeds that originated the overdensities were Gaussian distributed.

It has been found that a good approximation to such fields is a log-normally distributed random field (which can be derived from a Gaussian random field). In fact, log-normal fields are even able to produce reasonable approximations to the three-point and four-point statistics measured from N-body simulations (Hall & Taylor, 2022; Piras et al., 2023). For this reason, log-normal random fields are a common approximation used for matter overdensity fields (Coles & Jones, 1991; Böhm et al., 2017; Abramo et al., 2016, 2022) as well as for convergence fields (Hilbert et al., 2011; Clerkin et al., 2017; Giocoli et al., 2017; Gatti et al., 2020).

Consequently, we are interested in sampling log-normal fields to model the matter overdensity fields in each shell from Gaussian random fields. In theory, one

could obtain these by taking the angular power spectra of the shells and assuming that they describe the correlations between log-normal fields, $\delta^{(i)}_{\log}(\boldsymbol{\theta};\boldsymbol{\Theta})$, as follows

$$\langle \delta^{(i)}_{\log}(\boldsymbol{\theta};\boldsymbol{\Theta})\, \delta^{(j)*}_{\log}(\boldsymbol{\theta};\boldsymbol{\Theta})\rangle = C^{(ij)}_{\delta\delta}(\theta;\boldsymbol{\Theta}). \tag{3.11}$$

If we then define the angular power spectrum of a Gaussian random field, $G^{(ij)}_{\delta\delta}(\theta;\boldsymbol{\Theta})$, as follows

$$\langle \delta^{(i)}_{\mathrm{G}}(\boldsymbol{\theta};\boldsymbol{\Theta})\, \delta^{(j)*}_{\mathrm{G}}(\boldsymbol{\theta};\boldsymbol{\Theta})\rangle = G^{(ij)}_{\delta\delta}(\theta;\boldsymbol{\Theta}). \tag{3.12}$$

By defining the mean and the variance for $G^{(ij)}_{\delta\delta}(\theta;\boldsymbol{\Theta})$, one can then sample $\delta^{(i)}_{\mathrm{G}}(\boldsymbol{\theta};\boldsymbol{\Theta})$ for each shell. To achieve this with numerical methods, the two-dimensional spherical shell of each layer is discretised using HEALPIX pixels (Górski et al., 2005). This means that $\boldsymbol{\theta}$ is discretised such that it can be mapped to a linearised variable given by θ_m where $m \in \{1,2,3,...,12N^2_{\mathrm{side}}\}$ and N_{side} is the HEALPIX resolution parameter. To obtain the final matter overdensity fields, we use the following expression (Coles & Jones, 1991; Kayo et al., 2001; Hilbert et al., 2011; Xavier et al., 2016),

$$\delta^{(i)}_{\log}(\theta_m;\boldsymbol{\Theta}) = e^{\delta^{(i)}_{\mathrm{G}}(\theta_m;\boldsymbol{\Theta})} - \lambda, \tag{3.13}$$

where λ is the shift of the log-normal distribution which is assumed to be $\lambda = 1$ for matter fields.

In practice, the simulations do not actually calculate these quantities in the order that has been laid out here. In reality, in the numerical calculations, it is not possible to sum/integrate over an infinite range of angular scales, ℓ. When instead setting a maximum angular scale, $\ell_{\max}$, it is possible that with $G^{(ij)}_{\delta\delta}(\ell > \ell_{\max};\boldsymbol{\Theta}) = 0$, one can still obtain $|C^{(ij)}_{\delta\delta}(\ell > \ell_{\max};\boldsymbol{\Theta})| > 0$ which can bias the sampled log-normal fields (Xavier et al., 2016). To avoid this and make the sampling more efficient, within GLASS, an initial $G^{(ij)}_{\delta\delta}(\ell;\boldsymbol{\Theta})$ is guessed which is transformed using the previously described prescription. The associated log-normal angular power spectrum is then compared to the given $C^{(ij)}_{\delta\delta}(\ell;\boldsymbol{\Theta})$, which gives a residual. Subsequently,

the guessed $G^{(ij)}_{\delta\delta}(\ell;\boldsymbol{\Theta})$ is updated using the Gauss-Newton algorithm, such that the residual is minimised between iterations and the result converges to an accurate $G^{(ij)}_{\delta\delta}(\ell;\boldsymbol{\Theta})$ over a finite ℓ range. For more details about this calculation, see Tessore et al. 2023.

The layers of concentric log-normal matter overdensity fields within each shell give a full description of the large-scale structure. The resolution of this structure will of course be limited by the size of the shells along the line-of-sight, while being limited by the resolution of the discretised two-dimensional pixels on the sphere. However, as discussed with regards to the resolution of shells in Section 3.2.2, if one chooses sufficiently large $\ell_{\rm max}$ up to which the input angular power spectra are calculated, one can accurately sample $\delta^{(i)}_{\rm log}(\theta_m;\boldsymbol{\Theta})$ up to an $N_{\rm side} \sim 0.5\ell_{\rm max}$ (Leistedt et al., 2013; Alonso et al., 2019).

As we intend to make use of these simulations for weak gravitational lensing, the next step in the forward simulations is to model how the image of galactic sources within a given shell will be distorted by weak gravitational lensing due to the matter overdensities the light encounters along the line-of-sight to the observer. Here, the geometry of the simulations aids us again, since the concentric volumes in comoving distance that each shell makes up intrinsically allow one to trace all possible light-cones emanating from the observer.

We start from the definition of the convergence for a source located at $\boldsymbol{\theta}$ and z along a continuous line-of-sight under the Born approximation (Schneider, 2005),

$$\kappa(\boldsymbol{\theta},z;\boldsymbol{\Theta}) = \frac{3\Omega_{\rm m}}{2}\int_0^z {\rm d}z' \frac{f_k(z';\boldsymbol{\Theta})\,[f_k(z';\boldsymbol{\Theta}) - f_k(z;\boldsymbol{\Theta})]}{f_k(z;\boldsymbol{\Theta})}\frac{1+z'}{E(z';\boldsymbol{\Theta})}\,\delta(\boldsymbol{\theta},z';\boldsymbol{\Theta}), \quad (3.14)$$

where $\Omega_{\rm m}$ is the matter density fraction at $z=0$, and $\Omega_{\rm m} \in \boldsymbol{\Theta}$. We then discretise the continuous two-dimensional matter overdensity field along over $\boldsymbol{\theta}$ using HEALPIX pixels, while also discretising along the line-of-sight using Equation (3.6) and defining a mean redshift, $\overline{z}^{(i)}$, for a given weighted shell as follows

$$\bar{z}^{(i)}(\mathbf{\Theta}) = \frac{\int \mathrm{d}z\, z\, W^{(i)}(z;\mathbf{\Theta})}{\int \mathrm{d}z\, W^{(i)}(z);\mathbf{\Theta}}. \tag{3.15}$$

If we then assume that the integral over the geometric and the weight factors in Equation (3.14) is approximately equal to the factors evaluated at the mean of each weighted shell, $\bar{z}^{(i)}$, we can rewrite Equation (3.14) as follows

$$\kappa^{(i)}(\theta_m;\mathbf{\Theta}) = \frac{3\Omega_\mathrm{m}}{2} \sum_{j=0}^{i-1} \frac{f_k(\bar{z}^{(j)};\mathbf{\Theta})\,[f_k(\bar{z}^{(j)};\mathbf{\Theta}) - f_k(\bar{z}^{(i)};\mathbf{\Theta})]}{f_k(\bar{z}^{(i)};\mathbf{\Theta})} \frac{1+\bar{z}^{(j)}}{E(\bar{z}^{(j)};\mathbf{\Theta})}\, w^{(j)}(\bar{z}^{(j)};\mathbf{\Theta})\, \delta^{(j)}(\theta_m;\mathbf{\Theta}), \tag{3.16}$$

where $w^{(j)}(\bar{z}^{(j)};\mathbf{\Theta})$ is a lensing weight defined as

$$w^{(j)}(\bar{z}^{(j)};\mathbf{\Theta}) = [1/W^{(j)}(\bar{z}^{(j)};\mathbf{\Theta})] \int \mathrm{d}z W^{(j)}(z;\mathbf{\Theta}). \tag{3.17}$$

From now on we take $\delta^{(j)}(\theta_m;\mathbf{\Theta}) \equiv \delta^{(j)}_{\log}(\theta_m;\mathbf{\Theta})$. To accelerate this calculation within GLASS, it makes use of the fact that for all Robertson-Walker space-times, transverse comoving distances will scale in such a way that one can write down a recurrence relation from one interval to the next along the line-of-sight (Schneider, 2016). This gives a recurrence relation for the convergence field, $\kappa^{(i)}$, within shells where $i \geq 2$, which only depends on $\kappa^{(i-1)}$, $\kappa^{(i-2)}$ and $\delta^{(i-1)}$ (see Tessore et al. 2023 for details).

Such a relation allows one to calculate the convergence fields for all shells, while only holding three fields in memory at a given time. This greatly reduces the amount of computational resources needed in order to create the large amount of realisations at different cosmologies, $\mathbf{\Theta}$, needed for this analysis. Additionally, we now have all the ingredients to sample galaxies and to shear their shapes, while simultaneously having the flexibility to add any field-level systematics which may be relevant.

3.2.5 Intrinsic Alignments: Non-Linear Alignment Model

An important effect to consider when modelling weak gravitational lensing is intrinsic alignments (IAs). This refers to the fact that correlations between source galaxy shapes in different parts of the sky may not only be caused by weak gravitational lensing due to dark matter in the foreground. This happens due to two local processes which occurs irrespectively of weak gravitational lensing: tidal alignments and tidal torquing. Tidal alignments occur when the gravitational forces from neighbouring matter distributions cause the galaxies to align with the surrounding matter density field. This alignment leads to a coherent stretching or compression of the galaxy shapes as function of the local matter overdensity field. Hence, he intrinsic ellipticities of source galaxies are systematically aligned with the underlying large-scale structure within which they form. Thus, the measured cosmic shear signal will be biased by these intrinsic alignments (Heavens et al., 2000a; King & Schneider, 2002; Heymans & Heavens, 2003; Bridle & King, 2007).

To model this effect, we follow the prescription set out in Tessore et al. (2023). In this prescription, for a given intrinsic alignment model, we define an associated effective convergence field, $\kappa_{\rm IA}^{(i)}$. This $\kappa_{\rm IA}^{(i)}$ is a useful construct which describes the contribution to the observed weak lensing signal from IA under the Born approximation. Alternatively, an IA model could also directly be implemented into the dark matter distribution of the matter fields which could an interesting avenue for future work.

Based on the fact that for the IA model we take into consideration $\kappa_{\rm IA}^{(i)} \propto \delta^{(i)}$, we can assume that, for a given shell, this field can be added in a linear fashion to the underlying convergence field due to weak lensing, as given by

$$\kappa^{(i)}(\theta_m;\mathbf{\Theta}) \rightarrow \kappa^{(i)}(\theta_m;\mathbf{\Theta}) + \kappa_{\rm IA}^{(i)}(\theta_m;\mathbf{\Theta}). \tag{3.18}$$

For simplicity and consistency with typical modelling assumptions in weak lensing surveys such as KiDS-1000, we choose to model $\kappa_{\rm IA}^{(i)}$ with the Non-Linear Alignment (NLA) model (Catelan et al., 2001; Hirata et al., 2004; Bridle & King, 2007). The NLA model assumes that the bias in the shear signal from IAs is linearly de-

pendent on the projected local tidal field. The "non-linear" part of the NLA model then simply refers to the fact that it has been found that the modelling of IAs is more accurate when modelling the underlying large-scale structure using a non-linear matter power spectrum, rather than a linear one (Bridle & King 2007, similar to our approach described in Section 3.2.1). This means that it is also proportional to the local matter overdensity field, which is given by (Hirata et al., 2004),

$$\kappa_{\mathrm{IA}}^{(i)}(\theta_m;\mathbf{\Theta}) = -A_{\mathrm{IA}}\frac{C_1\,\Omega_{\mathrm{m}}\,\overline{\rho}_{\mathrm{cr}}(\overline{z}^{(i)};\mathbf{\Theta})}{D(\overline{z}^{(i)};\mathbf{\Theta})}\,\delta^{(i)}(\theta_m;\mathbf{\Theta}), \tag{3.19}$$

where A_{IA} is the intrinsic alignments amplitude (which we are treating as a nuisance parameter that is also sampled, see Section 3.3), C_1 is a normalisation constant which we set to $C_1 = 5\times10^{-14}h_0^{-2}M_\odot^{-1}\mathrm{Mpc}^3$ in accordance with the IA measurements at low redshifts by SuperCOSMOS (Brown et al., 2002), $\overline{\rho}_{\mathrm{cr}}(\overline{z}^{(i)};\mathbf{\Theta})$ is the mean critical matter density as a function of redshift and $D(\overline{z}^{(i)};\mathbf{\Theta})$ is the linear growth factor normalised to be unity at $z=0$. Note that in Equation (3.19), we do not consider an explicit redshift dependence which is often expressed as a power-law term (Joachimi et al., 2011a). We omit this term, as it not considered in the fiducial KiDS-1000 analysis (Asgari et al., 2021; van den Busch et al., 2022), and it has been found that at least for $z<1$, the IA signals do not vary largely with redshift (Fortuna et al., 2021).

Nevertheless, the NLA model still comes with some caveats. Mainly, it neglects non-linearities which may not be captured in the underlying matter power spectrum, such as source density weighing. The former can in fact produce a signal in the shape measurements comparable to the correction due to non-linear structure growth (Krause et al., 2016). Despite these limitations, the NLA model has been found to be accurate enough to model the IA signal seen by past and current weak lensing surveys (Joachimi et al., 2011a; Blazek et al., 2011; Heymans et al., 2013; Krause et al., 2016; Hilbert et al., 2017), including KiDS-1000 (Fortuna et al., 2021).

3.2.6 Galaxy Positions and Redshifts

With the cosmological dependence of the underlying large-scale structure characterised and astrophysical effects such as intrinsic alignments and baryonic feedback included, the forward simulations still need to include observational biases from our instruments. We choose to apply these on the level of galaxies, i.e. by biasing the sampled galaxy positions, redshifts and shapes/shears in accordance with our models for the observational systematic effects.

When sampling the galaxies' positions on the sky, we take into consideration the survey footprint, $\Omega_{\rm survey}$, the galaxy number density which the survey can observe, $n_{\rm gal}$, and the variability in the latter as a function of redshift and observational depth across the sky, i.e. as a function of the local observing conditions. When sampling the galaxies' redshifts, we consider the redshift distributions calibrated from photometry for a given survey, $n(z)$. Lastly, when determining the observed galaxy shapes, we consider the intrinsic ellipticity dispersion of the galaxies, σ_ϵ, the multiplicative and additive shear bias, while also including the effects caused by the variation in the point spread function of the instrument as a function of the position in the sky.

Since we are interested in simulating a photometric survey, we must model the resolution along the line-of-sight realistically. This means that rather than modelling large-scale structure observables across $N_{\rm shells}$ concentric shells, they ought to be modelled across $N_{\rm tomo}$ tomographic bins. The tomographic bins of a given survey are defined by finite usually non-overlapping domains in photometric redshift, $z_{\rm ph}$, in which all observed galaxies are grouped, such that $z_p \leq z_{\rm ph} < z_{p+1}$, $p \in \{1,2,3,...,N_{\rm tomo}\}$[11]. The measured photometric redshift of a galaxy does not necessarily agree with its true/spectroscopic redshift, z. This means that galaxies which are grouped within a given photometric redshift bin, p, may not actually be located in this redshift range. To account for this, one can characterise a conditional redshift distribution for each tomographic bin $P(z|z_{\rm ph})$ by mapping the measured $z_{\rm ph}$ to known spectroscopic redshift measurements which have a comparatively low

[11] Note that we use the indices p and q for tomographic bins and the indices i and j for underlying shells.

Bin	$\overline{n}_{\rm gal}$ [arcmin^{-2}]	$\overline{\sigma}_\varepsilon/\sqrt{2}$	$a_{n_{\rm gal}}$	$b_{n_{\rm gal}}$	a_{σ_ε} $\times 10^3$	b_{σ_ε}
S1	0.62	0.27	-0.035	0.72	1.81	0.267
S2	1.18	0.26	-0.042	1.30	1.30	0.257
S3	1.85	0.28	-0.243	2.58	-0.86	0.280
S4	1.26	0.25	-0.250	2.01	-0.62	0.255
S5	1.31	0.27	-0.416	2.56	2.39	0.261

Table 3.1: Parameters used to sample galaxies and their shapes in line with the expectations for KiDS DR4 for each tomographic bin (from S1 to S5). $\overline{n}_{\rm gal}$ is the mean galaxy number density for a given tomographic bin, $\overline{\sigma}_\varepsilon/\sqrt{2}$ is the mean per-component shape dispersion, $a_{n_{\rm gal}}$ and $b_{n_{\rm gal}}$ are the slope and y-intercept, respectively, for the linear interpolation of the galaxy density as a function of the root-mean-square of the background noise in the KiDS catalogue, $\sigma_{\rm rms}$, according to Equation (3.30), while a_{σ_ε} and b_{σ_ε} are the parameters to linearly interpolate σ_ε from $\sigma_{\rm rms}$, according to Equation (3.31).

bias in the redshift estimate (Hildebrandt et al., 2021; van den Busch et al., 2022). Therefore, a given galaxy is in an associated pixel, m, shell, i, and tomographic bin, p. The probability of the galaxy being in tomographic bin, p, while it is sampled in shell, i, then depends on $P(z|z_{\rm ph})$.

In line with the KiDS-1000 analysis, we define the galaxy number density as follows

$$n_{\rm gal}^{(p)} \equiv \frac{1}{|\Omega_{\rm survey}|} \frac{\left[\left(1+\overline{M}^{(p)}\right)\sum_{i\in(p)} w_i\right]^2}{\sum_{i\in(p)}\left(1+\overline{M}^{(p)}\right)^2 w_i^2}, \tag{3.20}$$

where $|\Omega_{\rm survey}|$ is the effective area covered by the survey footprint, $\overline{M}^{(p)}$ is the mean multiplicative shear bias in tomographic bin p as calibrated in (Giblin et al., 2021), and w_i is the LENSFIT weight for a given galaxy i. Similarly, we define the galaxy intrinsic ellipticity dispersion, σ_ε, within a given tomographic bin as follows

$$\sigma_\varepsilon^{(p)2} \equiv \left(1+\overline{M}^{(p)}\right)^{-2} \frac{\sum_{i\in(p)} w_i^2\left(\varepsilon_{{\rm obs},i,1}^2 + \varepsilon_{{\rm obs},i,2}^2\right)}{\sum_{i\in(p)} w_i^2}, \tag{3.21}$$

where $\varepsilon_{\rm obs,1}$ and $\mathrm{i}\varepsilon_{\rm obs,2}$ are the two components of the observed galaxy ellipticity. Applying Equations (3.20) and (3.21) to KiDS DR4, we obtain the values shown in Table 3.1.

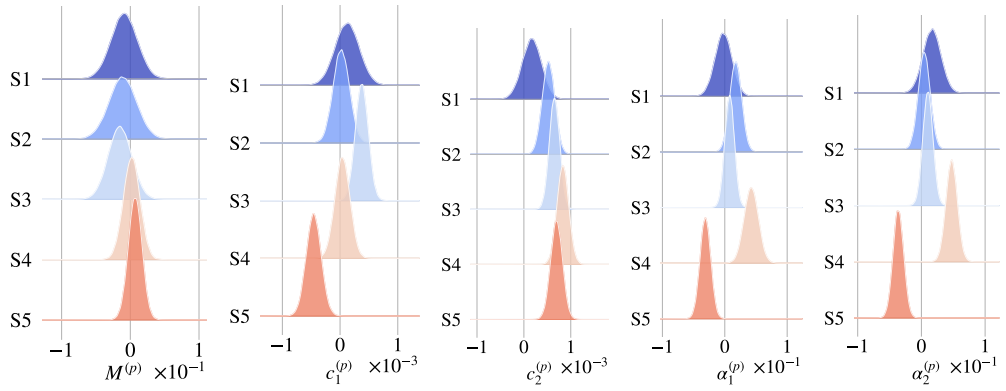

Figure 3.6: Plot of the distributions for each tomographic bin (S1 to S5) from which the shear bias parameters from KiDS-1000 (Giblin et al., 2021) shown in Equation (3.29) are sampled. The first panel from the left shows the multiplicative shear bias, $M^{(p)}$. The second and third panel show the real and imaginary part of the additive shear bias, $c_1^{(p)}$ and $c_2^{(p)}$, respectively. The fourth and fifth panel show the real and imaginary part of the amplitude of the shear bias due to variations in the point-spread function, $\alpha_1^{(p)}$ and $\alpha_2^{(p)}$, respectively.

Although it does not have detectable effects on the cosmic shear signal, the KiDS-SBI forward simulations sample galaxy positions such that they trace the underlying matter fields, so spatially varying systematics may be accurately applied. Using the SALMO[12] framework (Joachimi et al., 2021) for simulating galaxy catalogues, we Poisson-sample galaxies within each HEALPix pixel, m, on the sky using the following expectation for the galaxy counts

$$\langle N_m^{(i)(p)} \rangle (\mathbf{\Theta}) = w_m(\Omega_{\rm survey}) \left[1 + b^{(i)} \delta^{(i)}(\theta_m; \mathbf{\Theta}) \right] P_m(p|i)\, n_{{\rm gal},m}^{(p)} A_{{\rm pix},m}, \qquad (3.22)$$

where $w_m(\Omega_{\rm survey})$ is the weight for a given pixel m which we take to be unity if a pixel is within the survey's mask/footprint, $\Omega_{\rm survey}$, and zero if it is not, $b^{(i)}$ is galaxy bias parameter within a given shell which we assume to be $b^{(i)} = 1 \forall i$ for the purposes of this cosmic shear analysis, $P_m(p|i)$ is the probability of a galaxy in pixel m being detected within tomographic bin p while being within shell i (i.e. a discretised version of $P_m(z|z_{\rm ph})$), $n_{{\rm gal},m}^{(p)}$ is the observed galaxy number per unit area for a given tomographic bin p and a given pixel m (here the pixel dependence

[12] Speedy Acquisition for Lensing and Matter Observables; `https://github.com/Linc-tw/salmo`

accounts for the spatial variability of the detection limit of galaxies as a function of observing conditions, the number of exposures, etc.; see Section 3.2.8), and $A_{\mathrm{pix},m}$ is the area of pixel m. As $n_{\mathrm{gal},m}^{(p)}$ changes due to variations in observational depth, it is more likely/unlikely to observe high-redshift galaxies. Hence, a change in depth from one pixel to the next can change the observed redshift distribution (see Section 3.2.8).

Taking $w_m(\Omega_{\mathrm{survey}})$, $P_m(p|i)$ and $n_{\mathrm{gal},m}^{(p)}$ from the photometric survey's measurements, we can Poisson sample $N_m^{(i)(p)}(\boldsymbol{\Theta})$ given the expectation value shown in Equation (3.22) from which we can get the expected number of galaxy counts within each pixel of a tomographic bin p by taking

$$N_m^{(p)}(\boldsymbol{\Theta}) = \sum_{i=1}^{N_{\mathrm{shells}}} N_m^{(i)(p)}(\boldsymbol{\Theta}). \tag{3.23}$$

With that, we then randomly sample each galaxy's right ascension and declination within each pixel m assuming a uniform distribution. This means that below the scales of pixels, the simulations do not have any information on galaxy clustering.

To sample each galaxy's redshift, we can use the fact that we know exactly how many galaxies are in each shell i for a given tomographic bin p and pixel m from $N_m^{(i)(p)}(\boldsymbol{\Theta})$. We can therefore randomly assign an index i and p to each sampled galaxy in accordance with that number. We can then randomly sample a specific z_{spec} for each galaxy by assuming that the galaxies are uniformly distributed within a given shell. Again, this means that weak lensing and galaxy clustering measurements from the simulated galaxies do not contain any information about the large-scale structure along the line-of-sight below the scales of shells.

3.2.7 Galaxy Shears

In order to determine the *galaxy shears*, we begin by assigning intrinsic ellipticities. We define a given galaxy's shape through a complex ellipticity, ε, and any galaxy ellipticity in this work as follows

$$\varepsilon = \varepsilon_1 + \mathrm{i}\,\varepsilon_2, \tag{3.24}$$

where $\mathrm{i} \equiv \sqrt{-1}$. The intrinsic ε_1 and ε_2 of each galaxy are then sampled within SALMO as two independent normal random variates with zero mean and a variance equal to $\sigma^{(p)2}_{\varepsilon,m}$. As with the galaxy density, $n^{(p)}_{\mathrm{gal},m}$, the shape dispersion, $\sigma^{(p)}_{\varepsilon,m}$, can also vary with tomographic bin and as a function of the position on the sky due to the effects of the spatial variability of observational depth. See Section 3.2.8 for a more detailed discussion. The ellipticity of each source galaxy is then altered through weak gravitational lensing by the matter along the line-of-sight. The exact shape distortion is given by (Seitz & Schneider, 1996; Hu, 2000),

$$\varepsilon_{\mathrm{lensed}}(\boldsymbol{\Theta}) = \frac{\varepsilon_{\mathrm{int}} + g(\boldsymbol{\Theta})}{1 + g^*(\boldsymbol{\Theta})\varepsilon_{\mathrm{int}}}, \tag{3.25}$$

where $\varepsilon_{\mathrm{lensed}}(\boldsymbol{\Theta})$ is the lensed galaxy ellipticity, $\varepsilon_{\mathrm{int}}$ is the intrinsic galaxy ellipticity and $g(\boldsymbol{\Theta})$ is the reduced shear ($g \in \mathbb{C}$) which is given by

$$g(\boldsymbol{\Theta}) = \frac{\gamma(\boldsymbol{\Theta})}{1 - \kappa(\boldsymbol{\Theta})}, \tag{3.26}$$

where $\gamma(\boldsymbol{\Theta})$ is the shear factor ($\gamma \in \mathbb{C}$). The use of Equation (3.25) is an extension to the KiDS-1000 analysis which assumed the reduced shear approximation, i.e. $\varepsilon_{\mathrm{lensed}} \approx \varepsilon_{\mathrm{int}} + g$. In any case, this should cause a negligible change in the measured shear considering the sensitivity of KiDS-1000 (Joachimi et al., 2021). To calculate the shear field, it is useful to decompose the spin-0 convergence field into spherical harmonics as follows

$$\kappa^{(i)}_{\ell m}(\boldsymbol{\Theta}) = \int \boldsymbol{d}^2\boldsymbol{\theta}\, \kappa^{(i)}(\boldsymbol{\theta};\boldsymbol{\Theta})\, {}_0Y^*_{\ell m}(\boldsymbol{\theta}), \tag{3.27}$$

where $\kappa^{(i)}_{\ell m}(\boldsymbol{\Theta})$ are the harmonic coefficients of the convergence field of a given shell i. One can then use the following relation to define the shear field's harmonic coefficients, $\gamma^{(i)}_{\ell m}$, given by (Tessore et al., 2023),

$$\gamma^{(i)}_{\ell m}(\boldsymbol{\Theta}) = -\left(\frac{(\ell+2)(\ell-1)}{\ell(\ell+1)}\right)^{1/2} \kappa^{(i)}_{\ell m}(\boldsymbol{\Theta}). \tag{3.28}$$

The $\gamma_{\ell m}^{(i)}(\mathbf{\Theta})$ coefficients can then be used to calculate the spin-2 discrete shear field, $\gamma^{(i)}(\theta_m, \mathbf{\Theta})$. Having defined the effective $\kappa^{(i)}(\theta_m, \mathbf{\Theta})$ in Sections 3.2.4 and 3.2.5, while also defining $\gamma^{(i)}(\theta_m, \mathbf{\Theta})$, we can assign a value of κ and $\gamma = \gamma_1 + \mathrm{i}\,\gamma_2$ to each galaxy sampled according to the procedure described in Section 3.2.6. This is done by taking the value of $\kappa^{(i)}(\theta_m, \mathbf{\Theta})$ and $\gamma^{(i)}(\theta_m, \mathbf{\Theta})$ within the shell i and pixel m in which a given galaxy is located. Between neighbouring pixels, the values of $\kappa^{(i)}(\theta_m, \mathbf{\Theta})$ and $\gamma^{(i)}(\theta_m, \mathbf{\Theta})$ are linearly interpolated. We then use Equation (3.26) to calculate the associated reduced shear, $g(\mathbf{\Theta})$, as well as the lensed ellipticity, $\varepsilon_{\mathrm{lensed}}(\mathbf{\Theta})$, by combining g with the intrinsic ellipticity using Equation (3.25).

However, this lensed ellipticity, $\varepsilon_{\mathrm{lensed}}(\mathbf{\Theta})$, is not the observed shape measurement in a weak gravitational lensing survey such as KiDS-1000. Additionally, the shape measurement may be distorted by instrumentational effects or shape modelling inaccuracies. Some relevant effects are selection biases, noise biases from a low signal-to-noise ratio, biases in the galaxy weights, artefacts, non-linear CCD responses, asymmetries in the point spread function, etc. (see e.g. Mandelbaum 2018 for a review). To account for this, it is common to parametrically estimate any residual systematics which may be affecting the shape measurements (Hildebrandt et al., 2017, 2021; Zuntz et al., 2018; Giblin et al., 2021). Within the forward simulations, we use the first-order parametric expansion of the observed galaxy shapes given by Heymans et al. (2006) as follows

$$\varepsilon_{\mathrm{obs},i\in m}^{(p)} = \left(1 + M^{(p)}\right)\varepsilon_{\mathrm{lensed},i} + \alpha^{(p)}\varepsilon_{\mathrm{PSF},m} + \beta^{(p)}\delta\varepsilon_{\mathrm{PSF},m} + c_m^{(p)}, \tag{3.29}$$

where i is the index for a single galaxy within tomographic bin p and pixel m, $M^{(p)}$ is the multiplicative shear bias as measured for tomographic bin p, $\alpha^{(p)}$ gives the fraction of the PSF ellipticity which remains in the shear estimator, $\varepsilon_{\mathrm{PSF}}(m)$ is the local PSF measured within pixel m (see Figure 3.7 for a map of $\varepsilon_{\mathrm{PSF}}$ in KiDS-1000), $\beta^{(p)}$ gives the amplitude of the shear bias due to residuals which are not taken into account by the PSF model, $\delta\varepsilon_{\mathrm{PSF}}$ represents the residuals in question, and $c^{(p)}$ is the additive shear bias within a given tomographic bin p. Note that spatially vary-

ing multiplicative shear biases, $M^{(p)}$, can also arise, but previous work has found that the mean bias across all positions is sufficient to accurately model the effect (Kitching et al., 2019). In accordance with previous findings (Giblin et al., 2021), we take $\beta^{(p)} = 0 \forall p$. The amplitudes of the remaining systematics as calibrated by Giblin et al. (2021) are shown in Figure 3.6. It becomes apparent that these effects are relatively small which explains why only the multiplicative and additive bias is taken into consideration in the modelling on the fiducial KiDS-1000 analysis (Giblin et al., 2021; Joachimi et al., 2021; Asgari et al., 2021; Heymans et al., 2021). Additionally, the measured values of $M^{(p)}$, $\alpha^{(p)}$ and $c^{(p)}$ have associated uncertainties. To take this uncertainty into account and make sure it is represented in the effective likelihood, rather than using a fixed value, we randomly sample different values for $M^{(p)}$, $\alpha^{(p)}$ and $c^{(p)}$ for each realisation of the forward simulations from a normal distribution as shown in Figure 3.6. Note that we do not explicitly marginalise over these parameters, but as they are sampled randomly from predetermined probability density distributions, so that their associated uncertainty is propagated into the effective likelihood. This is similar to how the seed for each log-normal random field and the random galaxy positions which populate the associated matter fields is varied from one forward simulation to the next, so that cosmic variance enters the effective likelihood as well.

As a result, for a given set cosmological parameters, $\mathbf{\Theta}$, the KiDS-SBI forward simulations produce a galaxy catalogue containing the galaxies' position on the sky, their spectroscopic redshift, the tomographic bin in which they are detected and their observed ellipticities, while taking into account many relevant systematics which are considered in the modelling of weak gravitational lensing measurements.

3.2.8 Variable Depth

As alluded to in Sections 3.2.6 and 3.2.7, when randomly sampling galaxies and their intrinsic ellipticities, the parameters which charaterise the probability distributions depend on the location on the sky, i.e. the pixel m in question. Specifically, the number of galaxies which is sampled (i.e. "observed") within a given pixel depends on the average galaxy density, $n^{(p)}_{\mathrm{gal},m}$, measured within that pixel. The galaxy

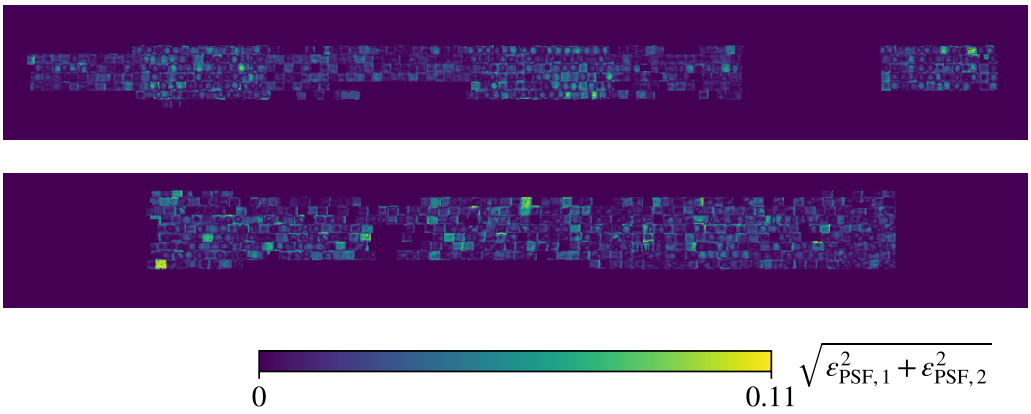

Figure 3.7: Cartesian spatial map ($N_{\mathrm{side}} = 1024$) of the observed magnitude of the point-spread function ellipticities, $|\varepsilon_{\mathrm{PSF}}| = \sqrt{\varepsilon^2_{\mathrm{PSF},1}+\varepsilon^2_{\mathrm{PSF},2}}$, throughout the KiDS-1000 North field in the upper panel and the KiDS-1000 South field in the lower panel. $\varepsilon_{\mathrm{PSF}}$ is added to the lensed galaxy shapes in the forward simulations within KiDS-SBI in accordance with Equation (3.29).

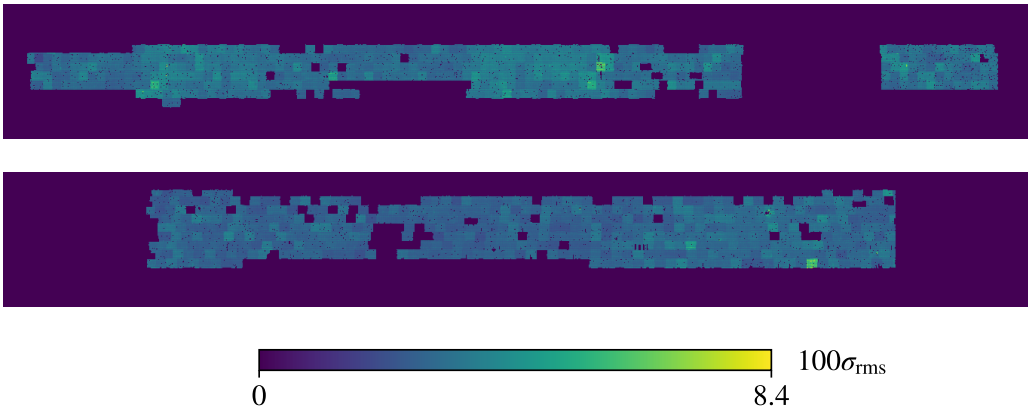

Figure 3.8: Cartesian spatial map ($N_{\mathrm{side}} = 1024$) of root-mean-square of the observed background noise, σ_{rms}, throughout the KiDS-1000 North field in the upper panel and the KiDS-1000 South field in the lower panel.

density varies from pixel to pixel as it depends on the observational depth of the survey in different parts of the sky, and on anisotropies and time variations in the atmospheric seeing. Since many parts of the sky are observed at different times, as is the case for ground-based telescopes such as the VLT Survey Telescope and the Visible and the Infrared Survey Telescope for Astronomy (VISTA) used for KiDS-1000, the atmospheric conditions and background light will be different for each pointing for a given exposure time. Thus, the depth of observations may change as seeing, the point spread functions and/or the background flux varies (for example,

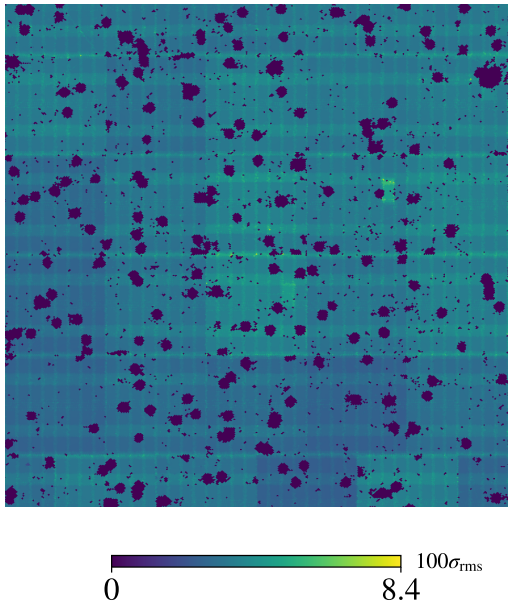

Figure 3.9: Spatial map ($N_{\rm side} = 4096$) in a Cartesian projection of root-mean-square of the observed background noise, $\sigma_{\rm rms}$, for a $5^\circ \times 5^\circ$ patch of the KiDS-1000 North field at a right ascension (RA) of 180° and a declination (DEC) of 0°.

due to zodiacal light or galactic absorption).

In addition, surveys such as KiDS often have overlapping pointings in order to ensure that the footprint is observed without gaps. However, this implies that galactic sources which happen to be located near the edge of a pointing will be observed more often in different overlapping fields than a source located in the centre of a pointing. Consequently, near the edge of a pointing there will tend to be a higher signal-to-noise ratio which allows for deeper observations. Since these systematic effects in the observational depth can happen at fixed scales and/or have certain periodicities, they can induce significant systematic effects into galaxy clustering and weak gravitational lensing signals. For cosmic shear in KiDS-1000, the bias can be near 1% in the signal and an average of $\sim 10\%$ in the standard deviation of cosmic shear observables (Heydenreich et al., 2020; Joachimi et al., 2021; Baleato Lizancos & White, 2023).

As the systematics modify the local observational depth, a selection bias occurs which changes the types of galaxies which can be observed within a given pixel. This will systematically bias the dispersion of intrinsic ellipticities observed, $\sigma_{\epsilon,m}^{(p)}$,

and further exacerbate the effect of observational depth variability on the cosmic shear signal.

To take all these effects into consideration within the forward simulations, we calibrate the spatial variability of $n^{(p)}_{\mathrm{gal},m}$ and $\sigma^{(p)}_{\varepsilon,m}$ from the KiDS-1000 measurements directly. Rather than defining maps over the entire survey footprint of both quantities for each tomographic bin, we use a map of a direct estimator of observational depth which also correlates with $n^{(p)}_{\mathrm{gal},m}$ and $\sigma^{(p)}_{\varepsilon,m}$: the root-mean square of the background noise, σ_{rms}. We found this quantity to be a good indicator of variable depth as it correlates well with the measured $n^{(p)}_{\mathrm{gal},m}$ and $\sigma^{(p)}_{\varepsilon,m}$, while being uncorrelated with the measured photometric redshifts and magnitudes in KiDS-1000.

A similar approach to modelling variable depth has been taken in Joachimi et al. (2021). However, in this analysis, the variable depth along the KiDS-1000 footprint was modelled with a direct estimate of the magnitude limit in the *r*-band (the band in which galaxy shape measurements are made in KiDS). We have found that the magnitude limit associated with a given galaxy in KiDS DR4 is correlated with the galaxy's observed r-band magnitude as well as the photometric redshift estimate for the galaxy. This means that a selection according to the magnitude limit in the *r*-band as was done in Joachimi et al. (2021) can bias cosmological estimates, as the selection is not independent from the cosmic shear signal. In contrast, σ_{rms} is uncorrelated with the *r*-band magnitude measurements and the photometric redshifts in KiDS DR4, while still being highly correlated with the local magnitude limit in the *r*-band. Thus, we conclude it is a more direct parametrisation of the variations across the survey footprint the galaxy selection.

As can be seen in Figured 3.8 and 3.9, a map of σ_{rms} across the KiDS North and South fields clearly shows variations over different pointings as well as with areas of increased overlap. In addition, Figure 3.10 shows that $n^{(p)}_{\mathrm{gal}}$ and $\sigma^{(p)}_{\varepsilon}$ vary linearly with σ_{rms} for KiDS-1000 data. Thanks to this fact, we do not need to model the variation of $n^{(p)}_{\mathrm{gal},m}$ and $\sigma^{(p)}_{\varepsilon,m}$ with the HEALPIX pixel m with individual maps for each tomographic bin p. Instead, using the linear relations shown in Figure 3.10, we may assign a value for $n^{(p)}_{\mathrm{gal},m}$ and $\sigma^{(p)}_{\varepsilon,m}$ for each $\sigma_{\mathrm{rms},m}$ (i.e. the mean value of

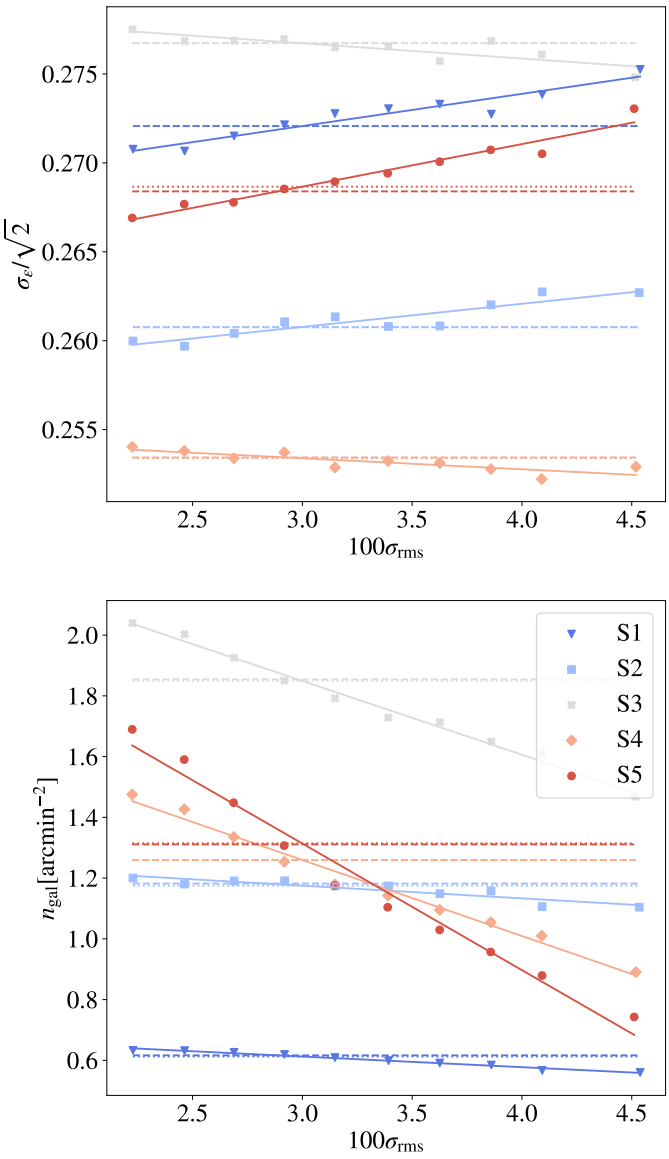

Figure 3.10: Plots showing the dependence of the per-component Gaussian shape dispersion, $\sigma_\varepsilon/\sqrt{2}$, (top panel) and the galaxy density, $n_{\rm gal}$, (bottom panel) on the root-mean square of the background noise, $\sigma_{\rm rms}$ in the KiDS-1000 DR4 data. For both panels, the data points represent the mean σ_ε or $n_{\rm gal}$ of ten equi-populated bins in $\sigma_{\rm rms}$ with their boundaries in $\{1.70, 2.33, 2.57, 2.80, 3.04, 3.28, 3.52, 3.76, 4.00, 4.23, 12.96\}$. The solid line shows the linear fit to the aforementioned data points of their respective tomographic bin according to Equations (3.30) and (3.31). The parameters obtained from this fit are given in Table 3.1. The dotted horizontal lines show the mean values of σ_ε and $n_{\rm gal}$ calculated from the galaxy samples with variable depth per tomographic bin, while the dashed horizontal lines show the values of σ_ε and $n_{\rm gal}$ for the respective galaxy samples without any spatial variations in the observational depth. Both of these lines agree exceptionally well by construction, so that for some source bins there is not any observable difference between them.

level within each pixel m contained within the footprint observed by KiDS-1000) given that

$$n_{\mathrm{gal},m}^{(p)} = a_{n_{\mathrm{gal}}}^{(p)} \sigma_{\mathrm{rms},m} + b_{n_{\mathrm{gal}}}^{(p)}, \tag{3.30}$$

$$\sigma_{\varepsilon,m}^{(p)} = a_{\sigma_\varepsilon}^{(p)} \sigma_{\mathrm{rms},m} + b_{\sigma_\varepsilon}^{(p)}, \tag{3.31}$$

where $a_{n_{\mathrm{gal}}}^{(p)}$ and $a_{\sigma_\varepsilon}^{(p)}$ are the slopes of the linear fits shown in Figure 3.10 of galaxy density and galaxy shape dispersion for a given tomographic bin, respectively, and $b_{n_{\mathrm{gal}}}^{(p)}$ and $b_{\sigma_\varepsilon}^{(p)}$ are the associated y-intercepts. The values of these parameters are shown in Table 3.1. This drastically helps the performance of the forward simulations, as this allows to just use a single map of σ_{rms} to model the variable depth rather than requiring $2N_{\mathrm{tomo}}$ maps for it. Such a linear relation may not exist in other weak lensing surveys, but it should still be possible to numerically calibrate the relations dependence of $n_{\mathrm{gal}}^{(p)}$ and $\sigma_\varepsilon^{(p)}$ on σ_{rms} to define an interpolation.

These inhomogeneities in the galaxy selection across the survey footprint also change the redshift distribution of galaxies. As σ_{rms} decreases, the local magnitude limit increases and fainter galaxies can be detected by the survey. As more distant galaxies with high redshifts also tend to be fainter, variable depth can shift the redshift distributions of the local galaxy population as can be seen in Figure 3.11. The redshift distributions shown in Figure 3.11 have been calculated using the same approach as in KiDS-1000 where the photometric redshifts are calibrated from spectroscopic samples using self-organising maps (Wright et al., 2019). The self-organising map has been applied to ten equi-populated subsamples of each tomographic bin of KiDS DR4 which are binned in σ_{rms} along the following boundaries: $\{1.70, 2.33, 2.57, 2.80, 3.04, 3.28, 3.52, 3.76, 4.00, 4.23, 12.96\}$. To account for variable depth in our forward model, depending on the value of σ_{rms} in a given pixel, m, on the survey footprint, we sample galaxy redshifts in that pixel from the associated redshift distribution from Figure 3.11. This is effect is important to model as an unaccounted shift in the redshift distribution can lead to significant biases in the cosmic shear signal (Heydenreich et al., 2020; Baleato Lizancos &

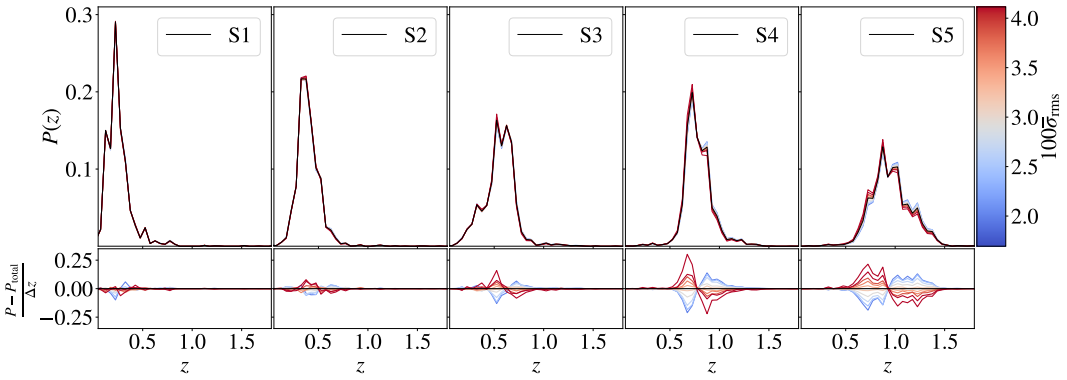

Figure 3.11: In the upper panels, plot of the normalised redshift distributions, $P(z)$, for each tomographic bin (S1 to S5). The redshift distribution from the entire KiDS-1000 DR4 galaxy sample, $P_{\rm total}(z)$, is shown in black, while the other ten redshift distributions are derived from 10 equi-populated subsamples of DR4 based on their observational depth (i.e. the mean value of the root-mean-square of the background noise, $\overline{\sigma}_{\rm rms}$) which is shown with its respective colour. The lower panels show the associated residual change in the redshift distributions with respect to $P_{\rm total}(z)$ per unit redshift. It is apparent that variable depth mostly affects the source distributions at high redshifts, while the effect tends to decrease the mean of the redshift distribution with increasing $\sigma_{\rm rms}$.

White, 2023).

To summarise, in KiDS-SBI, we account for the effects of spatial variations in the observational depth on galaxy number density, galaxy shape dispersion and redshift distributions. Moreover, we also consider anisotropies in the shear bias by including distortions to the lensed galaxy ellipticities calibrated from the spatial variation in the point-spread function ellipticity in KiDS DR4 (see Section 3.2.7).

3.2.9 Shape Measurements

Upon sampling galaxy shapes as well as shear, and applying all relevant instrumental systematics, the forward simulations create a realistic catalogue of galaxies for each tomographic bin containing the associated $\varepsilon_{\rm obs}$ for each galaxy. To fully simulate the KiDS-1000 catalogue, we apply the same corrections to the shear measurements as described in Giblin et al. 2021. The observed shape measurements are corrected as follows

$$\varepsilon_{{\rm obs},i}^{{\rm corr}\,(p)}(\mathbf{\Theta}) = \frac{1}{1+\overline{M}^{(p)}} \frac{\sum_{i\in p} w_i (\varepsilon_{{\rm obs},i}(\mathbf{\Theta}) - \langle \varepsilon_{\rm obs}(\mathbf{\Theta})\rangle)}{\sum_{i\in p} w_i}, \tag{3.32}$$

where w_i is the galaxy LENSFIT weight for galaxy i (for the case of the simulations, $w_i = 1 \forall i$), $\langle \varepsilon_{\rm obs}(\mathbf{\Theta}) \rangle$ compensates for any additive biases, i.e. the $c^{(p)}$ term and the mean of $\alpha^{(p)}$ in Equation (3.29), and it is defined as the mean observed shape of all galaxies within tomographic bin p, and $\overline{M}^{(p)}$ is the mean multiplicative shear bias measured for each tomographic bin p as shown in Figure 3.6. Note that $\overline{M}^{(p)}$ is not necessarily the same value as the value of $M^{(p)}$ applied in Equation (3.29). For each instance of the simulations, a different $M^{(p)}$ is drawn from a Gaussian probability distribution with mean $\overline{M}^{(p)}$ and the standard deviations shown in Figure 3.6. Regardless of the value of $M^{(p)}$ drawn, the shear measurements of each simulation are corrected by the same mean $\overline{M}^{(p)}$. Any discrepancy between these values is going to introduce noise in the simulations which accounts for the uncertainty on the shear bias measurements.

3.2.10 Pseudo-Cls

With a full catalogue simulated, we can map any set of cosmological and astrophysical parameters, $\mathbf{\Theta}$, to a corresponding set of cosmic shear measurements from the KiDS survey as required for simulation-based inference. Although theoretically possible, it is computationally impractical to conduct cosmological inference at the level of catalogues. For this reason, it is useful to compress the catalogues down to useful statistics which still retain most of the relevant cosmological information about the underlying large-scale structure.

We choose to compress the catalogue down to two-point statistics. Firstly, the forward-simulation pipeline is designed to be only percentage-level accurate in two-point statistics. Secondly, this assures that the only distinguishing feature of our analysis when compared to the fiducial cosmic shear KiDS-1000 analyses (Loureiro et al., 2021; Asgari et al., 2021) is the dropping of the assumption of a Gaussian likelihood, so we can carry out a direct comparison between the analyses.

This way we can ensure that any non-Gaussianities in our likelihood are either attributed to inherent non-Gaussianities in the likelihood of two-point statistics or non-Gaussianities induced by systematic effects.

Due to its computational efficiency when compared to spatial two-point cor-

relation functions, we choose pseudo-Cls as our two-point statistic. We decompose the observed shear field, $\varepsilon_{\rm obs}^{{\rm corr}\,(p)}(\theta_m;\mathbf{\Theta})$, into its curl-free E-modes and its divergence-free B-modes as follows

$$\varepsilon_{\rm obs}^{{\rm corr}\,(p)}(\theta;\mathbf{\Theta},\Omega_{\rm survey}) = \sum_{\ell=0}^{\ell_{\rm max}}\sum_{m=-\ell}^{\ell}(\tilde{E}_{\ell m}^{(p)}(\mathbf{\Theta}) + {\rm i}\tilde{B}_{\ell m}^{(p)}(\mathbf{\Theta}))_2Y_{\ell m}(\theta), \tag{3.33}$$

where ${}_2Y_{\ell m}(\theta)$ are the spin-2 spherical harmonic functions which define an orthonormal basis for the observed shear field, such that

$$\int {\rm d}^2\theta\, {}_{\pm2}Y_{\ell m}(\boldsymbol{\theta})\, {}_{\pm2}Y^*_{\ell' m'}(\boldsymbol{\theta}) = \delta_{\ell\ell'}\delta_{mm'}, \tag{3.34}$$

while $\tilde{E}_{\ell m}^{(p)}(\mathbf{\Theta})$ and $\tilde{B}_{\ell m}^{(p)}(\mathbf{\Theta})$ are spherical harmonic coefficients of the curl-free and the divergence-free observed shear fields within each tomographic bin, respectively. These coefficients are defined as follows:

$$\tilde{E}_{\ell m}^{(p)}(\mathbf{\Theta}) \;=\; \frac{1}{2}\int {\rm d}^2\theta[\varepsilon_{\rm obs}^{{\rm corr}\,(p)}(\boldsymbol{\theta};\mathbf{\Theta})_2Y^*_{\ell m}(\boldsymbol{\theta}) \;+\; \varepsilon_{\rm obs}^{{\rm corr}\,(p)*}(\boldsymbol{\theta};\mathbf{\Theta})_{-2}Y^*_{\ell m}(\boldsymbol{\theta})], \tag{3.35}$$

$$\tilde{B}_{\ell m}^{(p)}(\mathbf{\Theta}) \;=\; \frac{-i}{2}\int {\rm d}^2\boldsymbol{\theta}[\varepsilon_{\rm obs}^{{\rm corr}\,(p)}(\boldsymbol{\theta};\mathbf{\Theta})_2Y^*_{\ell m}(\boldsymbol{\theta}) \;-\; \varepsilon_{\rm obs}^{{\rm corr}\,(p)*}(\theta;\mathbf{\Theta})_{-2}Y^*_{\ell m}(\boldsymbol{\theta})]. \tag{3.36}$$

From these coefficients, we then calculate the pseudo-Cls, $\tilde{C}_{\varepsilon\varepsilon}^{(pq)}(\ell;\mathbf{\Theta})$, as follows

$$\tilde{C}_{\varepsilon\varepsilon,\mu}^{(pq)}(\ell;\mathbf{\Theta}) = \begin{pmatrix}\tilde{C}_{\varepsilon\varepsilon}^{EE(pq)}(\ell;\mathbf{\Theta})\\ \tilde{C}_{\varepsilon\varepsilon}^{EB(pq)}(\ell;\mathbf{\Theta})\\ \tilde{C}_{\varepsilon\varepsilon}^{BB(pq)}(\ell;\mathbf{\Theta})\end{pmatrix} \tag{3.37}$$

$$= \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell}\begin{pmatrix}\tilde{E}_{\ell m}^{(p)}\tilde{E}_{\ell m}^{(q)*}(\mathbf{\Theta})\\ \tilde{E}_{\ell m}^{(p)}\tilde{B}_{\ell m}^{(q)*}(\mathbf{\Theta})\\ \tilde{B}_{\ell m}^{(p)}\tilde{B}_{\ell m}^{(q)*}(\mathbf{\Theta})\end{pmatrix}, \tag{3.38}$$

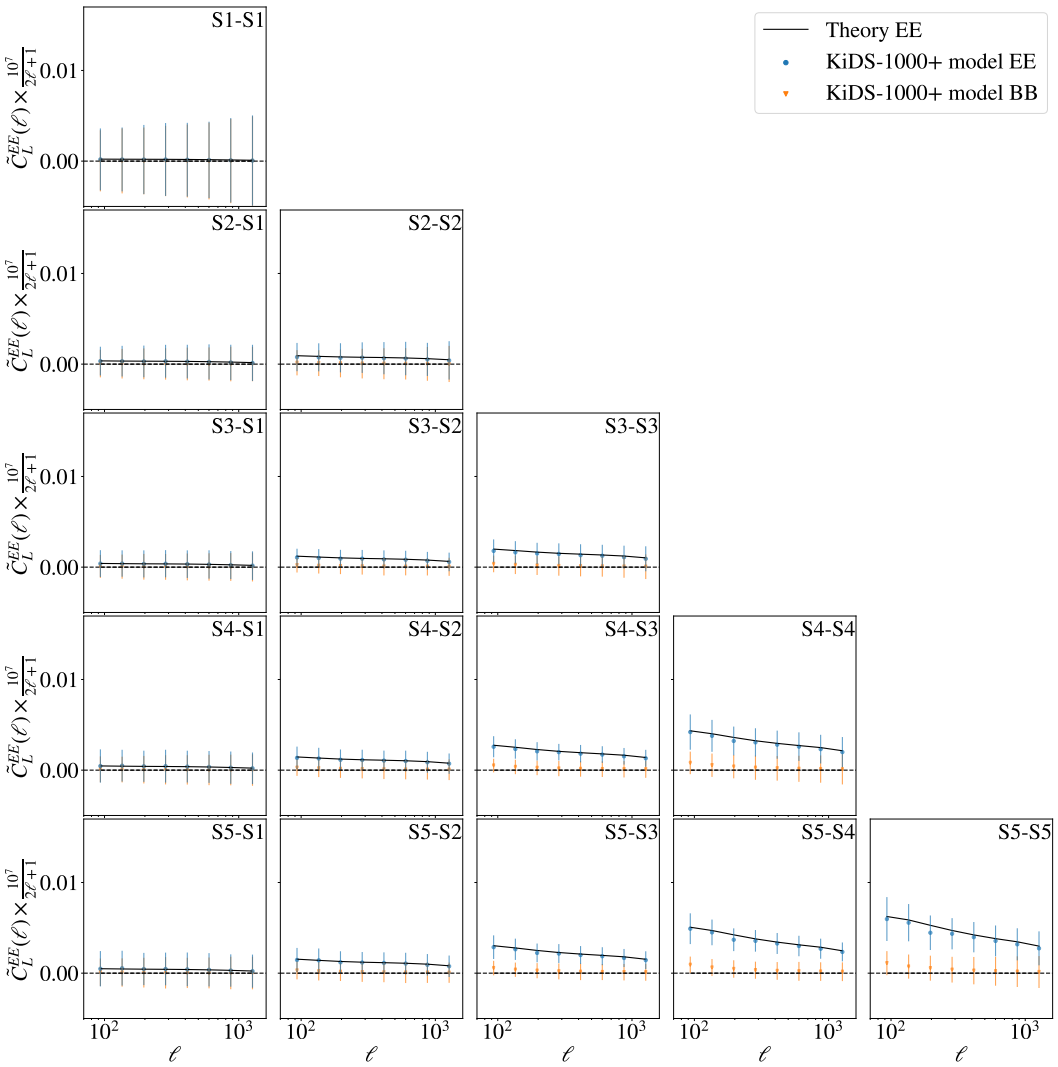

Figure 3.12: Signal of a single run of the KiDS-SBI simulations for all combinations for five tomographic bins (S1 to S5). The blue points show the pseudo-Cl measurements, $\tilde{C}^{(pq)}_{\varepsilon\varepsilon,L}(\mathbf{\Theta})$, for a realisation single realisation of the *KiDS-1000+* model. The orange triangles show the associated BB modes in the cosmic shear signal. Throughout we assume $\Omega_c = 0.05$, $\Omega_b = 0.28$, $\sigma_8 = 0.84$, $S_8 = 0.76$, $H_0 = 67$ km s^{-1} Mpc^{-1}, $A_{\rm bary} = 3.1$ and $A_{\rm IA} = 0.56$. The uncertainties on the measurements are derived from the covariance matrix described in Section 3.3.2. The solid black line shows the pseudo-Cls, $\tilde{C}_{\varepsilon\varepsilon}$, as derived from theory (see Appendix B for details on this).

where $\mu \in \{1,2,3\}$ such that 1 stands for the EE component, 2 for the EB component and 3 for the BB component. Note that $\tilde{C}^{EB}_{\varepsilon\varepsilon}(\ell;\mathbf{\Theta}) = \tilde{C}^{BE}_{\varepsilon\varepsilon}(\ell;\mathbf{\Theta})$, since $\tilde{E}_{\ell m}\tilde{B}^*_{\ell m}(\mathbf{\Theta}) = \tilde{B}_{\ell m}\tilde{E}^*_{\ell m}(\mathbf{\Theta})$.

Going forward, similar to other previous analyses (Hikage et al., 2019; Loureiro et al., 2021), we only take into consideration $\tilde{C}^{EE(pq)}_{\varepsilon\varepsilon}(\ell;\mathbf{\Theta})$, so any cosmo-

logical signal which may have ended up in $\tilde{C}^{EB(pq)}_{\varepsilon\varepsilon}(\ell;\mathbf{\Theta})$ or $\tilde{C}^{EB(pq)}_{\varepsilon\varepsilon}(\ell;\mathbf{\Theta})$ because of EE to EB, or EE to BB mode mixing will be lost. $\tilde{C}^{EE(pq)}_{\varepsilon\varepsilon}(\ell;\mathbf{\Theta})$ is expected to still contain all of the cosmological cosmic shear signal (Loureiro et al., 2021), but the autocorrelations will be dominated by the shape noise. To account for this, we subtract the mean shape noise power spectrum as follows

$$\tilde{C}^{(pq)}_{\varepsilon\varepsilon}(\ell;\mathbf{\Theta}) = \tilde{C}^{EE(pq)}_{\varepsilon\varepsilon}(\ell;\mathbf{\Theta}) - \delta_{pq}\langle\tilde{C}^{EE(pq)}_{\text{noise}}(\ell)\rangle, \tag{3.39}$$

where $\tilde{C}^{(pq)}_{\varepsilon\varepsilon}(\ell;\mathbf{\Theta})$ is the E-mode pseudo angular power spectrum for cosmic shear with the shape noise bias subtracted, while $\langle\tilde{C}^{EE(pq)}_{\text{noise}}(\ell)\rangle$ is the mean of the curl-free angular power spectrum of the shape noise. The latter is estimated as follows (Becker et al., 2016; Hikage et al., 2019; Nicola et al., 2021; Loureiro et al., 2021),

$$\tilde{C}^{EE(pp)}_{\text{noise}}(\ell) = \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell}\tilde{E}^{\text{rand}\,(p)}_{\ell m}(\mathbf{\Theta})\tilde{E}^{\text{rand}\,(p)*}_{\ell m}(\mathbf{\Theta}), \tag{3.40}$$

where $\tilde{E}^{\text{rand}\,(p)}_{\ell m}$ is the curl-free spherical harmonic coefficient of the randomly rotated shear values, $\varepsilon^{\text{rand}\,(p)}_{\text{obs}}(\theta_m;\mathbf{\Theta})$. In turn, we define this field from the following galaxy shear values as follows:

$$\varepsilon^{\text{rand}}_{\text{obs},i}(\mathbf{\Theta}) = \varepsilon^{\text{rand}}_{\text{obs},i,1}(\mathbf{\Theta}) + \mathrm{i}\varepsilon^{\text{rand}}_{\text{obs},i,2}(\mathbf{\Theta}), \tag{3.41}$$

$$\varepsilon^{\text{rand}}_{\text{obs},i,1} = \varepsilon^{\text{corr}}_{\text{obs},i,1}\cos(\theta_{\text{rand},i}) - \varepsilon^{\text{corr}}_{\text{obs},i,2}\sin(\theta_{\text{rand},i}), \tag{3.42}$$

$$\varepsilon^{\text{rand}}_{\text{obs},i,2} = \varepsilon^{\text{corr}}_{\text{obs},i,2}\cos(\theta_{\text{rand},i}) + \varepsilon^{\text{corr}}_{\text{obs},i,1}\sin(\theta_{\text{rand},i}), \tag{3.43}$$

where $\theta_{\text{rand},i}$ is a randomly drawn angle for each galaxy i from a uniform distribution where $\theta_{\text{rand},i} \in [0, 2\pi)$. To compute the mean shape noise bias, we have to take into consideration that each angular power, ℓ. To calculate the mean of the curl-free angular power spectrum of the shape noise, $\langle\tilde{C}^{EE(pq)}_{\text{noise}}(\ell)\rangle$, we take mean of all modes in $\tilde{C}^{EE(pq)}_{\text{noise}}(\ell)$ as follows

$$\langle\tilde{C}^{EE(pq)}_{\text{noise}}(\ell)\rangle = \frac{\sum_{\ell=\ell_{\text{min}}}^{\ell_{\text{max}}}(2\ell+1)\,\tilde{C}^{EE(pq)}_{\text{noise}}(\ell)}{\sum_{\ell=\ell_{\text{min}}}^{\ell_{\text{max}}}(2\ell+1)}, \tag{3.44}$$

where the mean is weighted by a factor of $2\ell+1$ to account for the fact that each angular scale, ℓ, contains $2\ell+1$ modes, m. We find setting $\ell_{\rm min} \sim 200$ yields optimal results, as high angular scales are typically shape noise-dominated, and thus put better constraints on the average shape noise bias. In fact, we find that the measured average shape noise bias, $\langle \tilde{C}^{EE(pq)}_{\rm noise}(\ell)\rangle$, agrees within $< 0.5\%$ with the underlying shape noise signal computed directly from the known intrinsic galaxy shapes. Additionally, we find that $\langle \tilde{C}^{EE(pq)}_{\rm noise}(\ell)\rangle$ is self-consistent with $\langle \tilde{C}^{BB(pq)}_{\rm noise}(\ell)\rangle$ within $< 0.01\%$. The introduction of a weighted mean given by Equation (3.44) allows us to achieve this level of precision with only a single random rotation of the galaxies' shapes. Instead, previous approaches relied on creating many different instances of random rotations and then computed $\langle \tilde{C}^{EE(pq)}_{\rm noise}(\ell)\rangle$ by taking the average over all the instances while weighting each angular scale, ℓ, equally (see e.g. Loureiro et al. 2021). When weighting each ℓ equally, the low-ℓ scales have an equally strong pull on the estimator as the high-ℓ scales which increases the random uncertainty on $\langle \tilde{C}^{EE(pq)}_{\rm noise}(\ell)\rangle$ as the large scales are mostly dominated by the large-scale structure signal. For this reason, it is common to take the mean over many realisation of the shape noise to reduce the variance of the mean. Although such approaches achieve similar levels of precision when estimating the shape noise bias as the approach shown in this work, computing the angular power spectra for hundreds of realisations of randomly rotated galaxy shape catalogues can become computationally expensive and time-consuming. Hence, it is not feasible for an SBI analysis which requires $\sim 10^4$ realisations and we opt to use the weighted mean of a single realisation of a random rotation to estimate the shape noise bias.

Furthermore, we bin the noise-free observed angular power spectrum, $\tilde{C}^{(pq)}_{\varepsilon\varepsilon}(\ell;\mathbf{\Theta})$, into 8 log-spaced bins between $\ell = 76$ and $\ell = 1500$ in line with Loureiro et al. 2021. For this, we choose the pseudo-Cl binning scheme described in Brown et al. 2005 given by

$$\tilde{C}^{(pq)}_{\varepsilon\varepsilon,L}(\mathbf{\Theta}) = \frac{1}{2\pi}\sum_{\ell}^{\ell_{\rm max}} \frac{\ell(\ell+1)}{(\ell_{L+1}-\ell_L)} \tilde{C}^{(pq)}_{\varepsilon\varepsilon}(\ell;\mathbf{\Theta}), \tag{3.45}$$

where ℓ_L and ℓ_{L+1} are the lower and upper limits of the $L^{\rm th}$ bin, respectively.

Note that we do not choose to deconvolve the pseudo-Cls to estimate the full-sky E-mode angular power spectra for cosmic shear. We find that with the sky-coverage of KiDS-1000 and the complexity of the geometry of the KiDS-1000 footprint, the mixing matrix (see Appendix B) is not necessarily invertable, and the deconvolution is not single-valued.

We therefore obtain for a given forward simulation a cosmic shear measurement like the one shown in Figure 3.12 for a single run of the forward simulations. The calculation of the data vector calculated from the theory prediction shown in Figure 3.12 is described in Appendix B. As can be seen from Figure 3.12, the measured cosmic shear signal is consistent with theory predictions which we find is the case at all cosmologies throughout the prior volume. This shows that the log-normal random fields accurately recover the two-point statistics of the galaxy populations throughout parameter space, as expected.

As seen in Figure 3.12, for the five tomographic bins in KiDS-1000, we obtain $8 \times N_{\rm tomo}(N_{\rm tomo}+1)/2$ data points in $\tilde{C}^{(pq)}_{\varepsilon\varepsilon,L}(\mathbf{\Theta})$, i.e. a 120-dimensional data vector. Each data vector depends on 12 parameters in $\mathbf{\Theta}$: 5 cosmological parameters (σ_8, the root-mean-square matter fluctuation over 8 Mpc$/h_0$; ω_b, baryonic matter density; ω_c, cold dark matter density; n_s, the scalar spectral index of the primordial density fluctuation power spectrum, $P_{\mathscr{R}}$; and h_0, the normalised Hubble constant), and 7 astrophysical/nuisance parameters related to systematics: $A_{\rm bary}$, the baryonic feedback amplitude within the non-linear three-dimensional matter power spectrum; $A_{\rm IA}$, the galaxy intrinsic alignment amplitude of the NLA model; and five correlated δ_z parameters which define the shift in the mean of the source redshift distribution, $P(z|z_{\rm ph})$, of each tomographic bin. All other parameters on which the simulation depends are fixed from run to run (e.g. $\omega_k = 0$, the equation of state of dark energy, $w = -1$, $\sum m_\nu = 0.06$ eV, etc.). The exceptions to this are the seed used to sample the matter overdensities and the galaxies, and the amplitudes of the shear biases as described in Section 3.2.7. These values are varied from run to run in order to simulate cosmic variance, shape noise as well as the uncertainty in the shear bias, respectively.

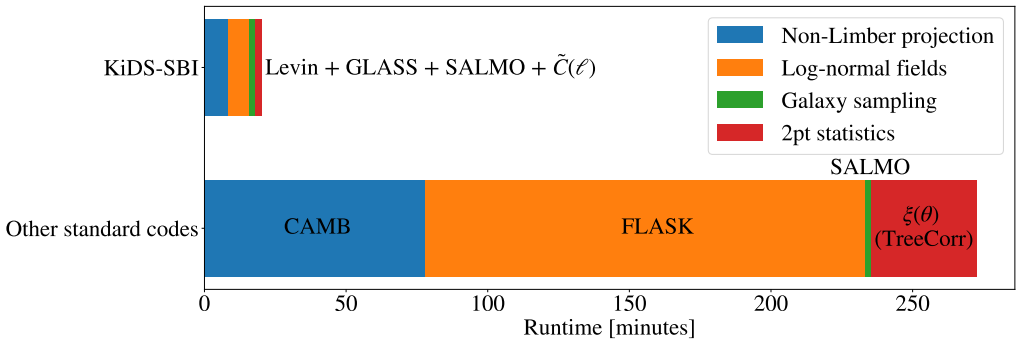

Figure 3.13: Bar chart comparing the run-time of a single evaluation of KiDS-SBI (above) versus a single evaluation of a simulation based on Joachimi et al. 2021 (below), both on a single core ($N_{\rm shells} = 19$, $N_{\rm tomo} = 5$ and $N_{\rm side} = 1024$). Both suites of simulations use CAMB (Lewis et al., 2000; Lewis & Challinor, 2002; Howlett et al., 2012) to compute the three-dimensional matter power spectrum. For the reference simulations, we use the non-Limber projection built into CAMB with *limber_phi_lmin* $= 1200$ rather than LEVIN with $\ell_{\rm max,nL} = 1200$. We run FLASK (Xavier et al., 2016) rather than GLASS (Tessore et al., 2023) to compute the underlying matter and convergence fields of each of the 19 shells. Subsequently, we sample galaxies using SALMO in both cases, and then calculate the spatial two-point correlation functions, $\xi_\pm(\theta)$, rather than calculating the angular power spectra, $\tilde{C}(\ell)$. To calculate $\xi_\pm(\theta)$ in the reference simulations, we use TREECORR (Jarvis et al., 2004).

A single such evaluation of the forward simulations runs within ∼20 minutes on a single core (with $N_{\rm shells} = 19$, $N_{\rm tomo} = 5$ and $N_{\rm side} = 1024$). We choose this spatial resolution $N_{\rm side}$, because this implies that the two-point statistics should be accurate up to an $\ell_{\rm max} \sim 2 \times N_{\rm side}$ (Leistedt et al., 2013; Alonso et al., 2019). In addition, in the context of KiDS-1000's galaxy number density, the pixel size at an $N_{\rm side} = 1024$ is sufficient for almost all pixels to contain at least one galaxy. This avoids mode mixing in the pseudo Cls due to random masking of empty pixels which were observed (see Appendix B).

In comparison, as can be seen in the bar chart in Figure 3.13, log-normal random field simulations similar to the ones presented in Joachimi et al. 2021, take ∼280 minutes to compute a single forward simulation when run with the same accuracy and precision settings on a single core. These gains in speed are driven by three main factors: the fast non-Limber integration using the Levin method, the use of recurrence relations to calculate the convergence within GLASS and the choice

of using a quick-to-compute summary statistic like angular power spectra, $\tilde{C}(\ell)$. This is a substantial improvement with respect to other typical codes used for such simulations, while also improving accuracy with respect to previous models by including systematics, improving the resolution along the line-of-sight and dropping common approximations like the Limber and reduced shear approximation. Thanks to this, it becomes feasible to increase the precision of our simulations as necessary, while also adding the realism discussed in Sections 3.2.6, 3.2.7 and 3.2.8, while still being able to compute the $\sim 10^4$ forward simulations needed to adequately characterise the effective likelihood L22.

3.3 Simulation-Based Inference (SBI)

The final data vector as defined by the forward simulations is used to determine the effective likelihood of the data using Density Estimation Likelihood-Free Inference (DELFI[13], Alsing et al. 2019) using the same analysis pipeline outlined in L22. To implement this, we create a bespoke sampler within CosmoSIS[14]. Initially, the sampler chooses a fixed number of points within the considered hyperparameter space based on the hypercube-generating algorithm from the PyDOE[15] with some minor modifications (see L22 for more details). Subsequently, as is outlined in Figure 3.14, the measured pseudo-Cls for each evaluation of the parameters on the hypercube are compressed further using score compression (see Sections 1.4.2 and 3.3.2). This reduces the dimensionality of the data vector to the size of the parameter vector (in this case, 12 dimensions). The compressed data is then used to train neural density estimators through DELFI (see Sections 1.4.3.2 and 3.3.3). As the ensembles of neural density estimators evaluate the effective likelihood throughout the hypercube, through the use of active learning, DELFI finds areas in parameter space near likelihood peaks to select new parameter vectors (see Alsing et al. 2019 for details). As found in L22, this can halve the number of simulations needed to accurately learn the effective likelihood when compared to a Latin hypercube

[13] `https://github.com/justinalsing/pydelfi`
[14] `https://github.com/joezuntz/cosmosis`
[15] `https://github.com/tisimst/pyDOE`

alone. Once the learned effective likelihood has converged, we calculate the posterior distribution based on an observed data vector (which can be a mock or measured from real data), and the priors. This has the advantage that the posteriors need not be re-sampled if a different data vector and/or prior are considered.

Going forward, we refer to the model choices that we make within KiDS-SBI as the *KiDS-1000+* model. The model follows the choices outlined in Section 3.2. It is designed to be consistent with previous analyses of KiDS-1000, specifically, with Loureiro et al. (2021). Both analyses assume a flat ΛCDM cosmology to model the cosmic shear signal, both use pseudo-Cls as their data vector of choice, and both consider systematics such as multiplicative shear bias, variable depth in the uncertainty and intrinsic alignments in the signal. The *KiDS-1000+* model differs slightly from Loureiro et al. (2021) in that it also includes variable depth in the signal modelling (as described in Section 3.2.8), it considers the effect of intrinsic alignments on the likelihood, and it takes into account the variance in the additive and PSF shear biases (as described in Section 3.2.7). This is a consequence of the fact that in KiDS-SBI the effects which are modelled in the signal are intrinsically considered in the uncertainty model. In Loureiro et al. (2021), these effects are only considered separately as they pertain to the signal or the uncertainty, not both at once. Having said that, these discrepancies between the two models are not expected to have a significant effect on the Gaussian likelihood.

3.3.1 Parameters and Priors

The parameters which are varied in the simulation-based inference in this work are shown in Table 3.2. The priors shown in Table 3.2 for each parameter are selected in accordance with the assumed priors for the previous KiDS-1000 cosmic shear analyses (Asgari et al., 2021; Heymans et al., 2021; Loureiro et al., 2021; van den Busch et al., 2022).

To avoid overinformative priors, most are flat top-hat functions. For the same reason, the top-hat priors in S_8, h_0, ω_b, A_{IA} and A_{bary} are chosen to be wide. The prior on n_s spans a smaller range around the theoretical value of unity for scale-invariant primordial fluctuations. This avoids artefacts within the prior volume as

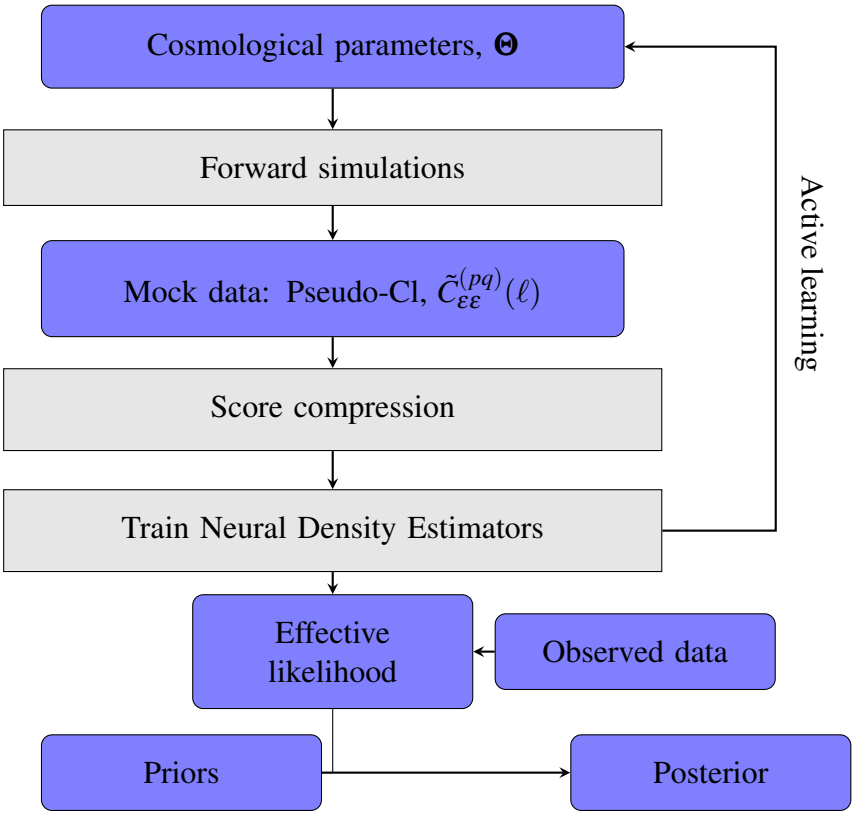

Figure 3.14: Flowchart describing the structure of the simulation-based inference pipeline. The dark blue rounded boxes represent the inputs and outputs which are given to the simulation-based inference pipeline. The grey rectangular boxes show steps in the inference pipeline.

this parameter is not well constrained by weak gravitational lensing. The prior on ω_c is defined to be consistent with a range in $\Omega_m \in [0.188, 0.408]$, where the limits are given by the $\pm 5\sigma$ intervals of the marginal constraints from independent measurements of luminosity distance to Type Ia Supernovae in Scolnic et al. (2018). See Joachimi et al. (2021) for more details on the motivations for the chosen priors. The priors on the nuisance parameters capturing any shifts in the mean of each of the five tomographic bins, $\boldsymbol{\delta}_z$, are given by a multivariate Gaussian, since the δ_z of a given tomographic bin is not independent of the shifts in the other bins. This is quantified by the covariance, $\boldsymbol{C}_z$, as estimated in Hildebrandt et al. (2021).

3.3.2 Score Compression

To improve the computational efficiency of the analysis and to facilitate the use of DELFI, we compress each of the 120-dimensional data vectors measured from

Parameter	Symbol	Prior type	Prior parameters
Density fluctuation amp.	S_8	Flat	[0.1, 1.3]
Hubble constant	h_0	Flat	[0.64, 0.82]
Cold dark matter density	ω_c	Flat	[0.051, 0.255]
Baryonic matter density	ω_b	Flat	[0.019, 0.026]
Scalar spectral index	n_s	Flat	[0.84, 1.1]
Intrinsic alignment amp.	A_{IA}	Flat	[-6, 6]
Baryon feedback amp.	A_{bary}	Flat	[2, 3.13]
Redshift displacement	$\boldsymbol{\delta}_z$	Gaussian	$\mathcal{N}(\mathbf{0}, \boldsymbol{C}_z)$

Table 3.2: Table showing the parameters which are varied within the simulation-based inference pipeline. The prior ranges are selected to be exactly in line with previous KiDS-1000 analyses (Asgari et al., 2021; Heymans et al., 2021; Loureiro et al., 2021; van den Busch et al., 2022). The upper five rows show the cosmological parameters of interest, while the lower three rows show the nuisance parameters which quantify systematic biases. For flat priors, the lower and upper limits of the normalised rectangular function which defines the prior. For the Gaussian prior on $\boldsymbol{\delta}_z$, we use a five-dimensional multivariate Gaussian with its mean at the zero vector and the covariance, $\boldsymbol{C}_z$, defined by the one estimated in (Hildebrandt et al., 2021). Note that for simplicity the $\boldsymbol{\delta}_z$ are implicitly marginalised throughout this analysis.

each forward simulation using nuisance-hardened score compression as described in Alsing et al. (2019). As discussed in Section 1.4.2, if the likelihood is known a priori, one can compress a given data vector to a summary of the same dimensionality as the degrees of freedom in the assumed model, such that Fisher information is conserved (Zablocki & Dodelson, 2016; Alsing & Wandelt, 2018; Alsing et al., 2018; Alsing et al., 2019). If the likelihood is Gaussian, this is equivalent to MOPED (Heavens et al., 2000b) or a linear compression based on Karhunen-Loéve eigenvalue decomposition (Tegmark et al., 1997).

In this work the exact form of the likelihood is not known a priori, as the main motivation is to characterise the form of the effective likelihood. However, it is still possible to perform score compression on the data by assuming an analytical form for the likelihood when defining the Fisher matrix. To this end, we create 1,000 forward simulations at a fiducial cosmology to characterise a numerical covariance which we use to define the Fisher matrix using Equation (1.112). This compression is optimal if the true likelihood and the chosen fiducial set of parameters equal to the true parameters. The downside of this is that the compression can lose information

if the latter is not the case. Nevertheless, it was found in L22 that, in KiDS-SBI, the score compression is robust to suboptimal choices of fiducial parameter values and of data covariance.

Additionally, we find that the 12 parameters shown in Table 3.2 are not necessary to characterise the degrees of freedom of the 120-dimensional cosmic shear pseudo-Cls. As all δ_z are broadly consistent with zero, we may compress all data vectors to a seven-dimensional summary which is still capable of capturing all the complexity in the data. In any case, the $\boldsymbol{\delta}_z$ parameters are still explicitly varied within the simulation to propagate any uncertainties on the mean of the tomographic bins. For simplicity, henceforth all posteriors are implicitly marginalised over the five δ_z parameters.

In Figure 3.15, we show the Fisher forecasts used for the score compression of the data in this analysis. The contours shown in Figure 3.15 are obtained from sampling a Gaussian likelihood characterised by the inverse of the Fisher matrix defined in Equation (1.112) as its covariance and the fiducial parameters as the mean. Besides this, the Fisher forecasts are also useful to estimate the constraining power that a Gaussian likelihood based on the *KiDS-1000+* model has. We note that the constraining power achieved by the *KiDS-1000+* model appears to be improved when including variable depth in the modelling. This is driven by variable depth reducing the estimated shape noise in KiDS-1000, which improves the constraining power on S_8 and $A_{\rm IA}$ (see Chapter 4 for a detailed discussion of this).

3.3.3 Density Estimation Likelihood-Free Inference

To estimate the effective likelihood for the *KiDS-1000+* model, we train neural density estimators using DELFI (Alsing et al., 2019) such that they learn the joint probability distribution between the compressed data and the input cosmological parameters. To this end, we initialise an ensemble of six independent conditional Masked Autoregressive Flows (MAFs). Each MAF is made up of between three and eight Masked Autoencoder for Density Estimations (MADEs) each with two hidden layers of 50 neurons. For a detailed definition and description of these network architectures, see Section 1.4.3.

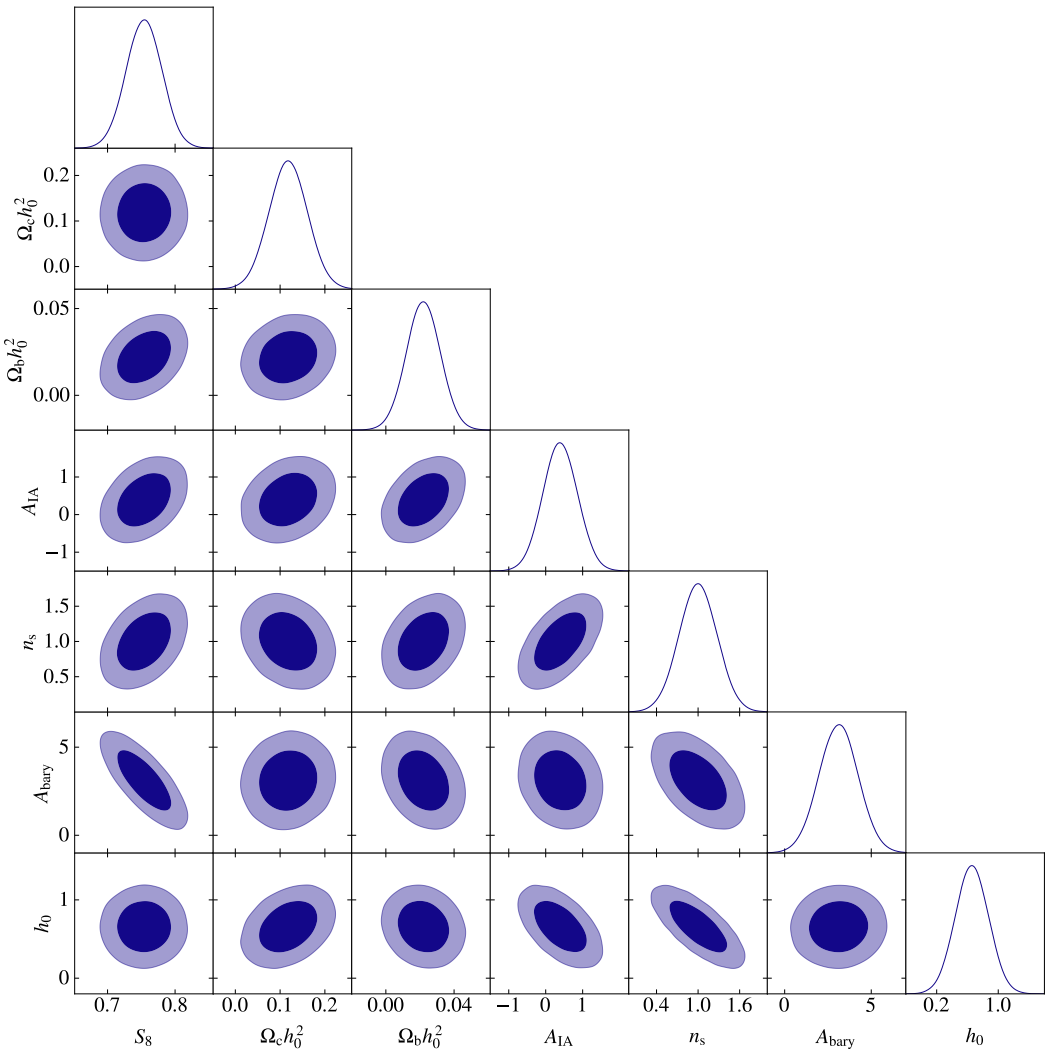

Figure 3.15: Contours showing the Fisher forecasts obtained from 1,000 realisations of the *KiDS-1000+* (in blue) model. The input fiducial cosmology is set to be $S_8 = 0.761$, $\omega_c = 0.118$, $\omega_b = 0.022$, $h_0 = 0.657$, $n_s = 1.0$, $A_{IA} = 0.396$, $A_{bary} = 3.113$.

Initially, the ensemble of NDEs learns the effective likelihood from a set 400 points on a Latin hypercube. Based on this, we then make use of the active learning feature within DELFI to sample additional points in parameter space which most efficiently contribute to learning the effective likelihood within prior space. This recursive sampling continues until the training loss as defined in Equation (1.128) plateaus at a minimum. We find that for the *KiDS-1000+* model, this occurs already after 5,000 realisations. We note that the NDEs require fewer realistic *KiDS-SBI* forward simulations to minimise the loss function as is the case in L22 where the forward simulations where idealised random samples from a Gaussian distribution. In any case, to ensure that the prior space is densely sampled, we choose to train the NDEs on 14,000 realisations. This is facilitated by the computational efficiency of the forward simulations making additional realisations relatively inexpensive.

Upon training each of the six MADs on a fraction of the forward simulation independently, the NDEs are stacked while weighting each according to their relative cross-validation losses. From the stacked NDE, we obtain the final posterior distributions.

3.4 Validation of the SBI Pipeline

The SBI pipeline presented in this work has previously been extensively tested in an idealised case. From the findings in L22, we conclude that the SBI pipeline based on DELFI implemented within KiDS-SBI robustly and accurately recovers the posterior distribution of ΛCDM cosmological parameters from KiDS-1000 cosmic shear data, even if the data compression loses information. However, L22 only considered simulated data whose likelihood is by construction Gaussian. For this reason, it becomes necessary to test the robustness of the SBI pipeline again for the *KiDS-1000+* forward model.

Figure 3.16 shows the posterior distribution of the seven main parameters varied in the analysis as learned from 14,000 realisations of the *KiDS-1000+* model. The data vector used for this posterior is a noisy mock data vector generated with the *KiDS-1000+* model while assuming the input parameters shown in Figure 3.12.

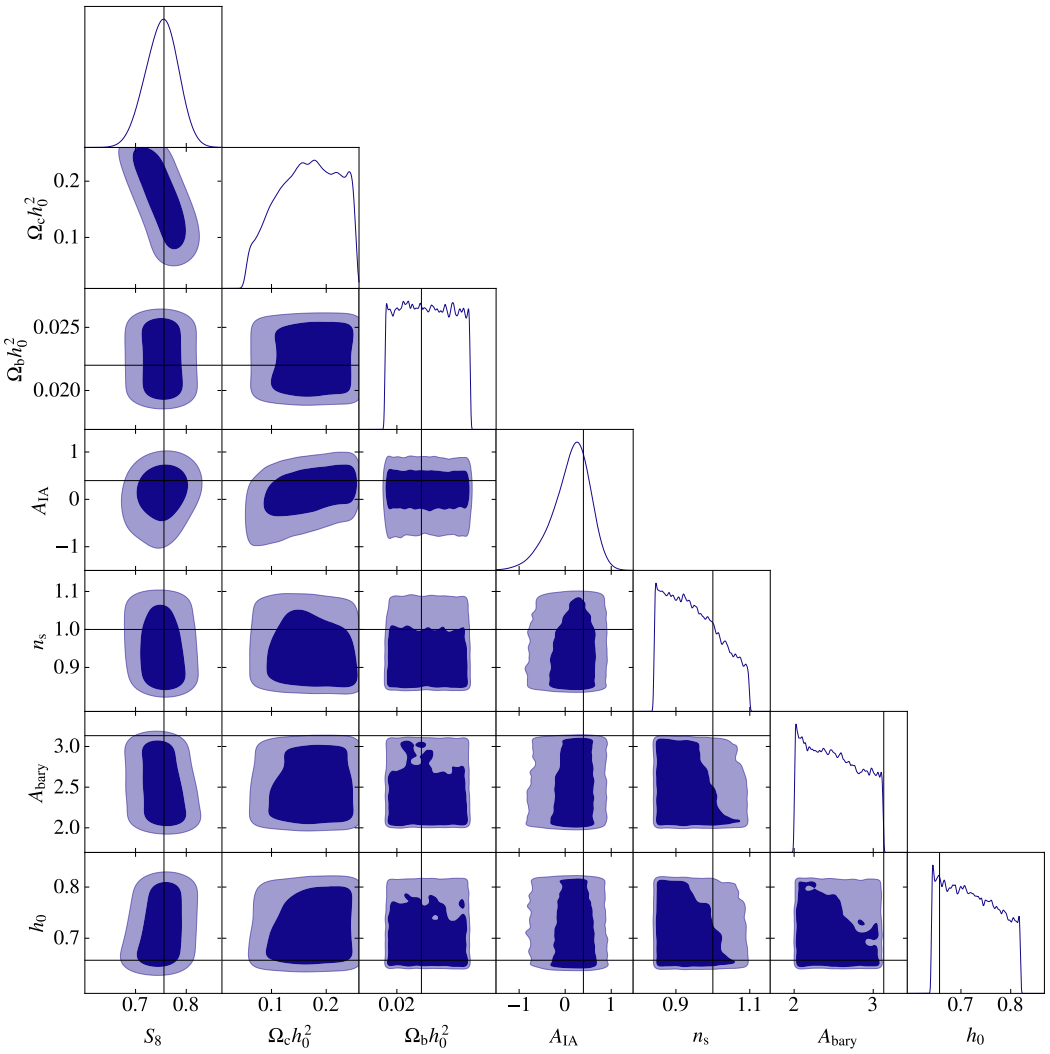

Figure 3.16: Posterior contours, in blue, of the seven cosmological and astrophysical parameters which are varied given the *KiDS-1000+* model within KiDS-SBI over the prior space shown in Table 3.2. The black solid lines indicate the true cosmology of the input mock data vector generated from the *KiDS-1000+* model while adding noise. All the aforementioned values are shown in Table 3.3. These posteriors are obtained from training neural density estimators in DELFI (Alsing et al., 2019) on 14,000 realisations of the forward simulations assuming the *KiDS-1000+* model, in line with the choices made in L22. The posterior is obtained from the combined posteriors of six independent conditional Masked Autoregressive Flows (MADs) each is made up of three to eight Masked Autoencoder for Density Estimations (MADEs) each with two hidden layers of 50 neurons.

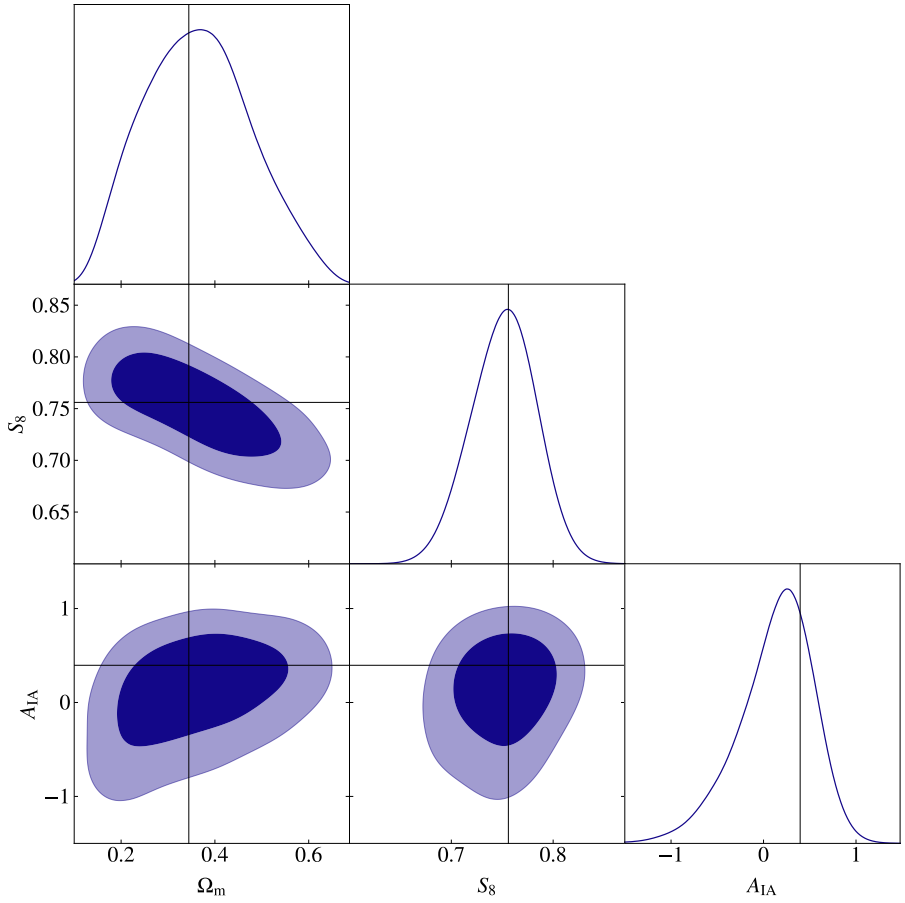

Figure 3.17: Posterior marginals, in blue, of the three cosmological parameters of interest which varied given the *KiDS-1000+* model within KiDS-SBI over the prior space shown in Table 3.2. The black solid lines indicate the true cosmology of the input mock data vector generated from the *KiDS-1000+* model while adding noise. All the aforementioned values are shown in Table 3.3. Note that the matter density parameter, $\Omega_{\rm m}$, is a derived parameter given by $\Omega_{\rm m} \equiv (\omega_{\rm c} + \omega_{\rm b})/h_0^2$, and the posterior shown here is derived from the posteriors shown in Figure 3.16.

We find that the posteriors from *KiDS-SBI* accurately recover the true input parameters (see Table 3.3 to see the inferred parameter values). As expected and shown in Figure 3.17, the *KiDS-1000+* posterior mostly only constrains S_8 and $A_{\rm IA}$. At the same time, it becomes apparent that the posterior distributions over $\omega_{\rm b}$, $n_{\rm s}$, $A_{\rm bary}$ and h_0 are mainly prior driven, i.e. the posterior distribution is mostly flat throughout prior space. As the cosmic shear signal is largely degenerate to small changes in these parameters, this is expected. However, in analyses which assume a Gaussian likelihood, such parameters which are unconstrained by the data, can appear con-

Parameter	Mock truth	Marginal $\pm 1\sigma$	MAP $\pm$ PJ-HPD
S_8	0.756	$0.752^{+0.034}_{-0.029}$	$0.753^{+0.022}_{-0.049}$
σ_8	0.706	$0.720^{+0.093}_{-0.19}$	$0.618^{+0.156}_{-0.138}$
Ω_m	0.344	$0.36^{+0.10}_{-0.14}$	$0.446^{+0.212}_{-0.150}$
h_0	0.657	< 0.747	$0.640^{+0.125}_{-0.001}$
ω_c	0.292	$0.162^{+0.075}_{-0.048}$	$0.160^{+0.095}_{-0.027}$
ω_b	0.022	—	$0.023^{+0.003}_{-0.003}$
n_s	1.0	< 0.981	$0.984^{+0.014}_{-0.131}$
A_IA	0.396	$0.11^{+0.48}_{-0.31}$	$0.257^{+0.360}_{-0.448}$
A_bary	3.133	—	$2.016^{+0.911}_{-0.016}$

Table 3.3: Table of the main inferred cosmological and astrophysical parameters varied within the *KiDS-1000+* model. For each parameter, we show the underlying true value which was input into the mock data vector, and the inferred value recovered by KiDS-SBI. The second column shows the marginal peaks as well as the upper and lower 68% confidence intervals, i.e. 1σ, of the marginals. h_0, n_s, ω_b, and A_bary are not well enough constrained in order to calculate a meaningful marginal parameter estimate. The third columns shows the Maximum A Posteriori (MAP), and the uncertainties are defined as the upper and lower 68% confidence intervals, i.e. 1σ, given by the projected joint highest posterior density, PJ-HPD (Robert et al., 2007; Joachimi et al., 2021). Note that cosmic shear is only expected to recover precisely the value of S_8 and A_IA.

strained in the posterior just because of the constraints placed by the assumption of a Gaussian likelihood. In contrast, the additional degrees of freedom in the effective likelihood learned by the NDEs can help avoid overconfident constraints on parameters.

To determine whether the learned posterior distribution from the *KiDS-1000+* model is representative of the true underlying posterior and is unbiased, we conduct the Tests of Accuracy with Random Points (TARP) shown in Figure 3.18. TARP (Lemos et al., 2023a) measures the expected coverage probability of a random sample of posterior samples within a given credibility level of the learned posterior. If the expected coverage probability is directly proportional to the credibility level, it would indicate that the learned posterior is unbiased and representative of the true underlying posterior. Any biases in the learned posterior would lead to deviations from linearity in the TARP. As can be seen from Figure 3.18, the learned posterior must be unbiased and accurate (Lin et al., in prep.). At low credibility levels, the posterior samples are narrower than the ideal expectation, so the learned posterior

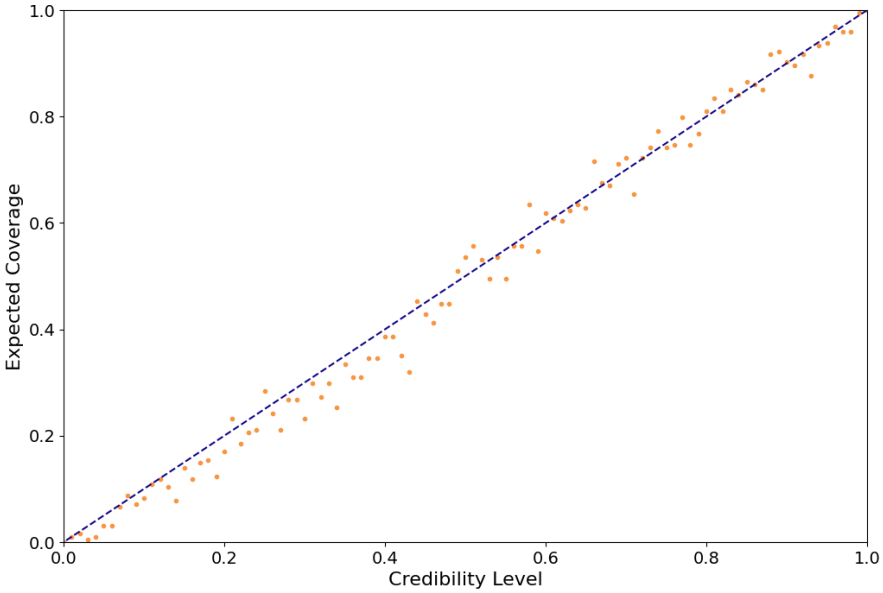

Figure 3.18: Plot of the expected coverage probability versus the credibility level as defined in the Tests of Accuracy with Random Points (TARP) described in Lemos et al. (2023a) for the posterior shown in Figure 3.16 assuming the *KiDS-1000+* model. The dashed line is a reference line for a perfectly linear relation. The credibility level gives the fraction of the total probability density of the learned posterior being considered, while the expected coverage probability measures the fraction of posterior samples which have a posterior probability smaller than the best estimate at a given credibility level. We note the relation in this case is highly linear which is a necessary and sufficient measure that the posterior estimate given in Figure 3.16 is accurate. Figure from Lin et al. (in prep.).

is slightly underconfident.

In summary, we find that the SBI inference pipeline within *KiDS-SBI* accurately and robustly recovers the underlying posterior distribution of ΛCDM parameters from 2PCF measurements of the cosmic shear signal in KiDS-1000 when considering a forward model based log-normal random fields and biased by relevant systematic effects.

3.5 Cosmological Inference from Mock KiDS-1000 Data

To put the constraining power of the results from KiDS-SBI into context, we compare the posterior contours shown in Figures 3.16 and 3.17 with the constraints from

other analyses: an analysis of the KiDS-1000 cosmic shear signal from pseudo-Cls assuming a Gaussian likelihood (Loureiro et al., 2021), and the cosmological constraints from the TT+TE+EE modes of temperature fluctuations in the cosmic microwave background (Planck Collaboration et al., 2020).

Figure 3.19 compares the marginalised posterior contours from these analysis over S_8, Ω_m and A_IA, where the latter is not inferred from CMB data, as it is a nuisance parameter specific weak lensing and galaxy survey probes. In Figure 3.19, the SBI posterior contours are consistent with the Gaussian likelihood contours from Loureiro et al. (2021) by construction, as the input data vector used for the SBI *KiDS-1000+* contours are set to near the MAP cosmology inferred from the KiDS-1000 data (Asgari et al., 2021). Despite this being the case, it is apparent that the constraints in S_8 and Ω_m from the SBI posteriors are noticeably more conservative than the constraints from the Gaussian likelihood analysis from Loureiro et al. (2021), while the opposite is the case for A_IA. In particular, the SBI posterior prefers higher values of S_8 and smaller values of Ω_m than the Gaussian likelihood analysis. Nevertheless, the *KiDS-1000+* model is still in good agreement with Loureiro et al. (2021). The marginal in S_8 from SBI *KiDS-1000+* is $S_8 = 0.752^{+0.034}_{-0.029}$ (i.e. a relative uncertainty of 8.4%), while in Loureiro et al. (2021) it was found to be $S_8 = 0.742^{+0.034}_{-0.023}$. At the same time, the SBI *KiDS-1000+* yields a MAP$\pm$PJ-HPD of $S_8 = 0.753^{+0.022}_{-0.049}$ (i.e. a relative uncertainty of 7.6%), while the MAP$\pm$PJ-HPD from Loureiro et al. (2021) is $S_8 = 0.754^{+0.027}_{-0.029}$. The latter implies that the dropping the assumption of a Gaussian likelihood increases the relative uncertainty on S_8 from 7.4% to 9.4%, while also changing the skew of the posterior distribution such that higher values of S_8 and Ω_m are preferred over low values of S_8 and Ω_m. We note that the shape of the posterior and the resultant constraints on S_8 can noticeably change depending on the mock realisation used as the measured data vector. Consequently, the uncertainties quoted here may not be entirely representative of the constraints obtained from the real KiDS-1000 cosmic shear data. Having said this, if this result persists in the real KiDS-1000 data, this would provide further insights into the S_8 or σ_8 "tension" observed between early-Universe

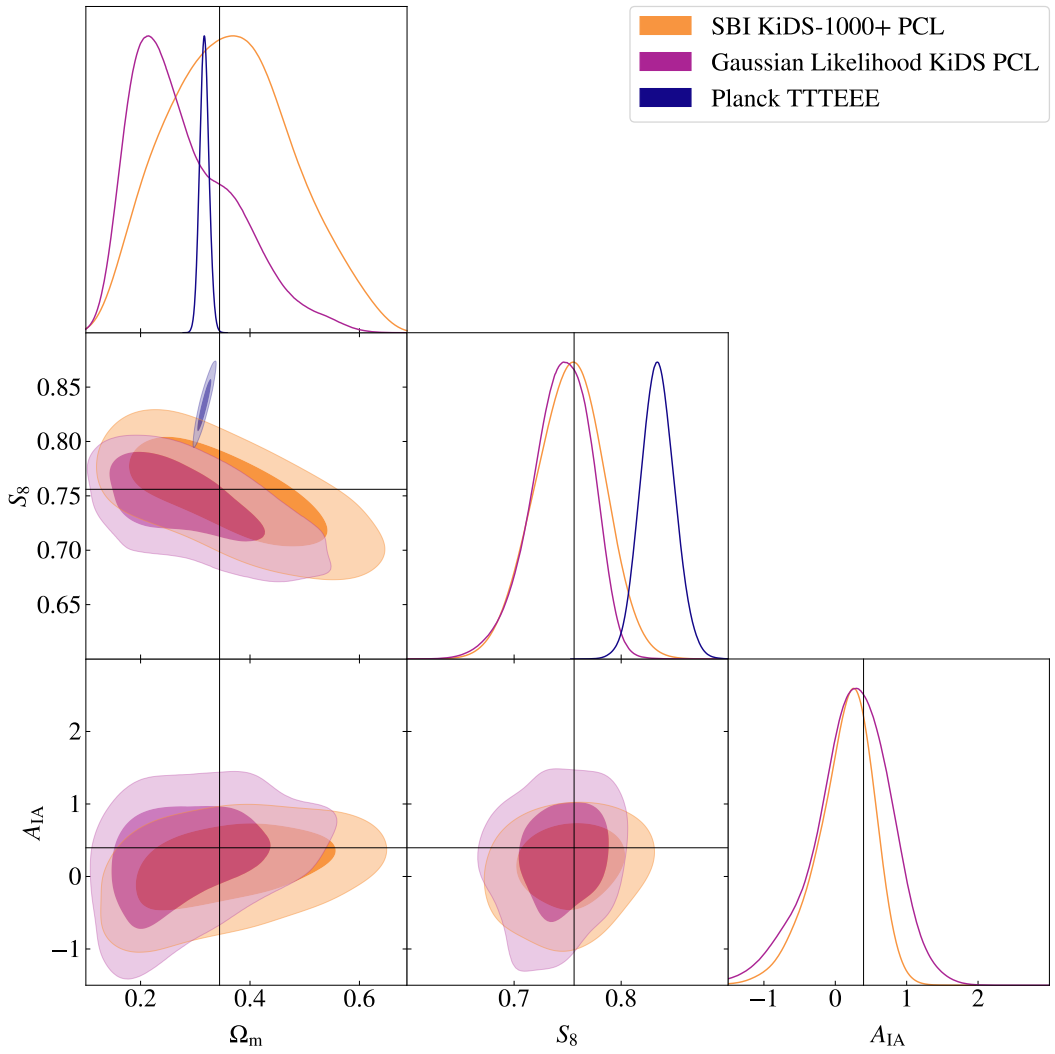

Figure 3.19: Posterior contours of the main constrained cosmological parameters from the KiDS-SBI analysis of a mock cosmic shear data vector assuming the *KiDS-1000+* model (in orange) compared against posterior contours from other analysis. The purple contour shows the posterior from an analysis of the cosmic shear signal measured with pseudo-Cls from KiDS-1000 data assuming a Gaussian likelihood (Loureiro et al. 2021; where the uncertainty model includes variable depth as is the case in the *KiDS-1000+* model), while the blue contour shows the posterior from the cosmic microwave background constraints from the TT+TE+EE modes (Planck Collaboration et al., 2020). The solid black lines show the true cosmology assumed in the mock data vectors used for the SBI contours (see Table 3.3). The true cosmology is based on the MAP from Asgari et al. (2021) which is close to the MAP from Loureiro et al. (2021), but noticeably different from the preferred cosmology from Planck Collaboration et al. (2020). Note that the Planck TT+TE+EE contours do not have any marginals in $A_{\rm IA}$ as the CMB is not sensitive to the IAs of galaxies.

and late-Universe large-scale structure probes (Hildebrandt et al., 2017; Planck Collaboration et al., 2020; Asgari et al., 2021; Amon et al., 2022; Li et al., 2023; Qu et al., 2023). To quantify this, we use the marginal tension estimate, τ, given by

$$\tau = \frac{|\hat{\theta}_{\rm A} - \hat{\theta}_{\rm B}|}{\sqrt{\sigma^2_{\rm A} + \sigma^2_{\rm B}}}, \tag{3.46}$$

where $\hat{\theta}_{\rm A}$ and $\hat{\theta}_{\rm B}$ are the best estimates from probe A and B, respectively, while $\sigma^2_{\rm A}$ and $\sigma^2_{\rm B}$ are the variance of probe A and B. Applying this metric to quantify the tension between the best estimate in S_8 marginal from SBI *KiDS-1000+* and from Planck Collaboration et al. (2020) TT+TE+EE+lowE ($S_8 = 0.832 \pm 0.013$), we find that the "tension" is 2.2σ as opposed to the 2.8σ tension between the S_8 estimates from Loureiro et al. (2021) and Planck Collaboration et al. (2020).

As the modelling assumptions in Loureiro et al. (2021) and in the *KiDS-1000+* model are constructed to be consistent, with the one exception being that the former assumes a Gaussian likelihood and the latter does not, the discrepancy in their constraints may point at a non-negligible non-Gaussianity in the likelihood of 2PCF of the cosmic shear signal. Such a result would be in tension with previous studies testing the Gaussianity of the likelihood of cosmic shear 2PCF (Schneider & Hartlap, 2009; Sellentin & Heavens, 2018; Sellentin et al., 2018; Taylor et al., 2019; Upham et al., 2021), so we investigate in detail the origin of the reduced constraining power observed in this work.

Firstly, we explore the possibility of low angular modes entering our summary statistic. Previous studies have found that the modes at angular frequencies ($\ell < 50$) in the 2PCF can have significant non-Gaussianities in their likelihood (Hamimeche & Lewis, 2008; Schneider & Hartlap, 2009; Sellentin et al., 2018; Lin et al., 2020). As we do not deconvolve the pseudo-Cls used as summary statistics for the cosmic shear signal in this analysis, such low-ℓ modes may enter the measured $\tilde{C}^{EE}(\ell)$ due to mode mixing (see Appendix B). To examine this in more detail, Figure 3.20 shows an example of the contributions of low-ℓ modes to the observed pseudo-Cls in the autocorrelation of the 5th tomographic bin in KiDS-1000, which has

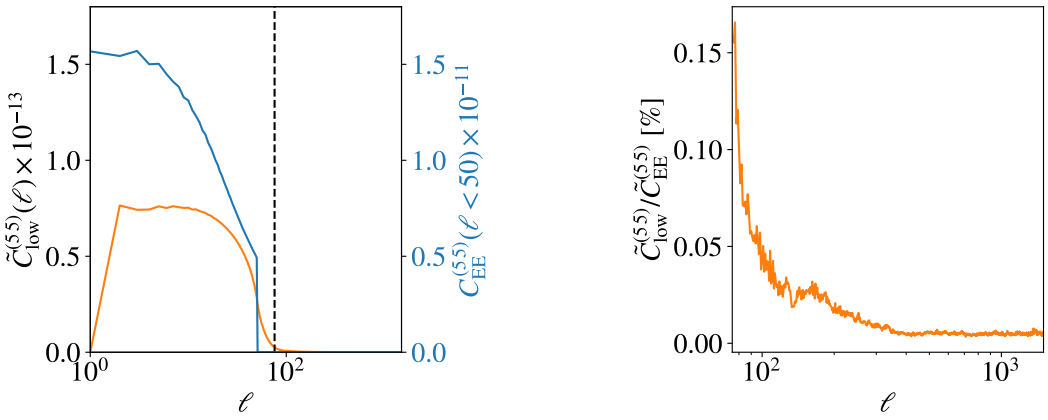

Figure 3.20: Plots showing the effect of mode mixing due to the KiDS-1000 survey footprint on the cosmic shear signal in the 5-5 tomographic bin combination. The left panel shows the expected angular power spectra from theory in blue, $C^{(55)}_{\rm EE}(\ell)$, truncated such that modes with $\ell \geq 50$ are set to zero, while also showing the corresponding pseudo-Cl, $\tilde{C}^{(55)}_{\rm low}(\ell)$, in orange which is derived from the mixing matrix shown in Figure B.1. The black dashed line marks lower limit in the domain of the pseudo-Cls considered in this analysis, i.e. $\ell = 76$. The right panel shows the ratio between the aforementioned pseudo-Cl derived from truncated theoretical angular power spectra over the measured pseudo-Cl for $\ell \in [76, 1500]$ from a single evaluation of the *KiDS-1000+* model with KiDS-SBI assuming the same cosmology.

the highest signal-to-noise ratio. We see that the cosmic shear signal from angular frequencies below 50 can contribute between 0.02% and 0.16% to the observed pseudo-Cl signal at $\ell > 76$ due to the mode mixing caused by the KiDS-1000 survey footprint. Consequently, we determine that this effect is too small to explain the observed increase of 13% in the relative uncertainty of S_8.

Another known effect which is implicitly ingrained in the error propagation of the SBI pipeline is the cosmology dependence of the uncertainty in the cosmic shear signal. It is known that cosmic variance varies with the assumed cosmological parameters (Eifler et al., 2009). Any such parameter dependence in the uncertainty of the cosmic shear signal is ignored in a Gaussian likelihood analysis, as it assumes that the covariance is fixed for all cosmologies. In KiDS-SBI, these parameter dependencies are implicitly carried through to the learned likelihood where they manifest changes in the uncertainty as a function of the cosmology across likelihood space.

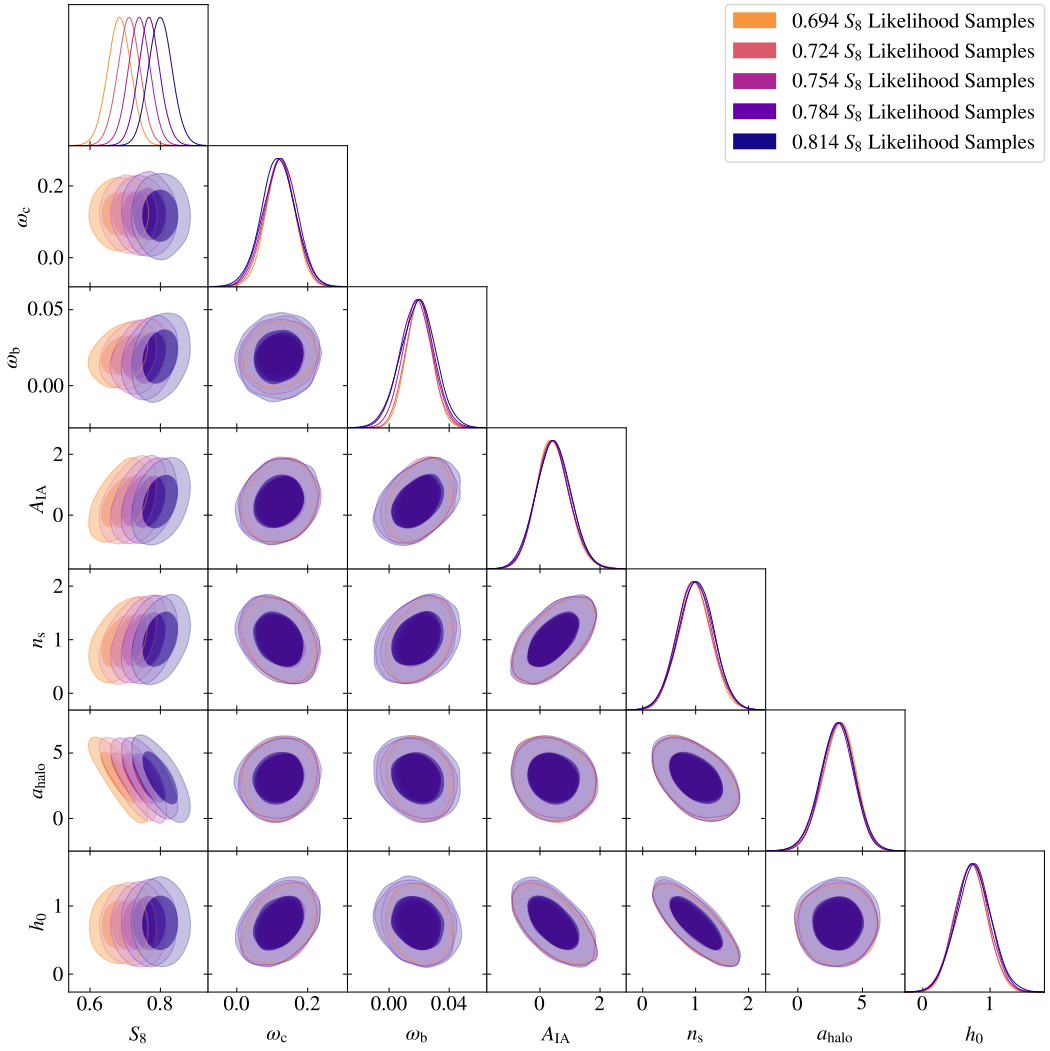

Figure 3.21: Likelihood marginals in the compressed data space for 5 different assumed sets of cosmological parameters given the *KiDS-1000+* model within KiDS-SBI over the prior space shown in Table 3.2. The compressed data values are labelled according to the cosmological parameter with which they are most correlated (see Sections 1.4.2 and 3.3.2 for details). For the orange contours, the input data vector is set to $S_8 = 0.694$, $S_8 = 0.724$ for the pink contours, $S_8 = 0.754$ for the purple contours, $S_8 = 0.784$ for the purple contours, and $S_8 = 0.814$ for the blue contours. All other cosmological parameters are taken to be the same as in Table 3.3. Table 3.4 shows the S_8 values and the associated standard deviations in the likelihood marginals.

S_8	σ_{ω_c}	σ_{ω_b}	σ_{S_8}	$\sigma_{A_{IA}}$	σ_{h_0}	σ_{n_s}	$\sigma_{A_{bary}}$
0.694	0.041	0.0086	0.034	0.55	0.24	0.32	1.3
0.724	0.043	0.0091	0.033	0.56	0.24	0.33	1.3
0.754	0.044	0.01	0.032	0.58	0.25	0.33	1.3
0.784	0.047	0.011	0.032	0.56	0.26	0.34	1.3
0.814	0.048	0.012	0.034	0.58	0.26	0.34	1.3

Table 3.4: Table of the standard deviations of the marginal likelihoods shown in Figure 3.21 as the assumed valued of S_8 is varied.

To determine whether the SBI likelihood is cosmology-dependent, we draw samples from the likelihood learned by the neural density estimators while assuming different values of S_8 and keeping the same mock data vector, while fixing all other inferred parameters. Figure 3.21 shows the resulting likelihood contours. Firstly, we note that at a given cosmology, the likelihood samples are consistent with a Gaussian distribution. At the same time, we find that as we increase the assumed value of S_8 in the likelihood in increments of approximately 0.025, the width of the likelihood contours steadily increases in the marginals of the data points most correlated with ω_b, ω_c, h_0, and n_s, as evidenced in Table 3.4. In fact, with each increment in S_8, these marginals widen by between 2% and 10%. This implies that the estimated likelihood in the cosmic shear 2PCF is dependent in S_8.

To determine the origin of the observed cosmology-dependence in the likelihood, we investigate whether there is a theoretical justification for the uncertainty in the cosmic shear signal in KiDS-1000 to vary significantly with S_8. First, we assume that the likelihood of a full-sky shear angular power spectrum, $C(\ell)$, at a given cosmology is Gaussian, i.e. $\mathscr{L} \propto \exp(-\frac{1}{2}\chi^2)$, where χ is the goodness of fit given by

$$\chi^{(pq)} = \frac{\hat{C}^{(pq)}(\ell) - C^{(pq)}_{\rm fid}(\ell)}{\sigma^{(pq)}_{\rm fid}(\ell)}, \tag{3.47}$$

where $\hat{C}^{(pq)}$ is the observed data vector, $C^{(pq)}_{\rm fid}$ is the modelled data vector at the fiducial cosmology, and $\sigma^{(pq)}_{\rm fid}$ is the standard deviation of the data at the fiducial cosmology. Assuming that the underlying shear field is described by a Gaussian random field, $\sigma^{(pq)}_{\rm fid}$ is given by

$$\left(\sigma_{\rm fid}^{(pq)}\right)^2 = \frac{\left[\left(C_{\rm fid}^{(pq)}(\ell)\right]^2 + C_{\rm fid}^{(pp)}(\ell)\,C_{\rm fid}^{(qq)}(\ell)\right.}{2\ell+1} + \left[N^{(pq)}(\ell)\right]^2, \tag{3.48}$$

where the first term on the right-hand side defined the contribution due to cosmic variance, while $N^{(pq)}$ is the shape noise contribution for a given tomographic bin combination (Joachimi et al., 2011a). To assess the impact of varying S_8 on such a likelihood distribution, we introduce a toy parameter, a, which linearly scales the model's fiducial data vector similar to S_8 as follows

$$C_{{\rm fid},a}^{(pq)}(\ell) = a\,C_{\rm fid}^{(pq)}(\ell). \tag{3.49}$$

Hence, we can define $\chi_a^{(pq)}$ as the the goodness of fit after rescaling the model by combining Equations (3.47), (3.48) and (3.49) as follows

$$\chi_a^{(pq)} = \frac{1 - a \pm \chi^{(pq)}\sqrt{\frac{1+\left[R^{(pq)}(\ell)\right]^{-2}}{2\ell+1} + \left[N^{(pq)}(\ell)/C_{\rm fid}^{(pq)}(\ell)\right]^2}}{\sqrt{a^2\,\frac{1+\left[R^{(pq)}(\ell)\right]^{-2}}{2l+1} + \left[N^{(pq)}(\ell)/C_{\rm fid}^{(pq)}(\ell)\right]^2}}, \tag{3.50}$$

where $R^{(pq)}(\ell) \equiv C_{\rm fid}^{(pq)}(\ell)/\sqrt{C_{\rm fid}^{(pp)}(\ell)\,C_{\rm fid}^{(qq)}(\ell)}$. By calculating the distribution of $\exp(-\frac{1}{2}\chi_a^2)$ over χ for different values of ℓ and a, we can assess the impact of rescaling the shear angular power spectrum directly on the Gaussian likelihood across different angular scales, ℓ. To facilitate this, one can analytically compute the standard deviation of this distribution, σ_{χ_a}, as follows

$$\sigma_{\chi_a}^2 = 1 + \frac{a^2-1}{1 + (2\ell+1)\,\frac{\left[N^{(pq)}(\ell)/C_{\rm fid}^{(pq)}(\ell)\right]^2}{1+\left[R^{(pq)}(\ell)\right]^{-2}}}. \tag{3.51}$$

Applying Equation (3.51) to the five KiDS-1000 tomographic bins, while assuming a shape noise contribution consistent with the values given in Table 3.1, we obtain Figure 3.22. We find that, as expected, at large angular scales ($\ell = 500$), the likelihood is dominated by the shape noise contribution, so its width does not scale with a under any circumstance. However, at $\ell = 70$ and $\ell = 100$, in some

tomographic bin combinations where the signal-to-noise ratio is large enough (i.e. the higher redshift bins), the width of the likelihood distribution scales linearly with the scaling applied to the shear angular power spectrum, a. This implies that the uncertainty of the KiDS-1000 cosmic shear measurements is dependent on the amplitude of the measurement itself, and therefore to the underlying cosmology. In fact, as shown in the upper right panel of Figure 3.22, the broadening of the analytical likelihood with a appears to be generally consistent with the widening observed in the likelihood marginals from the *KiDS-1000+* model in KIDS-SBI shown in Figure 3.21 and Table 3.4.

To confirm that the 2PCF cosmic shear likelihood is Gaussian at a given cosmology, but varies with the underlying cosmology, we repeat the KIDS-SBI analysis with the *KiDS-1000+* model shown in Figures 3.16, 3.17, and 3.19 with a different architecture for the neural density estimators within DEFLI. Instead of using an ensemble of MAFs, we train a single Mixture Density Network (MDN; Bishop 1994) which is made up of a single multivariate Gaussian (see Section 1.4.3.2 for details). This forces the likelihood to be Gaussian, while still allowing the cosmology to vary for each likelihood evaluation. As shown in Figure 3.23, the posterior from the MDN network is virtually identical to the posterior obtained from the SBI which allows for a non-Gaussian likelihood. Hence, we conclude that the likelihood of cosmic shear 2PCF conforms to a cosmology-dependent Gaussian likelihood.

Additionally, previous studies testing the Gaussianity of the cosmic shear likelihood do not consider all the systematics which are typically modelled in a cosmic shear analysis. Many neglect to incorporate some physical systematics, such as intrinsic alignments, or completely neglect observational systematics, such as the survey mask, shear biases or variable depth. For many of these systematics, their effect on the Gaussianity of the cosmic shear likelihood is currently unknown. Hence, each individually or in combination, these systematics may be introducing additional non-Gaussianities into the likelihood.

At the same time, it is also important to note that removing the assumption of a Gaussian likelihood can be interpreted as removing constraints on the allowed

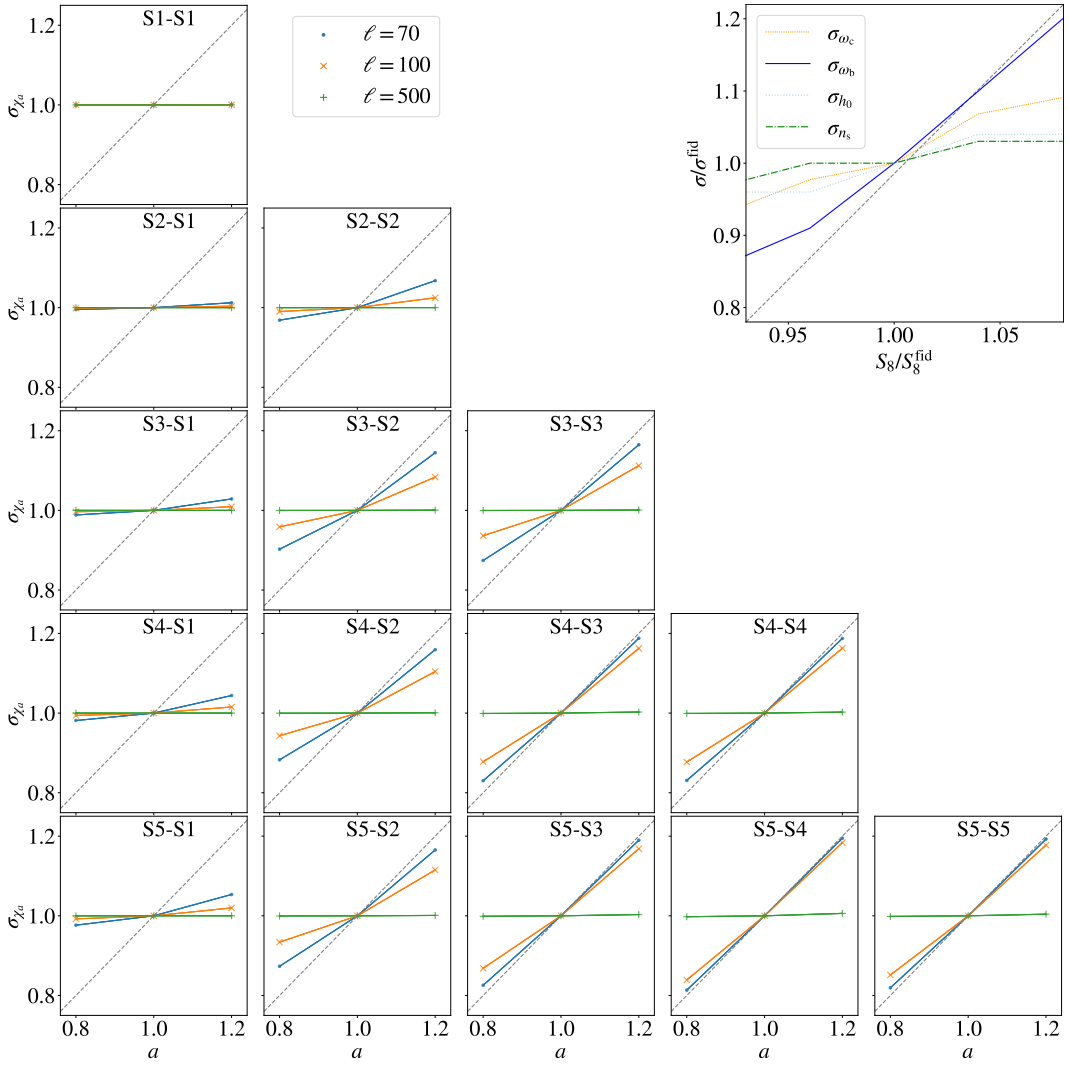

Figure 3.22: Plot of the standard deviation of the analytical likelihood distribution of χ_a (which is given by the goodness of fit of the cosmic shear 2PCF signal scaled by a factor of a) as a function of the factor a. Each of the panels shows the effect on the shear signal of a different combination of the five KiDS-1000 tomographic bins (S1 to S5). The blue dots represent the σ_{χ_a} values at $\ell = 70$ which is just below the scale cuts applied in the *KiDS-1000+* model. The orange crosses are evaluated at $\ell = 100$, while the green plus signs assume $\ell = 500$, where the uncertainty is dominated by the shape noise. The panel in the upper right corner shows how the 1σ intervals of the likelihood marginals from KiDS-SBI vary with respect to the change in S_8 relative to $S_8^{\rm fid} = 0.754$ as shown in Figure 3.21 and Table 3.4. The grey dashed line in each panel shows a direct proportionality for reference.

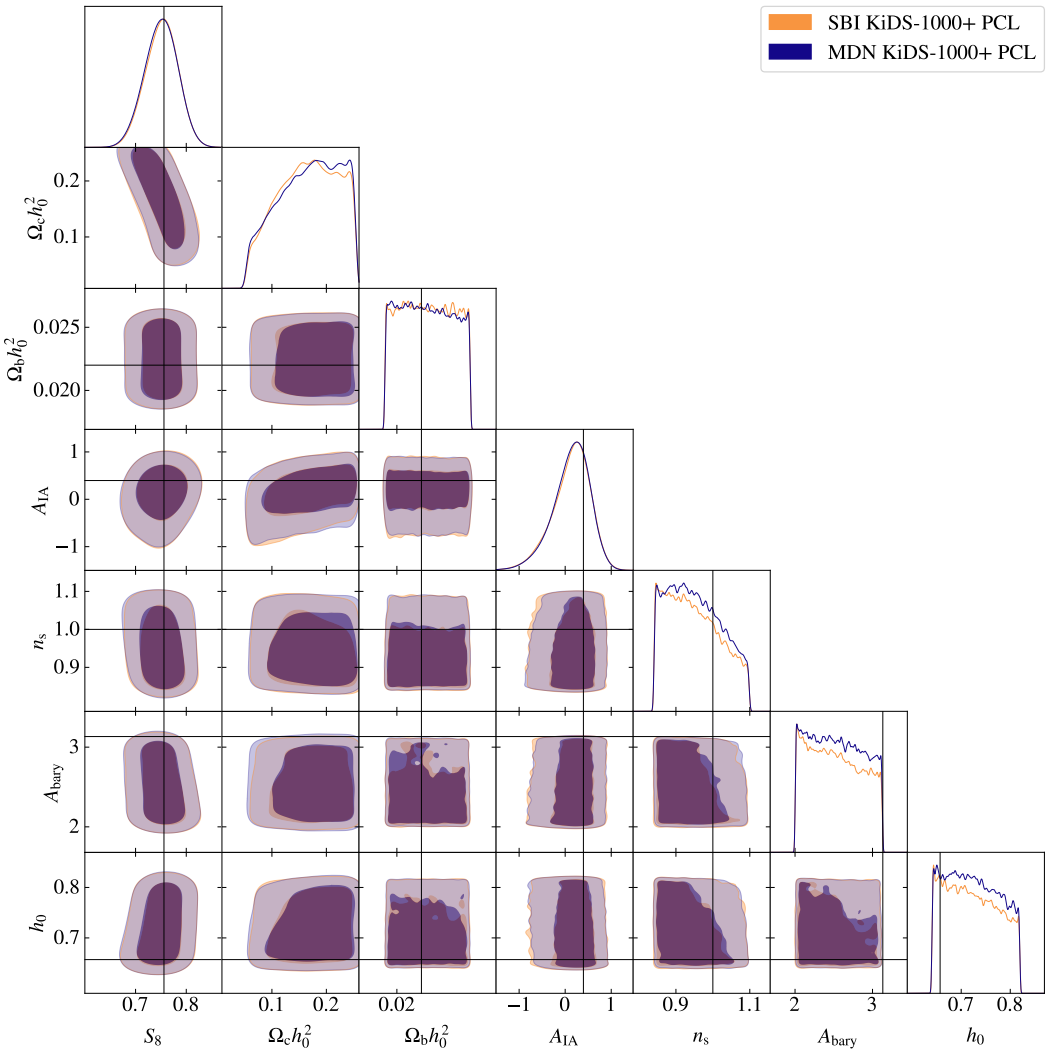

Figure 3.23: Posterior contours of the seven cosmological and astrophysical parameters which are varied given the *KiDS-1000+* model within KiDS-SBI over the prior space shown in Table 3.2. Both contours are obtained from training neural density estimators in DELFI (Alsing et al., 2019) on 14,000 realisations of the forward simulations assuming the *KiDS-1000+* model, in line with the choices made in L22. The blue posterior is calculated from a single Mixture Density Network which is made up of a single multivariate Gaussian. The orange countour is the combined posteriors of six independent conditional Masked Autoregressive Flows (MADs) each is made up of three to eight Masked Autoencoder for Density Estimations (MADEs) each with two hidden layers of 50 neurons. The black solid lines indicate the true cosmology of the input mock data vector generated from the *KiDS-1000+* model while adding noise. All the aforementioned values are shown in Table 3.3.

posterior distributions. Consequently, certain parameter combinations/volumes in prior space which were previously down-weighted, as they could not be adequately sampled by a distribution only quantified by a mean and a variance, can be characterised without any constraints. Such a picture appears to be consistent with the findings in Leclercq & Heavens (2021). In this study, it was found that, when inferring parameters of a toy model which characterises a log-normal random field from measurements of 2PCF, the constraints from SBI were found to be similarly conservative with respect to the constraints from a posterior assuming a Gaussian likelihood for the 2PCF. However, we do not find any indications that this is due to any intrinsic loss of constraining power caused by the simulation-based inference methodology (Lin et al., 2023).

3.6 Conclusions

In this chapter, we have presented a novel suite of stochastic forward simulations of cosmic shear observables, KiDS-SBI, which take into consideration all the systematic effects which are typically considered in a weak lensing analysis. At the same time, KiDS-SBI introduces considerable improvements in accuracy and performance with respect to previous similar simulation environments: non-Limber projections of angular power spectra, higher resolution along the line-of-sight, inclusion of intrinsic alignments, consideration of variable depth and anisotropic shear biases, and significant computational performance improvements.

On the basis of these forward simulations, we implement the first-of-its-kind simulation-based inference analysis of large-scale structure observables with the same level of complexity as other standard analyses in the field. Thus, this enables us to fully propagate all uncertainties from the data vector all the way to the inferred cosmological parameters in a Bayesian way. We conduct inference of the cosmological parameters in flat ΛCDM from cosmic shear two-point statistics (2PCFs) assuming a forward model consistent with previous KiDS-1000 analyses: the *KiDS-1000+* model.

We determine that KiDS-SBI recovers accurate and robust posteriors of the

cosmological parameters of interest. At the same time, we find that, when assuming a Gaussian likelihood, KiDS-SBI forward simulations give constraints consistent with previous analyses which made the same assumptions.

We also conclude that the constraints on S_8 and Ω_m from SBI are noticeably broader than constraints from equivalent analyses assuming a Gaussian likelihood. This suggests that the effective likelihood of the cosmic shear 2PCF learned through neural density estimators from the forward simulations, although approximately Gaussian at a given cosmology, is appreciably cosmology-dependent over the prior space. We exclude that any non-Gaussian contributions are driven by modes of $\ell < 50$ in the two-point statistics of the full-sky shear-shear field being mixed by the KiDS-1000 survey footprint such that they are measured in the observed pseudo-Cls at $\ell > 76$. Instead, we determine that our findings are consistent with cosmic variance scaling significantly with the underlying value of S_8, as has been observed in previous work (Eifler et al., 2009). The presence of such dependencies connotes cosmology-dependence in the likelihood which stand in contrast to a standard Gaussian likelihood analysis, where the cosmology dependence of the noise is ignored as the covariance is assumed to be fixed for all cosmologies. In the SBI analysis presented in this work, all such cosmology dependencies are implicitly considered through the forward model. Additionally, we expect that the deviations from the Gaussian likelihood analysis, at least partially, stem from the inclusion of systematics, such as intrinsic alignments, the survey footprint, and variable depth in the forward model which are not always included in other uncertainty models. Further investigations are necessary to fully determine whether the constraints observed in this SBI analysis can be reproduced with other forward models.

The next step in this analysis is conducting the same cosmological inference with the SBI pipeline with the real KiDS-1000 cosmic shear data vector. This result is not included in this work as we have followed a blinding strategy which consists in avoiding the evaluation of the effective likelihood of the real KiDS-1000 data until all tests on the SBI pipeline are finalised. Even though the KiDS-1000 cosmological analyses are already public, the non-Gaussian likelihood of the cosmic shear

2PCF is still unknown.

In any case, if the best estimate of S_8 and the uncertainties on S_8 obtained from the SBI analysis shown in this work remain the same, this would give important insights into the current σ_8 "tension" between early- and late-Universe large-scale structure probes. In this scenario, the measured tension in S_8 between KiDS-1000 and *Planck* 2018 would be within 2.2σ, rather than the 2.8σ found in an equivalent Gaussian likelihood analysis (Loureiro et al., 2021) or the 3.4σ found in the most recent KiDS-1000 analysis (van den Busch et al., 2022). This may imply that, depending on the modelling choices, current weak lensing surveys, such as KiDS, DES or HSC, may be overestimating the constraining power of their cosmic shear measurements when neglecting to model the non-Gaussianities in the likelihood caused by systematics or by cosmology-dependence in the cosmic shear uncertainty.

This work stresses the importance of testing the common assumption of Gaussian likelihoods in realistic conditions, even for data vectors which in theory have a Gaussian uncertainty. At the same time, we also highlight the potential of simulation-based inference to rigorously perform Bayesian uncertainty propagation even when the models on the uncertainty of the signal or the systematics are analytically intractable. On top of this, SBI with likelihood-based neural density estimation comes with the added advantage that it does not require the expensive reevaluation of the posterior distribution if the input measurements or priors are updated, as would be the case for a standard MCMC analysis. Since the likelihood for a given forward model is learnt independently of the input data and the priors, SBI can significantly improve the efficiency of inference analysis where the forward model is not changed between measurements.

We conclude that SBI is a powerful tool which may be able to tackle many of the physical and statistical challenges which future galaxy surveys will face, as their observations become more precise, less limited by statistical noise, and more limited by the accurate modelling of systematics. To address this need, future work should expand KiDS-SBI to also include other probes of large-scale structure such as galaxy clustering and galaxy-galaxy lensing, while incorporating the modelling

for the relevant systematic effects, such as magnification bias, non-linear galaxy bias and field-level baryonic feedback.

Chapter 4

Consistency Testing of KiDS-Legacy Modelling

Recent weak lensing analyses, such as the Kilo-Degree Survey (Asgari et al., 2021), Dark Energy Survey (Amon et al., 2022) or Hyper-Suprime Cam survey (Li et al., 2023), have relied on the assumption a of Gaussian likelihood for the cosmic shear two-point statistics. Despite Chapter 3 indicating that the Gaussian likelihood assumption might have some limitations in the context of cosmic shear, it is currently the standard assumption for any likelihood analysis of cosmic shear or galaxy clustering data (Schneider & Hartlap, 2009; Sellentin & Heavens, 2018; Sellentin et al., 2018; Taylor et al., 2019; Upham et al., 2021). Under such a presupposition, the Gaussian likelihood is characterised by a covariance matrix (see Section 1.4.1) which can be defined analytically or numerically.

Analytical error modelling has the advantage of being relatively efficient to compute when making some idealising assumptions, while providing a clear understanding of how different effects are contributing to the noise. In contrast, numerically calculated covariance matrices are often computationally expensive to obtain and may not offer clear insights into separate contributions to the noise. However, they allow us to account for all known sources of statistical and systematic noise which can be modelled numerically through simulations that may be intractable for analytical models. The analyses of KiDS-1000 (Joachimi et al., 2021; Asgari et al., 2021; Heymans et al., 2021) and DES-Y3 (Friedrich et al., 2020; Amon et al., 2022; Abbott et al., 2022) opted for an analytical covariance validated against numerical

estimates, while HSC-Y3 (Li et al., 2023; Miyatake et al., 2023) opted for a numerical approach derived from mock data vectors (Shirasaki et al., 2019) validated against analytical estimates.

For upcoming galaxy surveys, this dilemma persists. As weak lensing measurements become more precise with increasing sky coverage and depth of galaxy surveys, such as *Euclid* (Laureijs et al., 2011), *Rubin* (LSST Science Collaboration et al., 2009) and *Roman* (Spergel et al., 2015), the random uncertainty in the cosmic shear measurements decreases. Then, previously unimportant systematic uncertainties may become relevant in the error modelling. This can lead to previously sufficient analytical models becoming intractable, while increasing the complexity needed in numerical simulations. Regardless of the approach taken, it is always necessary to cross-validate the error model with an independent approach.

This work aims to test the analytical model signal and noise modelling of the cosmic shear signal as observed by the Kilo-Degree Survey against numerical simulations. Such an analysis is necessary in order to determine whether the assumptions made for the analytical covariance modelling in KiDS-1000 (Joachimi et al., 2021) are sufficient in the context of the final data-release from the Kilo-Degree Survey, also known as DR5 or KiDS-Legacy (see Li et al. 2022 for some preliminary discussions on this data-release). KiDS-Legacy improves upon KiDS-1000 by adding an additional 350 deg^2 of survey area (increasing the total to 1,350 deg^2), while also aiming to include source galaxies up to a photometric source redshift of $z \sim 2$ thanks to deeper i-band observations (Li et al., 2022). This is combined with improvements in the data reduction, masking, galaxy weights, shear calibration, redshift calibration, etc. These improvements alter the sensitivity of the cosmic shear measurements to systematic uncertainties, so it becomes important to reevaluate the assumptions made in the signal and uncertainty modelling of KiDS-1000. Two important such systematics which are sometimes ignored in the analytical covariance model in KiDS-1000 is the geometry of the survey footprint area (see e.g. Peebles 1973; Scharf & Lahav 1993) in the Gaussian covariance terms (see Section 4.3 for details) and any observational depth anisotropies, i.e. variable depth

(Heydenreich et al., 2020; Joachimi et al., 2021; Baleato Lizancos & White, 2023). To test these, I employ the statistical forward simulation framework of KiDS-SBI described in Chapter 3 to numerically characterise the covariance matrix with some minor modifications.

This chapter is structured as follows. In Section 4.1, I discuss the expected properties of the KiDS-Legacy data and the creation of the necessary mock data. Section 4.2 explains the scope of the forward simulations and highlights any changes made to the pipeline with respect to the simulations presented in Chapter 3. Section 4.3 outlines the analytical and numerical uncertainty models for a KiDS-Legacy-like analysis. Section 4.4 shows the results of the testing the signal modelling, while Section 4.5 proposes a novel measure of changes in the uncertainty model, and then proceeds to show the results of testing the uncertainty models. I give some concluding remarks in Section 4.6.

4.1 KiDS-Legacy-like Data and Calibration

The testing of the cosmic shear signal and noise modelling for the final KiDS data release, DR5 also known as KiDS-Legacy, requires many realisations of the data in order to characterise the covariance. To achieve this, the forward simulations need to be calibrated in a representative fashion. As KiDS DR5 is currently not available, it is also necessary to create a realistic mock DR5 catalogue for the purposes of calibration. Additionally, it is important to use a mock data catalogue to calibrate the simulations, as to not undermine the blinding strategy for future cosmological analyses of KiDS-Legacy cosmic shear measurements.

To address this, rather than using the KiDS DR4 catalogue (Kuijken et al., 2019), the forward simulations' inputs are instead calibrated from mock catalogues generated from organised randoms (Johnston et al., 2021). Organised randoms employ Self-Organising Maps (SOMs; Kohonen 1990) in order to learn mappings between the anisotropies in the number of observed galaxies in a given part of the sky to measures of different systematic effects in KiDS (e.g. atmospheric seeing, stellar density, dust extinction, variations in the point-spread function, etc.). For a

more detailed explanation, see Johnston et al. (2021). Once this mapping has been made, this allows to account for systematics in the data more efficiently. However, another more notable use for the application shown in this work is that the organised randoms allow to create random realisations of a galaxy catalogue which have the same underlying systematics as the real data on which it is trained. However, unlike the real data catalogue, this allows to choose the underlying cosmological model which avoids unblinding.

As the main goals of this analysis are to quantify the effects of the footprint and the variable depth on the cosmic shear signal and noise as observed by KiDS-Legacy, it is important to calibrate the relevant inputs to the simulations accordingly. Firstly, one of the main expected improvements in KiDS-Legacy with respect to KiDS-1000 is the increase in depth along the line-of-sight from a maximum source photometric redshift of $z = 1.2$ to $z \sim 2$ through the inclusion of an additional tomographic bin over that redshift range. This would boost the sensitivity of the measurements to cosmic shear, but it also can significantly influence the sensitivity of the shear measurements to systematic effects, such as variable depth. For this reason, it is important that the forward simulations account for the additional depth. KiDS-1000 (Asgari et al., 2021) used five tomographic bins for the source galaxy population with their boundaries in $z \in \{0.1, 0.3, 0.5, 0.7, 0.9, 1.2\}$. To forecast any possible increases in the depth in the analysis of KiDS-Legacy cosmic shear data, we choose the same tomographic binning scheme as for KiDS-1000, while including an additional bin between $z = 1.2$ and $z = 2.0$. To incorporate the high-redshift galaxies which populate this tomographic bin into the mock KiDS-Legacy catalogue, the organised randoms have been trained on a version of the KiDS-1000 DR4 galaxy catalogue without the GOLD redshift selection applied, such that galaxies with $z > 1.2$ may be included.

Applying this binning scheme to the mock KiDS-Legacy galaxy catalogue, we obtain the redshift distributions shown in Figure 4.1 for each of the six tomographic bins. These redshift distributions have been obtained through a separate SOM which maps the observed photometric redshifts to spectroscopic redshifts based on refer-

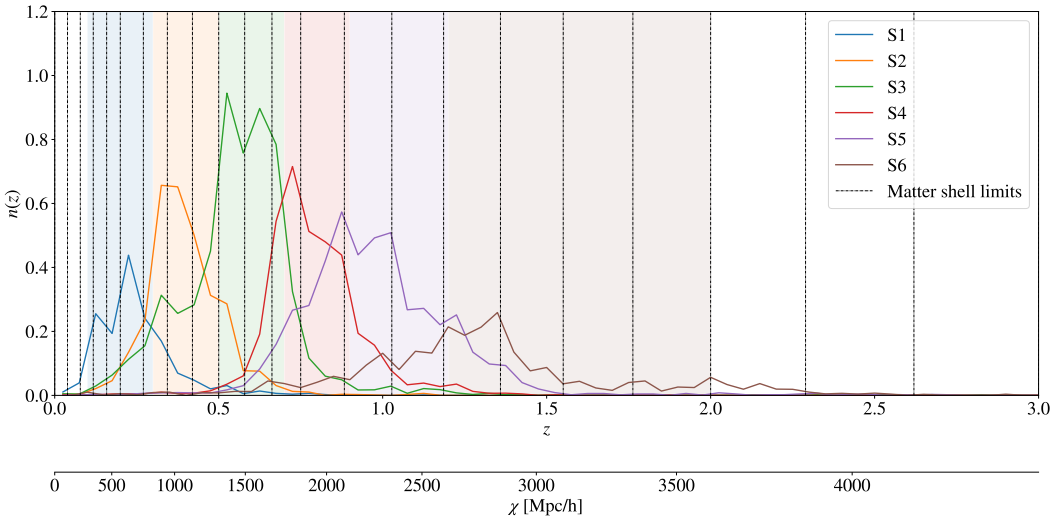

Figure 4.1: Plot of the redshift distributions of the six KiDS-Legacy-like tomographic bins (S1 to S6). The shaded areas show to limits of each tomographic bin, while the solid lines show the $n(z)$ of the source galaxies in each tomographic bin as a function of both redshift, z, and comoving distance, χ (the latter is derived assuming a Planck 2018 cosmology; Planck Collaboration et al. 2020). The black dashed lines show the limits of the spherical matter shells in our forward simulations.

ence spectroscopic samples, in accordance with the approach taken for KiDS-1000 (Wright et al., 2020a; Hildebrandt et al., 2021). We notice that the redshift distribution of the sixth tomographic bin, S6, is exceptionally broad which can significantly impact its response to systematic effects when compared to narrower bins. All other redshift distributions are virtually identical to the fiducial KiDS-1000 redshift distributions as defined in Hildebrandt et al. (2021) with only minor differences due to sampling noise from the realisation of the organised randoms and the subsequent retraining of the SOMs for redshift calibration.

With regards to the footprint, an accurate mask or footprint of the observed area on the sky by KiDS-Legacy does not currently exist. To still be able to produce useful predictions for KiDS-Legacy, we opt for the KiDS-1000 mosaic mask as shown in Figure 3.1. This implies that the footprint of the forward simulations only spans 1,000 deg^2 (with an effective area of 773.3 deg^2) instead of the expected 1,350 deg^2 which are shown in Figure 4.2. This means that we expect the amplitude of the noise on the cosmic shear signal to be overestimated as the effective area

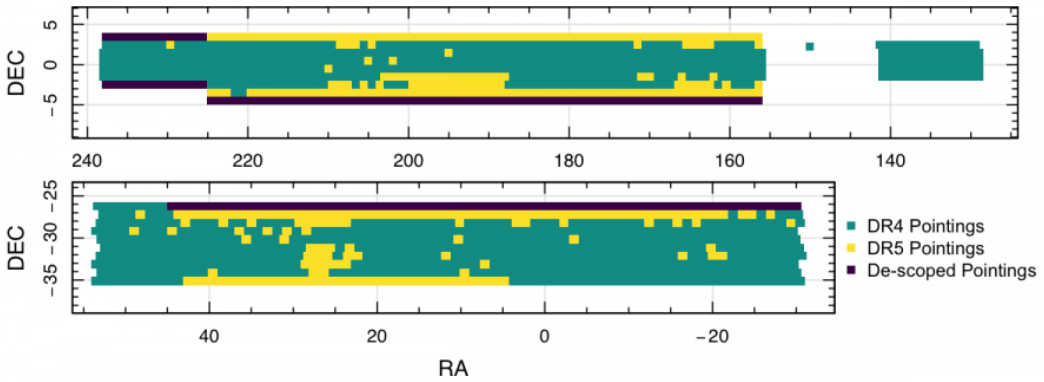

Figure 4.2: Spatial map of the expected KiDS DR5, i.e. KiDS-Legacy, and the KiDS DR4/KiDS-1000 footprints. The green areas show the pointings which are included in KiDS DR4, while both the green and the yellow pointings are included in KiDS DR5. The purple pointings show the pointings which were excluded, i.e. "de-scoped", from the final KiDS data release. The top and bottom panels show Cartesian projections of KiDS-North and KiDS-South fields, respectively. Figure courtesy of Angus H. Wright and the KiDS team.

covered by the mocks is less than the expected KiDS-Legacy area. Nevertheless, it is still sufficient to explore the effects of the geometry of the footprint. As can be seen from Figures 3.1 and 3.9, the KiDS-1000 mosaic mask has many holes, which can be of the size of whole pointings, stars or a few pixels. It is expected that the KiDS-Legacy footprint will not feature any pointing-sized holes (shown in yellow in Figure 4.2). In addition to this, the KiDS-Legacy footprint is going to have a similar density of masked stars, transients or diffraction spikes at the sub-degree scale as seen in KiDS-1000. With the expectation that the geometry of the KiDS-Legacy data will only become less fractured than in KiDS-1000, the KiDS-1000 footprint provides a conservative estimate for the complexity of the footprint edges and holes. Hence, any estimates on the relative impact of the footprint on the cosmic shear signal or noise can be viewed as upper limits for an equivalent KiDS-Legacy analysis. Similarly, since the variance of the cosmic shear signal is expected to approximately scale with the inverse of the survey footprint area, the estimated uncertainties from the forward-simulations based on the KiDS-1000 footprint are expected to be smaller than the true uncertainties of the cosmic shear signal as it will be observed by KiDS-Legacy.

4.2 Forward Simulations

The forward simulations pipeline in this work is based on the KiDS-SBI suite of simulations described in Chapter 3. The pipeline allows to simulate realistic galaxy shape and redshift catalogues based on log-normal random fields (see Section 1.4.3.1) which encode cosmological large-scale structure in an efficient manner. For the purposes of testing the signal and noise modelling for KiDS-Legacy, the calibration of the inputs and the scope of the forward simulations from KiDS-SBI differs significantly from the ones shown in Chapter 3 for the purposes of a simulation-based inference analysis of KiDS-1000 cosmic shear.

In terms of the cosmic shear signal, two-point correlators are constructed such that they are insensitive to the survey footprint geometry and size. Thus, the typical assumption of a contiguous square footprint does not have an effect on the mean signal. Another typical assumption in signal and noise models is that the sensitivity of the galaxy survey is homogeneous and isotropic throughout the survey's footprint. In reality, as discussed in Chapter 3, the sensitivity and therefore observational depth can have spatial variations. This can be caused by variations in the survey's seeing, changes in sky transparency, variations in the amount of overlapping pointings for a given position on the sky, etc. (see Section 3.2.8 for a more detailed explanation). As these variations may have coherent spatial patterns, they can induce additional modes into the two-point estimators which may have significant effects on the cosmic shear signal (Heydenreich et al., 2020; Joachimi et al., 2021; Baleato Lizancos & White, 2023).

In the case of single-visit surveys, such as KiDS, this variability in depth can be exacerbated further as a survey progresses over time, and obtains observations over a larger time-span such that long-period transient changes in the depth imprint themselves on the spatial variations of the galaxy observations. Note that for surveys such as DES or LSST which visit each pointing in the sky at multiple points in time, such variations decrease with time. Additionally, if one expands the redshift depth of the galaxies taking into consideration for the cosmic shear analysis as is the case for KiDS-Legacy, the sensitivity of the two-point correlators to the variations in

depth may change. The inclusion of an additional sixth tomographic bin in KiDS-Legacy implies the inclusion of more faint galaxies in the galaxy sample which are affected by variations in depth. Due to selection effects, high-redshift galaxies also tend to be closer to the magnitude limit of KiDS as they are more likely to be distant and faint. As discussed in Section 4.1, the variations in depth are correlated with the magnitude limit, so they may have strong effects on the galaxy population in a high-redshift tomographic bin. Hence, despite previous analyses finding that the effect of depth variability on the signal in KiDS-1000 is below its noise ceiling (Heydenreich et al., 2020), this is not necessarily the case for KiDS-Legacy and reevaluation is required.

The same applies for the uncertainty modelling for the cosmic shear measurements in KiDS. The variation of the depth does not only introduce additional random spatial variance across the sky, it also alters the local density and shape dispersion of galaxies which directly set the shape noise. Moreover, contrary to the signal, the uncertainty of cosmic shear is not insensitive to the geometry of the survey footprint. The KiDS-1000 and KiDS-Legacy footprints, as seen in Figure 4.2, are made up of two disjoint fields which are elongated along right ascension and only cover about 10 deg each in declination. The KiDS-1000 mask contains many holes of pointings. Although KiDS-Legacy is expected to include these missing pointings, as with KiDS-1000, it is going to include many subdegree-sized holes due to masked stars, diffraction spikes and other contaminants. For this reason, we expect the KiDS-1000 mask to allow us to obtain reasonably compatible estimates of the effect of survey geometry when incorporating an additional source bin, S6, as will be the case in KiDS-Legacy.

To accomplish this, we put forth KIDS-SBI which employs log-normal random matter fields to model the underlying cosmic shear signal and contaminate it with the relevant systematics. These are adequate for this purpose, as log-normal random fields accurately recover two-point statistics (Tessore et al., 2023), while also including higher-order statistic information (Hall & Taylor, 2022) relevant for covariance estimates. It is important to mention that the KIDS-SBI simulations

Parameter	Symbol	Value
Hubble constant	h	0.670
Cold dark matter density	Ω_c	0.276
Baryonic matter density	Ω_b	0.050
Curvature density	Ω_k	0
Scalar spectral index	n_s	0.960
Matter density fluctuation amplitude at 8 Mpc h^{-1}	σ_8	0.786
Dark energy equation of state	w	-1
Sum of the neutrino masses	$\sum m_\nu$	0.06 eV

Table 4.1: Fiducial choice of cosmological parameters within ΛCDM used for the forward simulations to test the signal and noise modelling for KiDS-Legacy. The given cosmology is equivalent to $S_8 \equiv \sigma_8[(\Omega_b + \Omega_c)/0.3]^{1/2} = 0.82$. The parameters are set in accordance with the constraints from KiDS-1000 (Asgari et al., 2021), while σ_8 is set to be the mid-value between KiDS-1000 and Planck (Planck Collaboration et al., 2020).

are constructed to reproduce two-point statistics, such that they are not necessarily allow accurate validation of model for higher-order statistics.

4.2.1 Structure and Modelling Choices

The pipeline used in this analysis is based on the KiDS-SBI pipeline described in Chapter 3. The pipeline is implemented within the Cosmological Survey Inference System (CosmoSIS; Zuntz et al. 2015) and based on the KiDS Cosmology Analysis Pipeline (KCAP; Joachimi et al. 2021; Asgari et al. 2021; Heymans et al. 2021). Within this framework, we run approximately 5,000 realisations of three different survey settings following similar analysis choices as in Joachimi et al. (2021) while assuming a fixed flat ΛCDM cosmology (see Table 4.1 for the chosen fiducial cosmology), and only varying the random seed between realisations in order to accurately sample cosmic variance.

As shown in the diagram in Figure 4.3, the cosmology dependence of all instances of the forward simulations is set by calculating the linear three-dimensional matter power spectrum, $P_{\delta,l}$, with CAMB (Lewis et al., 2000; Lewis & Bridle, 2002; Howlett et al., 2012) based on the parameters shown in Table 4.1. Although analytical covariance models can consider non-linear matter power spectra, most analytic covariance models (including the analytical model used in KiDS) do not take into consideration the uncertainty in the non-linearities in the power spectrum

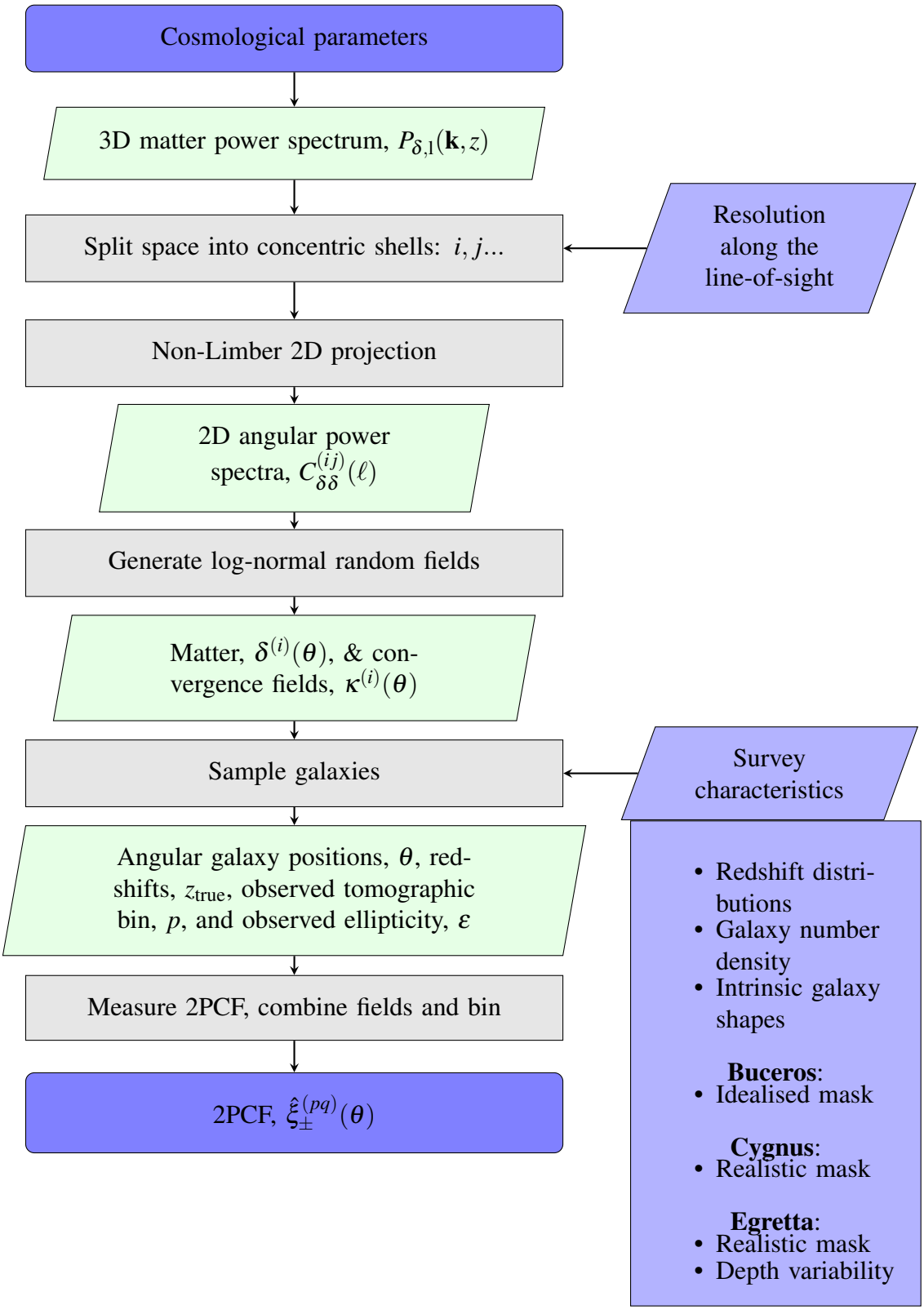

Figure 4.3: Flowchart describing the steps in a single forward simulation of cosmic shear observables from cosmological parameters used for the KiDS-Legacy signal and uncertainty modelling. The dark blue rounded boxes represent the inputs and outputs which are given to the simulation-based inference pipeline. The green slanted boxes represent relevant quantities which are calculated during the simulation. The grey rectangular boxes show steps in the calculations, while the blue slanted boxes show any (systematic) effects which are included.

due to baryonic feedback. Despite KiDS-SBI being capable of modelling the uncertainty on the cosmic shear signal while including the contributions due to baryonic feedback, we evaluate both the analytical and numerical models assuming a linear matter power spectrum. This ensures consistency between the models on the treatment of the uncertainty on small scales. Hence, it allows us to isolate the effects of observational systematics on the uncertainty model, even at small scales.

We note that the assumption of a linear matter power spectrum does not allow us to account for any potential dependence of observational systematics, such as the survey geometry and variable depth, on the non-linear contributions to the cosmic shear signal from baryonic feedback and gravitational non-linearities. Moreover, it also important to note that, in general, baryonic feedback cannot be ignored when modelling realistic cosmic shear observations at small scales. Therefore, in order to realistically model the signal and uncertainty expected in KiDS-Legacy, a necessary extension to the work shown in this chapter is the re-evaluation of the analytical and numerical models while assuming a non-linear matter power spectrum, as is the case in the analysis shown Chapter 3. Having said this, due to limitations in the available computational resources, this extension is outside of the scope of this work.

Upon setting the input matter power spectrum, we set the geometry of the simulations as a sphere made up of concentric shells which are discretised along the line-of-sight and which place the observer at the centre. The dashed vertical lines in Figure 4.1 show the boundaries in redshift/comoving distance space of the 22 shells between $z = 0$ and $z = 3$ which make up each simulation. As can be seen in the figure, these shells are chosen such that they cover the entirety of the probability density mass of the KiDS-Legacy redshift distributions. At the same time, the shells' widths are determined such that they are narrow enough to accurately represent the matter overdensity field along the line-of-sight, while also being wide enough such that the assumption of log-normality of the fields remains accurate (for widths of comoving distance, χ, less than $100\,\mathrm{Mpc}/h$ this breaks down; Hall & Taylor 2022; Piras et al. 2023).

Taking the kernels of each concentric shell of a fixed comoving volume, we

project $P_{\delta,\mathrm{l}}$ into two-dimensional angular power spectra correlating the matter fields, $C_{\delta\delta}^{(ij)}$, within given shell pairs, (ij). $C_{\delta\delta}^{(ij)}$ is calculated using LEVIN (see Section 3.2.3 for details). These angular two-point correlators between the 22 shells set the cosmology dependence for the log-normal random matter fields within each shell. To this end, KIDS-SBI implements the Generator for Large-Scale Structure (GLASS; Tessore et al. 2023) to efficiently sample the matter overdensity field in each shell, $\delta^{(i)}$, and the associated matter convergence fields, $\kappa^{(i)}$ (see Sections 1.4.3.1 and 3.2.4 for details on this calculation). As with non-linearities, at this stage, we do not include any intrinsic alignments in the model as this not typically fully considered in analytical uncertainty modelling. Up to this point, each of the 5,000 realisations only differs from the others through its seed used when sampling the log-normal random matter field.

Once the matter and convergence fields are set, KIDS-SBI samples galaxy positions and shapes in accordance with KiDS-Legacy's survey characteristics with the aid of Speedy Acquisition for Lensing and Matter Observables (SALMO; Joachimi et al. 2021 and see Sections 3.2.6 and 3.2.7 for details). Here, we distinguish between three sets of forward simulations which each make different choices about sample selections: *Buceros* (a KiDS-1000-sized idealised footprint with homogeneous and isotropic sensitivity), *Cygnus* (a realistic KiDS-1000 footprint with homogeneous and isotropic sensitivity), and *Egretta* (a realistic KiDS-1000 footprint with realistic spatial variability in the selection of galaxies, i.e. variable depth). The comparison between each of these simulation settings then allows to determine the effects of the survey footprint geometry and variable depth on the cosmic shear signal and uncertainty.

In the case of *Buceros*, we construct an idealised survey footprint which is a square of approximately 28 deg $\times$ 28 deg with an equivalent area to the KiDS-1000 footprint (773.3 deg^2). This is in line with the underlying assumption of survey area of the cosmic variance terms in the analytic uncertainty model (the shape noise and super-sample covariance terms in the analytical model take into account the realistic KiDS-1000 footprint, see Section 4.3.1). Along the line-of-sight, the galaxies

Bin	$\overline{n}_{\rm gal}$ [arcmin^{-2}]	$\overline{\sigma}_\varepsilon/\sqrt{2}$	$a_{n_{\rm gal}}$	$b_{n_{\rm gal}}$	a_{σ_ε} $\times 10^3$	b_{σ_ε}
S1	0.62	0.27	-0.036	0.73	1.90	0.267
S2	1.18	0.26	-0.040	1.29	1.36	0.256
S3	1.85	0.27	-0.244	2.60	-0.90	0.280
S4	1.26	0.25	-0.254	2.03	-0.68	0.256
S5	1.31	0.27	-0.411	2.53	2.36	0.261
S6	0.57	0.29	-0.241	1.29	4.53	0.276

Table 4.2: Parameters used to sample galaxies and their shapes in line with the expectations for KiDS-Legacy for each tomographic bin (from S1 to S6). $\overline{n}_{\rm gal}$ is the mean galaxy number density for a given tomographic bin, $\overline{\sigma}_\varepsilon/\sqrt{2}$ is the mean per-component shape dispersion, $a_{n_{\rm gal}}$ and $b_{n_{\rm gal}}$ are the slope and y-intercept, respectively, for the linear interpolation of the galaxy density as a function of the root-mean-square of the background noise in the KiDS catalogue, $\sigma_{\rm rms}$, according to Equation (3.30), while a_{σ_ε} and b_{σ_ε} are the parameters to linearly interpolate σ_ε from $\sigma_{\rm rms}$, according to Equation (3.31).

within each tomographic bin are sampled from the redshift distributions shown in Figure 4.1 with a fixed galaxy density, $n_{\rm gal}$ for each tomographic bin. The same is the case when sampling the galaxies' intrinsic ellipticities which are sampled as Gaussian ellipticities assuming a fixed shape dispersion, σ_ε, for each tomographic bin. Table 4.2 shows the values obtained from the mock KiDS-Legacy catalogue created with organised randoms.

For *Cygnus*, we follow the exact same procedure as for *Buceros*, assuming the same redshift distributions, galaxy densities and shape dispersions. The only difference is that *Cygnus* selects angular galaxy positions on the sky based on the KiDS-1000 mask as shown in Figure 3.1.

With *Egretta*, the same mask as for *Cygnus* is used, but the assumption of the homogeneity and isotropy of the galaxy selection function is dropped (see Section 3.2.8). To model this, we choose the $\sigma_{\rm rms}$ value for each galaxy as given in the mock KiDS-Legacy catalogue which gives the root-mean-square of the background noise signal in a given field. This we found to be a good indicator of the local seeing, atmospheric transparency and signal-to-noise, so it can be considered a good estimator of variable depth. In particular, as can be seen in Figure 4.4, $\sigma_{\rm rms}$, is uncorrelated with the *r*-band magnitude measurements of galaxies and their pho-

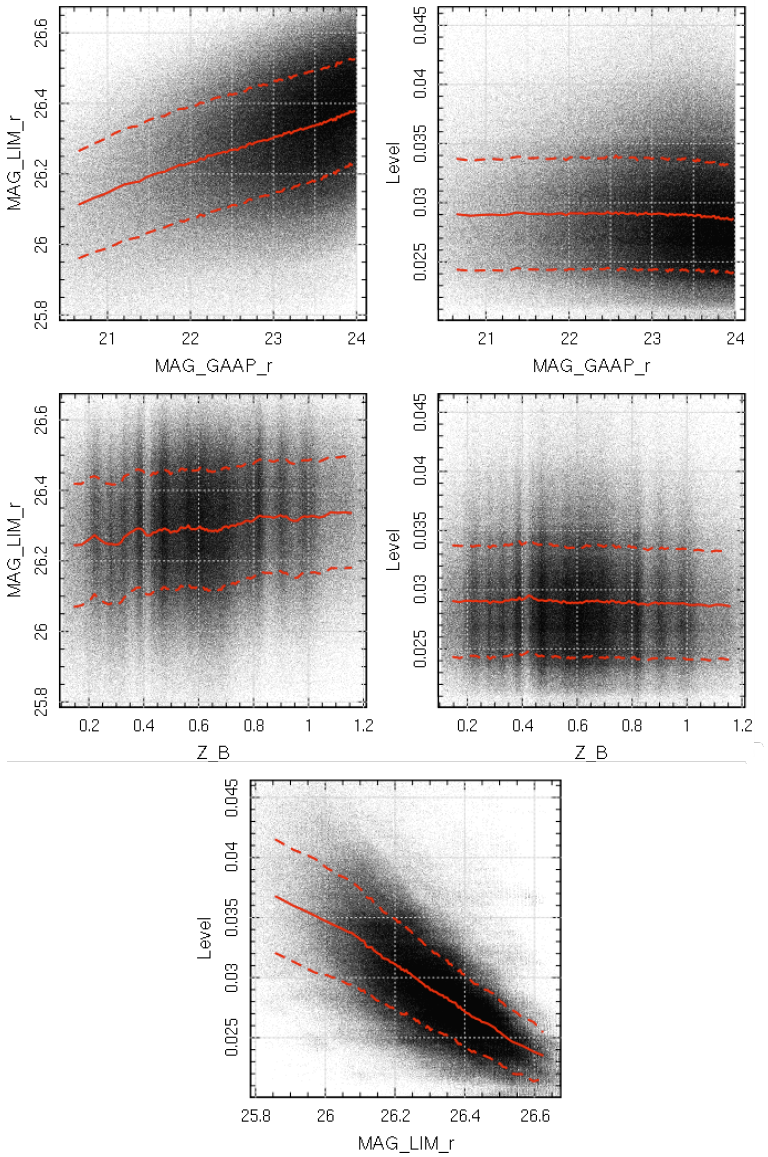

Figure 4.4: Plots showing the underlying correlations of systematic effect parameters with the magnitude measurements in the *r*-band, *MAG_GAAP_r*, the photometric redshifts, *Z_B*, and each other. These variables are the magnitude limit in the *r*-band, *MAG_LIM_r*, and the root-mean square of the background noise in the observations, "Level" or $\sigma_{\rm rms}$. The black dots show the values of galaxies in KiDS DR4 (Kuijken et al., 2019). The solid red lines show the running average, while the dashed lines show the 1σ upper and lower bounds of the running average. Figures courtesy of Angus H. Wright and the KiDS team.

tometric redshift, while being highly correlated with the magnitude limit in the *r*-band which KiDS uses for shape measurements. As $\sigma_{\rm rms}$ increases, the background noise in a given observation is larger, so the local magnitude limit decreases and less galaxies are observed. Hence, $\sigma_{\rm rms}$ can robustly predict the local observational depth independently of the galaxy position and shape measurements.

To calibrate the depth variability from the mock KiDS-Legacy catalogue, we bin each tomographic bin into ten equi-populated bins in $\sigma_{\rm rms}$ with their boundaries in $\{1.70, 2.33, 2.57, 2.80, 3.04, 3.28, 3.52, 3.76, 4.00, 4.23, 12.96\}$. Then, we re-calibrate $n_{\rm gal}$ and σ_ε from the galaxy population within each $\sigma_{\rm rms}$ independently. As shown in Figure 4.5, for both parameters, the dependence on $\sigma_{\rm rms}$ is highly linear for all six tomographic bins. Thanks to this, it is possible to linearly interpolate the value $n_{\rm gal}$ and σ_ε at every point within the KiDS footprint for every tomographic bin from a single spatial map of $\sigma_{\rm rms}$ as the one shown in Figure 3.8. Simultaneously, the variability in the galaxy selection does not only imply that the galaxy density and shape dispersion varies, but also the local redshift distribution. To account for this, we reapply the SOM (Wright et al., 2020a) to estimate the photometric redshift distribution for each $\sigma_{\rm rms}$ bin and tomographic bin independently. This yields 60 redshift distributions which are shown in Figure 4.6. The galaxy number in a pixel and the true redshift of these galaxies is sampled in accordance with a different redshift distribution depending on the associated local $\sigma_{\rm rms}$ value in Figure 3.8. From Figure 4.6, it is apparent that as observational conditions worsen and $\sigma_{\rm rms}$ increases, the number of high-redshift galaxies detected decreases with respect to the number of low redshift galaxies detected. Hence, this shifts the mean of the source redshift distributions which can affect cosmological parameter estimates. In addition, particularly for S1 and S6, the depth variability adds additional fluctuations into the redshift distributions beyond just shifting the mean, which may impact the signal and the noise.

It is important to highlight that, despite the forward simulations presented in this work resembling the ones used in the analysis shown in Joachimi et al. (2021) in scope, they bring forth significant differences and improvements. Firstly, thanks

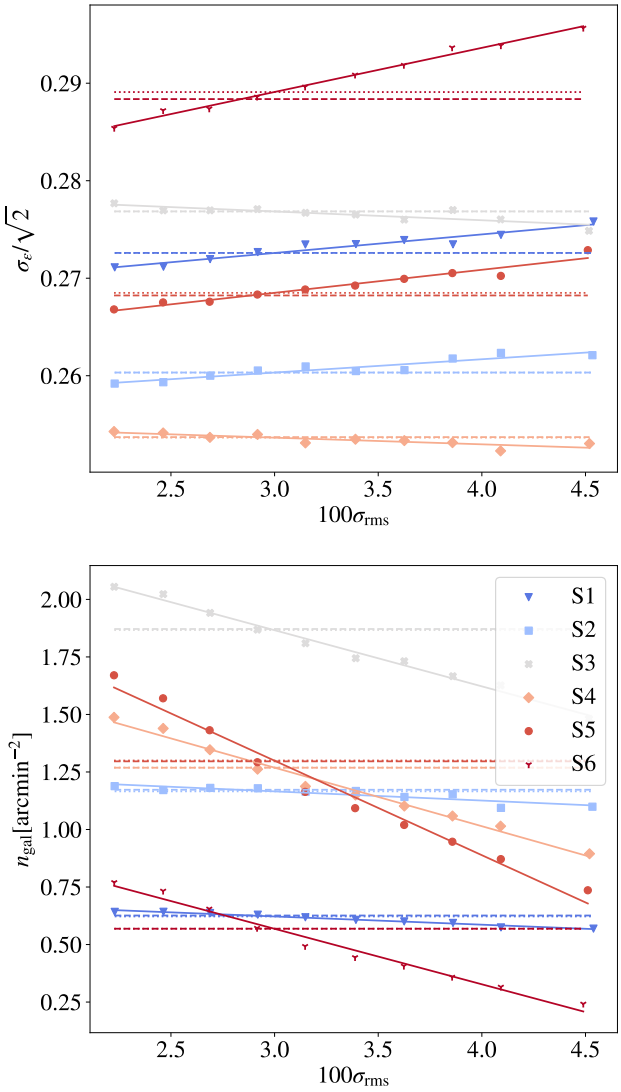

Figure 4.5: Plots showing the dependence of the per-component Gaussian shape dispersion, $\sigma_\varepsilon/\sqrt{2}$, (top panel) and the galaxy density, $n_{\rm gal}$, (bottom panel) on the root-mean square of the background noise, $\sigma_{\rm rms}$ in the KiDS-Legacy-like mock catalogue. For both panels, the data points represent the mean σ_ε or $n_{\rm gal}$ of ten equi-populated bins in $\sigma_{\rm rms}$. The solid line shows the linear fit to the aforementioned data points of their respective tomographic bin according to Equations (3.30) and (3.31). The parameters of this fit for each tomographic bin are shown in Table 4.2. The dotted horizontal lines show the mean values of σ_ε and $n_{\rm gal}$ calculated from the galaxy samples with variable depth per tomographic bin, while the dashed horizontal lines show the values of σ_ε and $n_{\rm gal}$ for the respective galaxy samples without any spatial variations in the observational depth. Both of these lines agree exceptionally well by construction, but some negligible deviations may occur due to rounding errors.

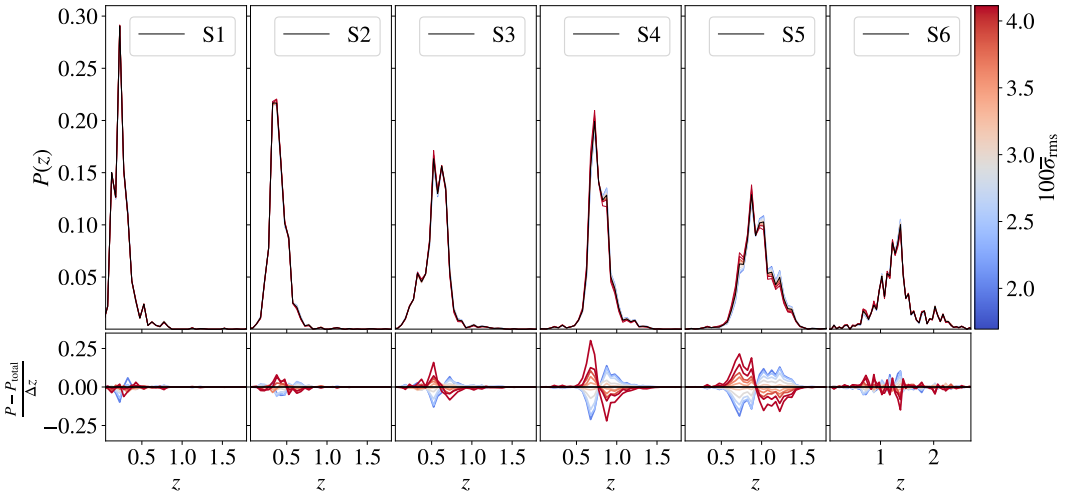

Figure 4.6: In the upper panels, plot of the normalised redshift distributions, $P(z)$, for each tomographic bin (S1 to S6). The redshift distribution from the entire KiDS-Legacy-like mock catalogue constructed from the organised randoms (Johnston et al., 2021), $P_{\rm total}(z)$, is shown in black, while the other ten redshift distributions are derived from 10 equi-populated subsamples of the mock catalogue based on their observational depth (i.e. the mean value of the root-mean-square of the background noise, $\overline{\sigma}_{\rm rms}$) which is shown with its respective colour. The lower panels show the associated residual change in the redshift distributions with respect to $P_{\rm total}(z)$ per unit redshift. It is apparent that variable depth mostly affects the source distributions at high redshifts, while the effect tends to decrease the mean of the redshift distribution with increasing $\sigma_{\rm rms}$.

to better computational efficiency, the resolution along the line-of-sight is increased from 18 shells to 22. Furthermore, instead of using FLASK (Xavier et al., 2016), KiDS-SBI integrates GLASS (Tessore et al., 2023) to compute log-normal random matter fields which are more efficient, while achieving better accuracy in the two-point statistics of the simulated fields. When it comes to the galaxy sampling, the redshift distributions as well as the values of $n_{\rm gal}$ and σ_ε have been recalibrated in line with KiDS DR4 rather than the values used in Joachimi et al. (2021) which are based on KiDS DR3. This also implies that the redshift distributions in this work are based on the SOM (Wright et al., 2020a) rather than the $k^{\rm th}$ nearest neighbour approach used in KiDS DR3 (Hildebrandt et al., 2017) which has been found to be less robust (Wright et al., 2020b). Lastly, the most substantial change with respect to previous analyses (Heydenreich et al., 2020; Joachimi et al., 2021; Johnston et al., 2021) is the abandonment of the magnitude limit in the r-band, *MAG_LIM_r*, as the main parameter to quantify variable depth. *MAG_LIM_r* is measured r-band

magnitude at 5σ above background in a 2" aperture at a given object's location. As is shown in Figure 4.4, although *MAG_LIM_r* is a direct measure of the magnitude limit in the band in which galaxies are selected, the measurement of *MAG_LIM_r* is highly correlated with the *r*-band magnitude measurement of galaxies, while also being correlated with the galaxies' photometric redshift. This implies that *MAG_LIM_r* does not only depend on the survey characteristics and the observational conditions, but it depends also on cosmological measurements. Thus, calibrating the anisotropy of the galaxy selection function using *MAG_LIM_r* artificially amplifies the cosmological signal in the galaxy sample. Instead, "Level", $\sigma_{\rm rms}$, is a more unbiased estimator of the anisotropy of any survey-specific systematic effect. "Level", $\sigma_{\rm rms}$, is the root-mean-square of the background pixel counts measured at a given object's location. We find that $\sigma_{\rm rms}$ does not correlate significantly with the measured *r*-band magnitude or redshift, while being highly correlated with the *r*-band magnitude limit. Consequently, in the case of KiDS, "variable depth" is not an entirely accurate term, instead it can be considered a "variable background noise" which leads to associated variations in the depth.

For all of the above mentioned reasons, it is not only necessary to reevaluate the signal and noise modelling in KiDS due to the improved fidelity of KiDS-Legacy with respect to KiDS-1000, but also because it is necessary to determine whether the improvements with respect to previous tests allow to reach consistent conclusions.

4.2.2 Summary Statistics

Upon generating 15,000 KiDS-Legacy-like galaxy position and shape catalogues (5,000 realisations of each setting: *Buceros*, *Cygnus* and *Egretta*), we measure the shear two-point correlation functions (2PCF, $\hat{\xi}_{\pm}$) as defined in Equation (1.96) using TREECORR (Jarvis et al., 2004). For this we follow the approach laid out in Giblin et al. (2021) and Joachimi et al. (2021) to ensure consistency. For each realisation, $\hat{\xi}_{\pm}$ is initially measured for the KiDS north (N) and south (S) fields separately with 300 log-spaced bins between $\theta = 0.1$ arcmin and $\theta = 300$ arcmin. The $\hat{\xi}_{\pm}$ for each field are then combined as follows

$$\hat{\xi}_{\pm}^{(ij)}(\theta) = \frac{N_{\mathrm{pairs,N}}^{(ij)}\,\hat{\xi}_{\pm,\mathrm{N}}^{(ij)}(\theta) + N_{\mathrm{pairs,S}}^{(ij)}\,\hat{\xi}_{\pm,\mathrm{S}}^{(ij)}(\theta)}{N_{\mathrm{pairs,N}}^{(ij)} + N_{\mathrm{pairs,S}}^{(ij)}}, \tag{4.1}$$

where $N_{\mathrm{pairs,N}}^{(ij)}$ and $N_{\mathrm{pairs,S}}^{(ij)}$ are the number of galaxy pairs between galaxy population i and j within the north and the south fields, respectively, while $\hat{\xi}_{\pm,\mathrm{N}}^{(ij)}(\theta)$ and $\hat{\xi}_{\pm,\mathrm{S}}^{(ij)}(\theta)$ are the measured shear two-point correlation functions between samples i and j from the north and south fields, respectively. $\hat{\xi}_{\pm}^{(ij)}(\theta)$ is re-binned such that the domain covers nine log-spaced bins between $\theta = 0.1$ arcmin and $\theta = 300$ arcmin giving the measurements shown in Figure 4.7. All three suites of simulations produce highly consistent two-point statistic measurements which all agree that the overall signal amplitude in the S6 tomographic bin can be expected to be approximately twice as large as the signal seen in the correlations with the S5 bin. This hints at KiDS-Legacy having the potential of significantly improving cosmological constraining power when compared to KiDS-1000.

4.3 Uncertainty Modelling

4.3.1 Analytical Uncertainty

Analytical uncertainty modelling is a powerful and widely adopted approach across weak lensing surveys (Asgari et al., 2021; Amon et al., 2022). Here, we test the approach adopted in KiDS-1000 (Joachimi et al., 2021) to determine whether it is apt for KiDS-Legacy. The analytical model operates under the assumption of a Gaussian likelihood (see Section 1.4.1). In that framework, all the uncertainty modelling is contained within the covariance matrix which relates the variances and cross-correlations of the values in the measured data vector. For a two-point statistic, such as $\hat{\xi}_{\pm}^{(ij)}(\theta) \equiv \langle \varepsilon^{(i)} \varepsilon^{(j)} \rangle$, the covariance is defined as follows

$$\mathrm{Cov}(\hat{\xi}_{\pm}^{(ij)}, \hat{\xi}_{\pm}^{(mn)}) \equiv \langle (\varepsilon^{(i)}\varepsilon^{(j)} - \langle \varepsilon^{(i)}\varepsilon^{(j)} \rangle)(\varepsilon^{(m)}\varepsilon^{(n)} - \langle \varepsilon^{(m)}\varepsilon^{(n)} \rangle) \rangle \tag{4.2}$$

$$= \langle \varepsilon^{(i)}\varepsilon^{(m)} \rangle \langle \varepsilon^{(j)}\varepsilon^{(n)} \rangle + \langle \varepsilon^{(j)}\varepsilon^{(m)} \rangle \langle \varepsilon^{(i)}\varepsilon^{(n)} \rangle + U_4\Big(\varepsilon^{(i)}, \varepsilon^{(j)}, \varepsilon^{(m)}, \varepsilon^{(n)}\Big). \tag{4.3}$$

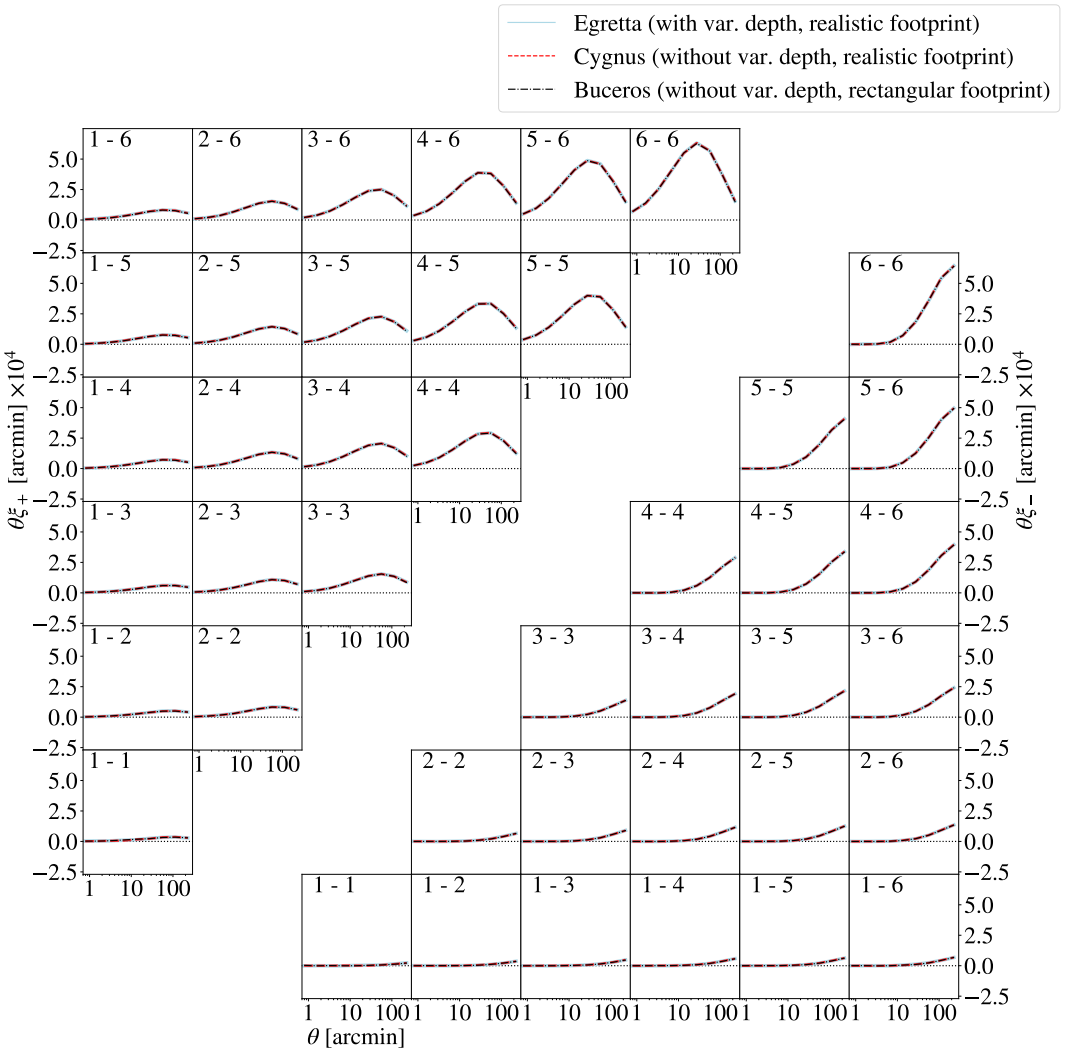

Figure 4.7: Measurements of the mean two-point correlation functions, $\xi_{\pm}$, as a function of angular separation, θ, between 5,000 realisations of KiDS-Legacy-like mock catalogues with the same underlying cosmology. Each line represents different choices for modelling survey characteristics such as the footprint and variable depth. The black dot-dahsed lines show the measurement for *Buceros*, the red dashed lines the measurement for *Cygnus* and the blue solid lines show *Egretta*.

where U_4 is a Ursell or connected correlation function of four random variables.

The first two terms in Equation (4.3) are commonly known as the Gaussian covariance terms as they are the only terms which remain in the covariance of the two-point statistics describing a Gaussian field. It is important to distinguish nomenclature at this stage, when referring to the Gaussanity of terms in the covariance of two-point statistics, it is a separate characterisation than the assumption of a Gaussian likelihood. Both Gaussian and non-Gaussian terms in the covariance assume that the likelihood distribution is a multivariate Gaussian with respect to the two-point data vector. The distinction between Gaussian and non-Gaussian (NG) terms in the covariance, such as the third term in Equation (4.3), refers to whether they depend on two-point correlators alone or on higher-order correlators, respectively. Here, the term Gaussian is used because a two-point correlator fully defines a Gaussian random field.

Taking into consideration that $\hat{\xi}_{\pm}^{(ij)}$ includes intrinsic shape noise as given by Equation (1.97), one can decompose a given Gaussian term as follows

$$\langle \varepsilon^{(i)} \varepsilon^{(m)} \rangle \langle \varepsilon^{(j)} \varepsilon^{(n)} \rangle \propto \mathscr{O}(\xi^2) + \mathscr{O}(\xi^1) + \mathscr{O}(\xi^0), \tag{4.4}$$

where the terms of order $\mathscr{O}(\xi^0)$ are known as shot noise (SN) terms, the $\mathscr{O}(\xi^1)$ terms is known as the mixed terms, and the $\mathscr{O}(\xi^2)$ terms account for sampling variance (SVA) contribution due to observing a finite volume of the Universe.

In addition to these terms, there are some additional effects which are typically considered in the analytical covariance model. The measured shear values, ε, are affected by a multiplicative bias which, when calibrated (Giblin et al., 2021), has associated uncertainties. To propagate these uncertainties into the likelihood, an additional term is added to the covariance. However, since the sky position dependence of the multiplicative shear bias scales most strongly with the local signal-to-noise ratio, its position dependence is already captured directly by the random noise. Any additional contribution from the multiplicative shear bias should not vary with the survey footprint, this term is ignored for the purposes of this analysis. Lastly, due to correlations between modes within the survey and the modes on scales larger than

the survey footprint, there is an additional non-Gaussian contribution to the covariance known as the super-sample covariance (SSC; see e.g. Takada & Hu 2013; Linke et al. 2023).

To summarise, the analytical covariance model can be decomposed as follows

$$\mathrm{Cov}(\hat{\xi}_{\pm}, \hat{\xi}'_{\pm}) = \mathrm{Cov}_{\mathrm{SN}} + \mathrm{Cov}_{\mathrm{mixed}} + \mathrm{Cov}_{\mathrm{SVA}} + \mathrm{Cov}_{\mathrm{NG}} + \mathrm{Cov}_{\mathrm{SSC}}. \tag{4.5}$$

For details on the definitions and parametrisations made to define the terms in Equation (4.5), see Joachimi et al. (2021). We note that the shape noise and super-sample covariance terms take into consideration the realistic KiDS-1000 footprint, while the other cosmic variance terms only take into consideration the effective survey area and neglect any boundary effects due to the survey geometry. Consequently, the analytical model is most comparable to the *Cygnus* forward simulations with the exception of the cosmic variance terms which are more consistent with the model in *Buceros*.

4.3.2 Numerical Uncertainty

To model the uncertainty for KiDS-Legacy numerically, we compute the numerical covariance for each of the simulation suites described in Section 4.2 as follows

$$\mathrm{Cov}(\hat{\xi}_{\pm}, \hat{\xi}'_{\pm}) = \frac{1}{N_{\mathrm{sim}} - 1} \sum_{i}^{N_{\mathrm{sim}}} (\hat{\xi}_{\pm,i} - \langle \hat{\xi}_{\pm} \rangle)(\hat{\xi}'_{\pm,i} - \langle \hat{\xi}'_{\pm} \rangle), \tag{4.6}$$

where N_{sim} is the total number of realisations of a given simulation suite, $\hat{\xi}_{\pm,i}$ is the measured two-point correlation function of a single realisation and $\langle \hat{\xi}_{\pm,i} \rangle$ is the measurements from all N_{sim} realisations.

Applying this to the 5,000 realisations of the three simulation settings, *Buceros*, *Cygnus* and *Egretta*, we obtain the covariance matrices shown in Figures 4.9, 4.10, and 4.11, respectively. As expected, the matrices are dominated by the diagonal terms which are dominated by the shape noise. It is also of note that for all three settings the terms in the covariance are noticeably larger by up to a factor of two in correlations with the S6 bin when compared to correlations with S5 or below.

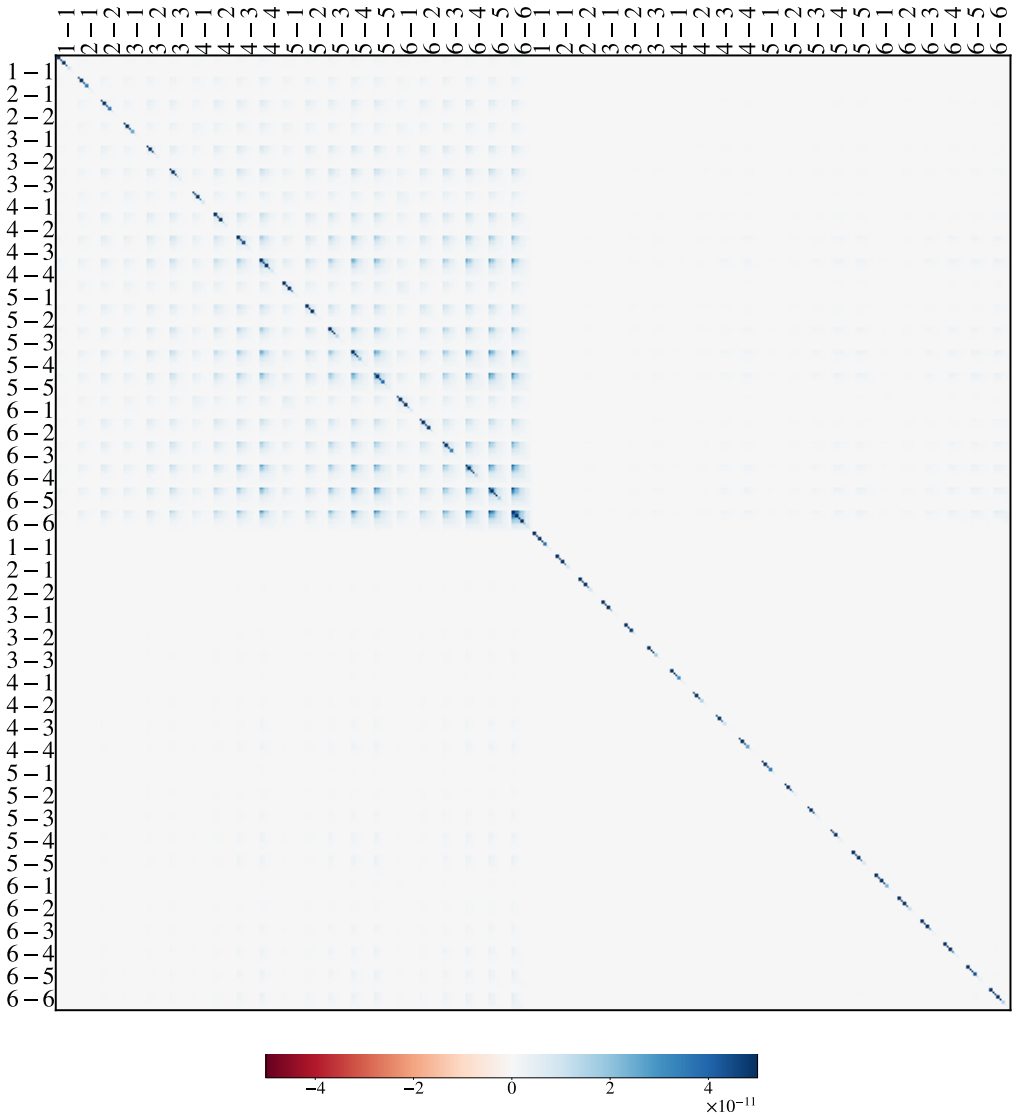

Figure 4.8: Bitmap of the KiDS-Legacy-like analytic covariance matrix. The upper left panels show the ξ_+-ξ_+ covariance for all tomographic bin combinations, with each pixel showing the value for a single log-spaced bin in angular separation, θ, between $\theta = 0.1$ arcmin and $\theta = 300$ arcmin. The upper right and lower left panels show the same, but for the covariance between ξ_+ and ξ_-, while the lower right panels show the covariance of ξ_- with itself.

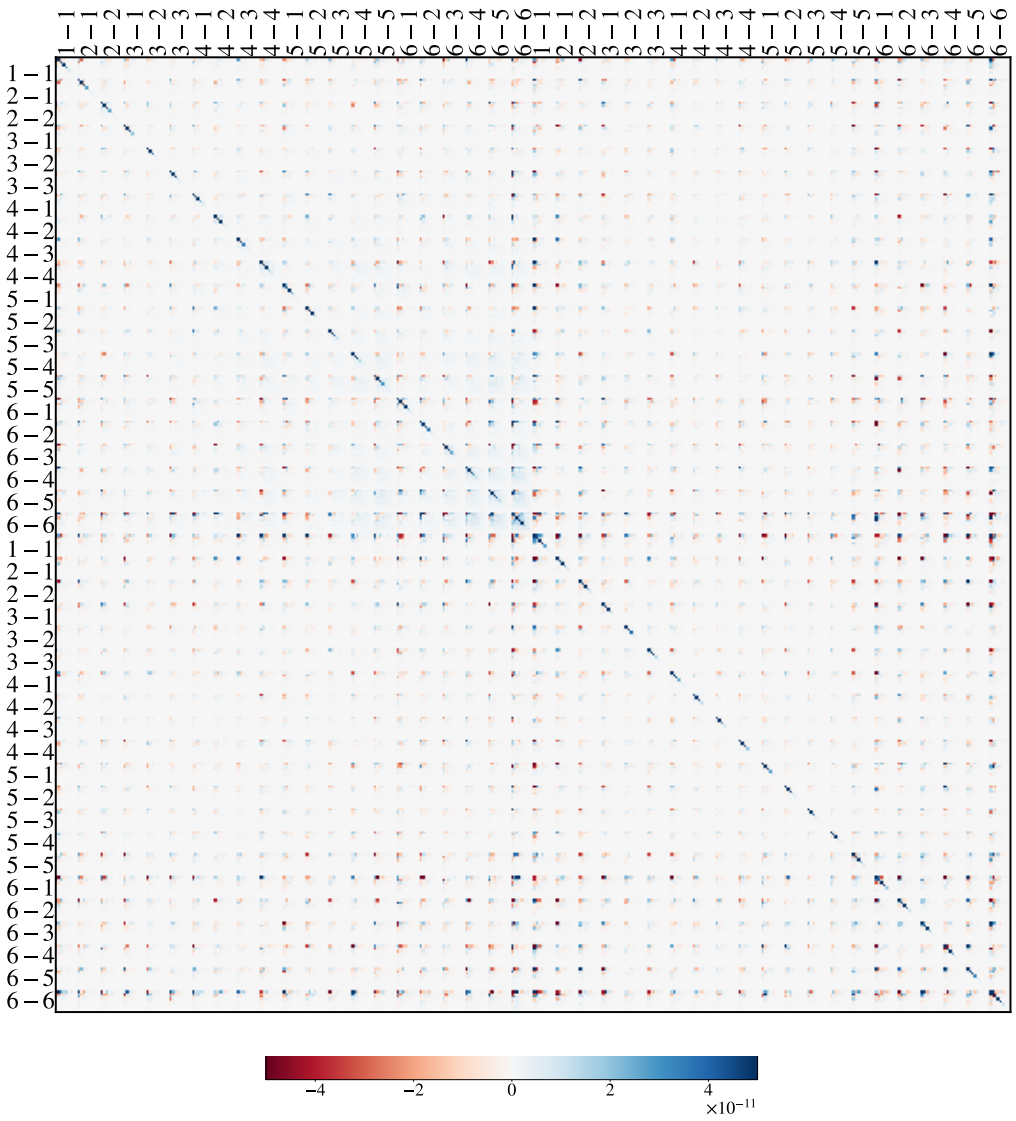

Figure 4.9: Bitmap of the covariance matrix terms as computed from 5,000 numerical realisations of KiDS-Legacy-like *Buceros* (idealised footprint and homogeneous galaxy selection). The upper left panels show the ξ_+-ξ_+ covariance for all tomographic bin combinations, with each pixel showing the value for a single log-spaced bin in angular separation, θ, between $\theta = 0.1$ arcmin and $\theta = 300$ arcmin. The upper right and lower left panels show the same, but for the covariance between ξ_+ and ξ_-, while the lower right panels show the covariance of ξ_- with itself.

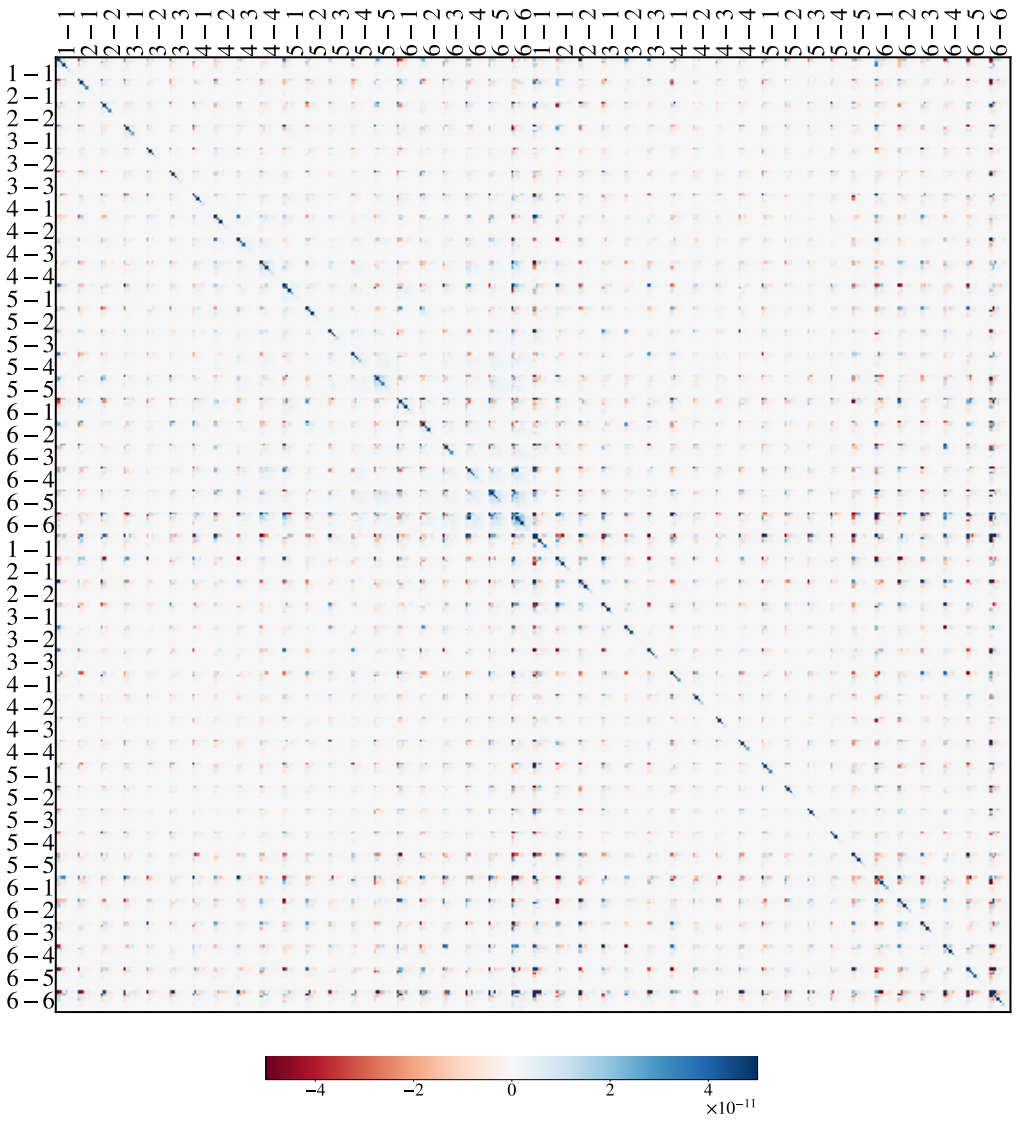

Figure 4.10: Bitmap of the diagonal covariance matrix terms as computed from 5,000 numerical realisations of KiDS-Legacy-like *Cygnus* (realistic footprint and homogeneous galaxy selection). The upper left panels show the covariance between ξ_+ measurements from tomographic bins S1 to S6, with each pixel showing the value for a single log-spaced bin in angular separation, θ, between $\theta = 0.1$ arcmin and $\theta = 300$ arcmin. The upper right and lower left panels show the same, but for the covariance between ξ_+ and ξ_-, while the lower right panels show the covariance of ξ_- with itself.

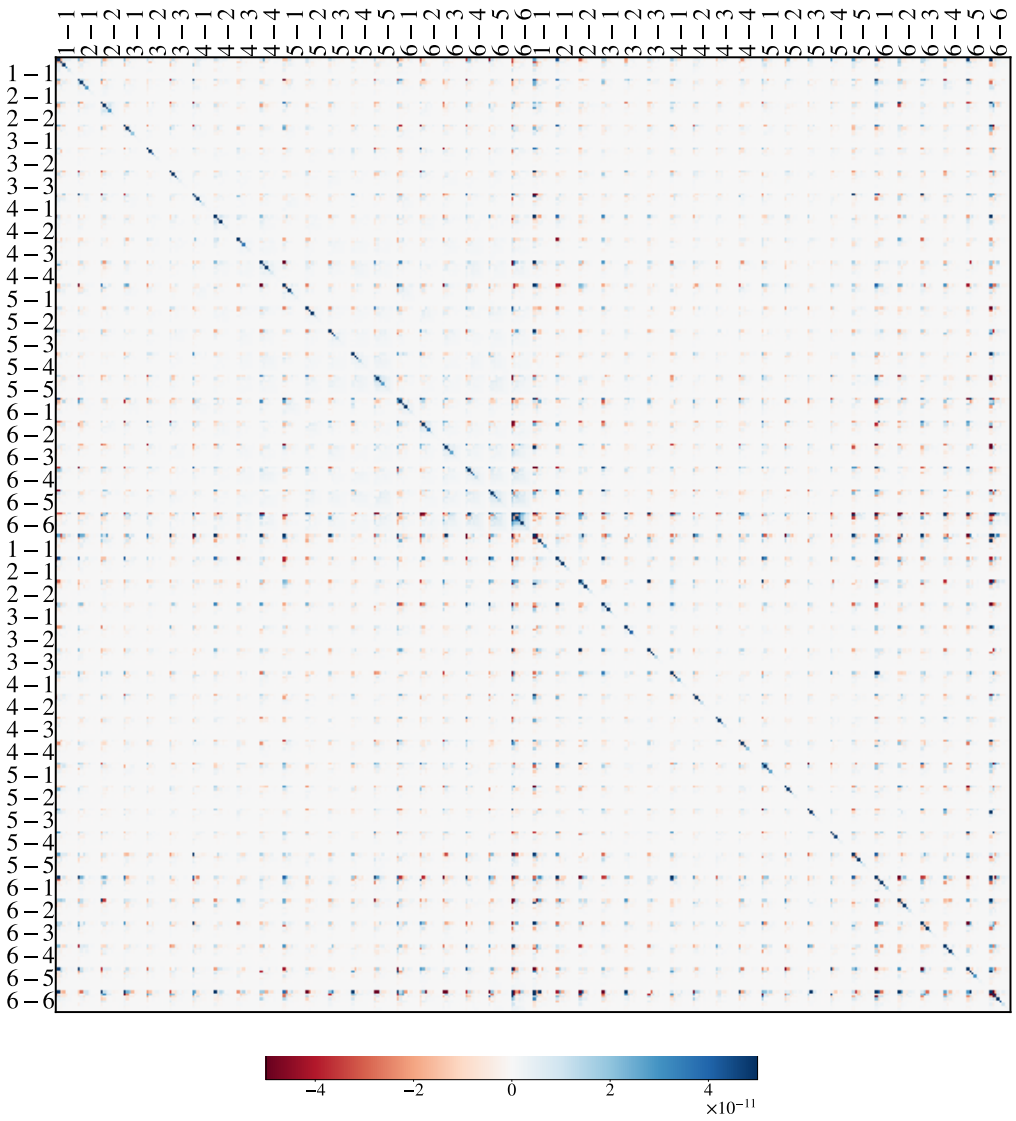

Figure 4.11: Bitmap of the diagonal covariance matrix terms as computed from 5,000 numerical realisations of KiDS-Legacy-like *Egretta* (realistic footprint, and inhomogenous and anisotropic galaxy selection). The upper left panels show the covariance between ξ_+ measurements from tomographic bins S1 to S6, with each pixel showing the value for a single log-spaced bin in angular separation, θ, between $\theta = 0.1$ arcmin and $\theta = 300$ arcmin. The upper right and lower left panels show the same, but for the covariance between ξ_+ and ξ_-, while the lower right panels show the covariance of ξ_- with itself.

4.4 Testing the Signal Model

4.4.1 Impact of the Spatial Footprint

Comparing the signal estimates obtained from the *Buceros* and *Cygnus* simulations as defined in Section 4.2, one can verify whether the two-point statistics are truly robust to the geometry of the footprint as they are expected to be from their construction. Figure 4.12 shows the difference in the mean of two-point correlation function measurements between *Cygnus* and *Buceros*. For all tomographic bins, the signal difference is consistent with zero up to angular separations of $\theta \sim 20$ arcmin. Beyond these scales, the signal difference appears to increase as a function of the signal amplitude up to a $\sim 5\%$ effect at scales near 300 arcmin. Based on the construction of two-point correlation functions, this should not be possible, and as a matter of fact, this is not a physical effect. This is an artifact of the precision settings chosen when computing the 2PCF.

When computing the 2PCF as defined in Equation (1.96) using TREECORR, one can adjust the maximum possible error in the angular separation a galaxy pair can have as a fraction of the bin size (bin_slop). The larger this error is allowed to be, the faster is the computation of the 2PCF, but the accuracy is reduced. In the case of the simulations shown in this work, for the sake of computational efficiency, this parameter has been set to 1.5 times the angular separation between bins (which can imply an error in angular separation of up to 4.7 degrees in the high angular separation bins). This is necessary because computing the 2PCFs for a single realisation with a given selection takes ~ 15 minutes with multi-threading on 32 cores. When scaling this up to 15,000 realisations, this implies about 4,000 CPU hours with 32 cores. For comparison, running TREECORR for a single realisation while setting the allowed error threshold in angular separation to zero entails a run-time of 2.6 hours per realisation using 64 cores for multi-threading. Consequently, increasing the fidelity of the 2PCF calculations for all realisations is currently limited by the efficiency and availability of computational resources.

Upon measuring the 2PCF from 50 realisations of *Cygnus* and *Buceros* while setting $bin\_slop = 0$, we find that the excess in signal from *Cygnus* with respect to

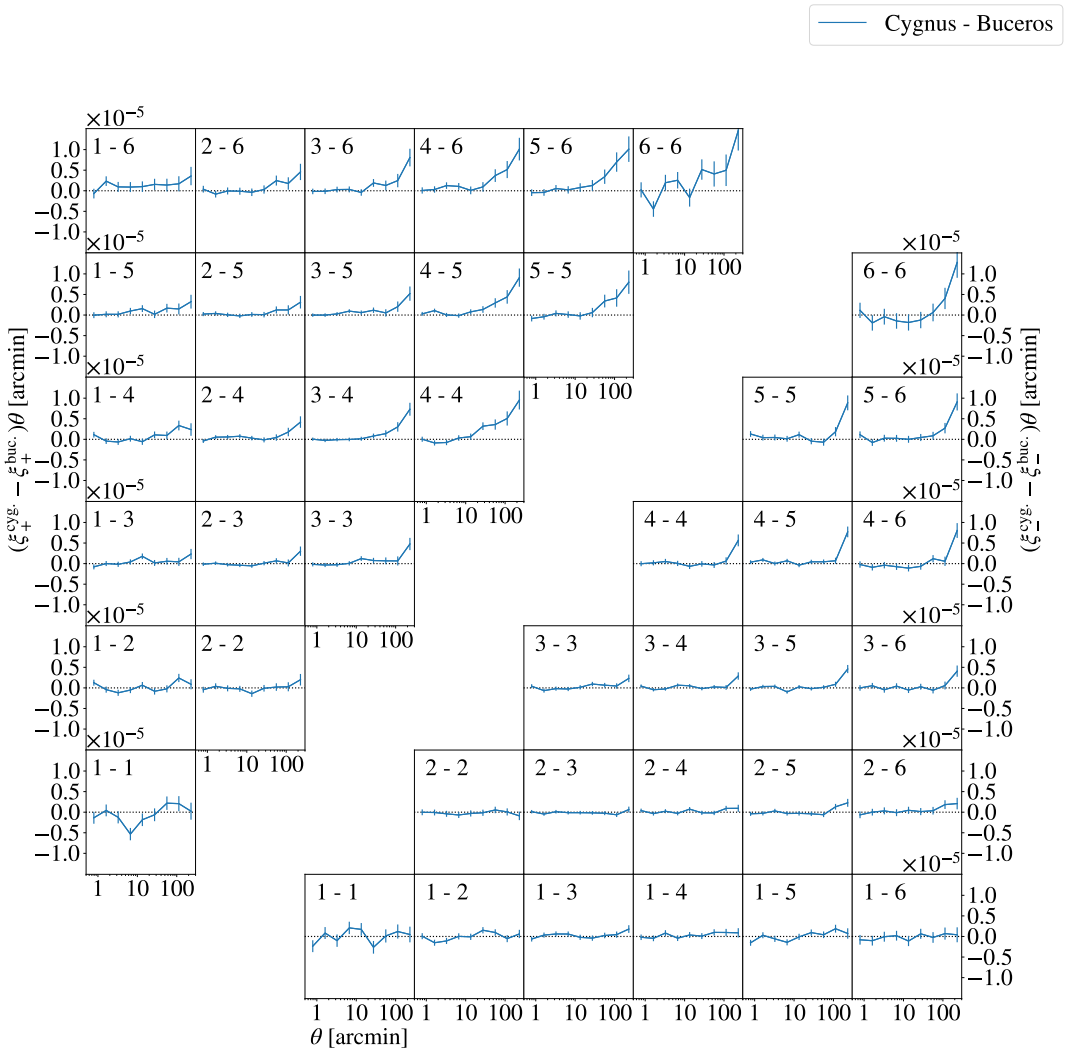

Figure 4.12: Difference between the mean shear two-point correlation functions, $\xi_{\pm}(\theta)$, as measured from 5,000 realisations of *Cygnus* and from 5,000 realisations of *Buceros* with the same underlying cosmology and seeds. The uncertainties shown are the propagated uncertainties of the difference of the means. Each panel shows the difference in signal for a given tomographic bin pair of the bins from S1 to S6.

Buceros at $\theta > 20$ arcmin vanishes. We can conclude that 2PCF are robust against the choice of survey geometry as expected, but this may not necessarily be the case for numerical biases when computing the 2PCF. *Buceros* is defined by a footprint given by a 28 deg $\times$ 28 deg square and *Cygnus* is defined by two disjoint fields of approximately 10 deg $\times$ 90+ deg each. In the former case, the maximum separation between galaxies in the sample is approximately 40 deg, while for the latter it is about 90 deg. Hence, the sensitivity to the choice of bin_slop can vary significantly between samples, and induce significant biases in the measured 2PCF signal.

4.4.2 Impact of Spatial Variability

Conducting a similar comparison between the measured 2PCF from the *Egretta* and *Cygnus* forward simulations, one can isolate the effect of variable depth on the measured signal for a given survey footprint. This is shown in Figure 4.13. We find that for almost all tomographic bin combinations and scales, the difference in signal is roughly consistent with zero. There are some 1σ to 2σ excesses and reductions in the measured signal at some scales, for a few bin combinations, due to the inclusion of variable depth. However, these could just be due to random noise or could be due to inaccuracies from the choice of bin_slop as discussed in Section 4.4.1.

From Figure 4.13, we also conclude that the observed effect of variable depth on the signal is broadly consistent with the predictions from the semi-analytical model from Heydenreich et al. (2020) which is parametrised with the same choices as described in Section 4.2. Hence, we conclude that depth variability is an insignificant effect for the signal modelling for KiDS-Legacy, since the effect on the signal is less than 1% which is below the expected relative random uncertainties.

4.5 Testing of the Uncertainty Model

4.5.1 Quantifying Changes in the Uncertainty

To test the uncertainty modelling for KiDS-Legacy, we compare the analytical model described in Section 4.3.1 with the numerical uncertainty modelling as defined in Section 4.3.2. A simple way to compare covariance matrices is to concentrate on the diagonal elements of the covariance which define the standard deviation

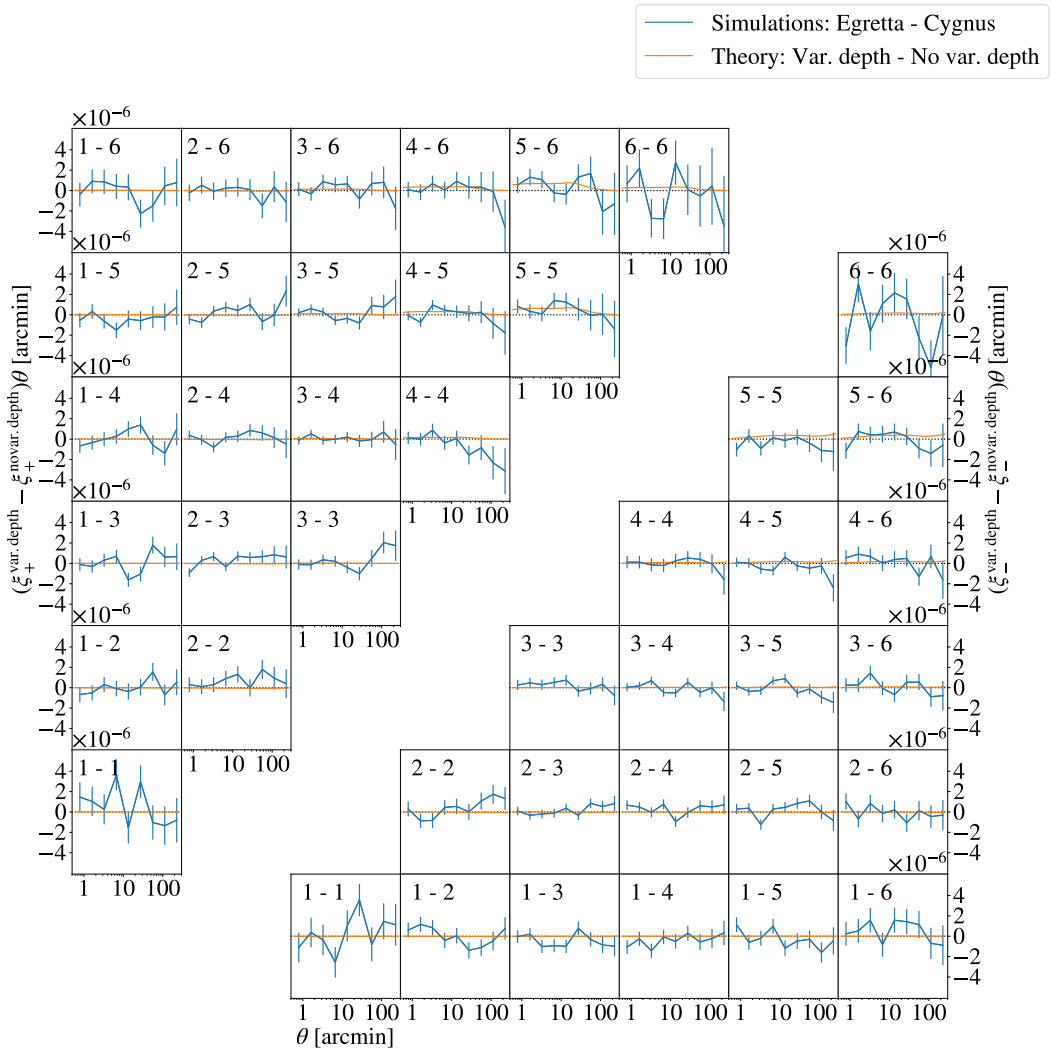

Figure 4.13: In blue, the difference between the mean shear two-point correlation functions, $\xi_\pm(\theta)$, as measured from 5,000 realisations of *Egretta* and from 5,000 realisations of *Cygnus* with the same underlying cosmology and seeds. The uncertainties shown are the propagated uncertainties of the difference of the means. The orange lines show the expected difference in the measured signal due to the inclusion of equivalent depth variability as given by the semi-analytical estimates form the model presented in Heydenreich et al. (2020). Each panel shows the difference in signal for a given tomographic bin pair of the bins from S1 to S6.

of a given data vector element with itself. However, when attempting to compare off-diagonal elements in the covariances, it is often difficult to quantitatively assess how changes in the correlation factor of the covariance are going to affect the final uncertainty. Visual assessment of the changes in the individual elements between two covariance matrices is of course useful, but they do not necessarily give an understanding of the impact such a shift might have in the final constraints, particularly, for high-dimensional covariance matrices. For this purpose, we propose a novel summary statistic to compare covariance matrices based on their relative Förstner-Moonen distance (Förstner & Moonen, 2003).

Förstner-Moonen distance, d_{FM}, is a measure of the distance between two symmetric positive definite matrices, $\boldsymbol{A}$ and $\boldsymbol{B}$, on a Riemannian manifold of $n \times n$-dimensional real, symmetric, positive definite matrices, and is defined as follows

$$d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B}) = \sqrt{\sum_{i}^{n} \ln^2[\lambda_i(\boldsymbol{A},\boldsymbol{B})]}, \tag{4.7}$$

where $\lambda_i(\boldsymbol{A},\boldsymbol{B})$ is the i^{th} root of the following equation

$$\det(\lambda\boldsymbol{A} - \boldsymbol{B}) = 0. \tag{4.8}$$

The Förstner-Moonen distance has properties which make it useful for the purpose of comparing covariance matrices. Firstly, it assumes that the matrices are square, real, symmetric and positive definite, as we expect to be the case for covariance matrices. Secondly, d_{FM} is always positive and it is only zero if $\boldsymbol{A} = \boldsymbol{B}$. Additionally, d_{FM} is invariant under commutation, i.e. $d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B}) = d_{\mathrm{FM}}(\boldsymbol{B},\boldsymbol{A})$, it is invariant under congruence transformations, i.e. $d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B}) = d_{\mathrm{FM}}(\boldsymbol{X}\boldsymbol{A}\boldsymbol{X}^T,\boldsymbol{X}\boldsymbol{B}\boldsymbol{X}^T)$ for any real, $n \times n$ matrix, $\boldsymbol{X}$, and it is invariant under inversion, i.e. $d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B}) = d_{\mathrm{FM}}(\boldsymbol{A}^{-1},\boldsymbol{B}^{-1})$. Hence, any change in d_{FM} can be attributed to translations of $\boldsymbol{A}$ with respect to $\boldsymbol{B}$ along the eigenvector basis of $\boldsymbol{B}\boldsymbol{A}^{-1}$ on a Riemannian manifold of all symmetric, real, positive definite matrices with the same dimensions. As $d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B})$ increases, the two matrices are less compatible and sit further away from each other in the hyperspace of all possible real, symmetric and positive definite

matrices. In the context of covariance matrices, such changes are effectively equivalent to variations in the principal components basis between the two matrices, such that, if $d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B}) \sim 0$, the covariance matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ have similar information contents and constraining power. More details and proofs of the properties of d_{FM} can be found in Förstner & Moonen (2003).

Despite these useful properties of $d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B})$, it can be difficult to interpret whether the nominal value of $d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B})$ is significant or not. To account for this, we propose a measure which relies on the relative Förstner-Moonen distance as follows

$$\Delta d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B};k) \equiv d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B}) - d_{\mathrm{FM}}[\boldsymbol{A},\boldsymbol{A}^{'}(\boldsymbol{A},k)], \tag{4.9}$$

where $\boldsymbol{A}^{'}(\boldsymbol{A},k)$ is a modified version of matrix $\boldsymbol{A}$ as a function of the scalar, $k \in \mathbb{R}$, which is defined as

$$\boldsymbol{A}^{'}(\boldsymbol{A},k) \equiv (1+k)\boldsymbol{A}, \tag{4.10}$$

Therefore, $\boldsymbol{A}^{'}(\boldsymbol{A},k)$ is a matrix with the same eigenvectors as $\boldsymbol{A}$, but all of its eigenvalues are scaled by a factor of $1+k$. By substituting Equation (4.10) into Equation (4.8), we obtain

$$\det\left(\lambda\boldsymbol{A} - \boldsymbol{A}^{'}\right) = 0 \tag{4.11}$$

$$\det\left([\lambda - (1+\mathrm{k})]\boldsymbol{A}\right) = 0 \tag{4.12}$$

$$[\lambda - (1+k)]^n \det(\boldsymbol{D}) = 0 \tag{4.13}$$

$$[\lambda - (1+k)]^n \prod_i^n \lambda_{\boldsymbol{A},i} = 0, \tag{4.14}$$

where $\boldsymbol{D}$ is the diagonalised matrix containing the real eigenvalues of $\boldsymbol{A}$ as its diagonal elements and $\boldsymbol{P}$ is the matrix of the associated unique eigenvectors of $\boldsymbol{A}$, such that $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$, $\lambda_{\boldsymbol{A},i}$ are the eigenvalues of $\boldsymbol{A}$ which make up the diagonal elements of $\boldsymbol{D}$. For non-zero $\lambda_{\boldsymbol{A},i}$, the solution to Equation (4.14) is that $\lambda = 1+k$

for all eigenvalues. Therefore, the Förstner-Moonen distance between $\boldsymbol{A}$ and $\boldsymbol{A}^{'}$ becomes

$$d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{A}^{'}) = \sqrt{\sum_{i}^{n} \ln^2[1+k]} = \sqrt{n\ln^2(1+k)}. \tag{4.15}$$

Note that Equation (4.15) is independent of the eigenvalues and eigenvectors of $\boldsymbol{A}$ and $\boldsymbol{A}^{'}$. Hence, for a given k, $d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{A}^{'}) = d_{\mathrm{FM}}(\boldsymbol{B},\boldsymbol{B}^{'})$. Instead, it only depends on the factor k and the dimensionality of $\boldsymbol{A}$. Consequently, Equation (4.9) becomes

$$\Delta d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B};k) = d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B}) - \sqrt{n\ln^2(1+k)}. \tag{4.16}$$

With this, $\Delta d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B};k)$ can be interpreted as a measure of the excess FM distance between two matrices with respect to the FM distance between two matrices in the same basis with eigenvalues scaled by a factor of $1+k$.

This is useful as it allows to measure the relative information content of the covariance matrix. For a covariance matrix, $\boldsymbol{A}$, one can show that the entropy, S, contained by the Gaussian multivariate probability distribution defined by the covariance, scales as $S \propto \log_2|\boldsymbol{A}| = \log_2|\boldsymbol{D}| = \log_2|\prod_i^n \lambda_{\boldsymbol{A},i}| = \sum_i^n \log_2|\lambda_{\boldsymbol{A},i}|$ (Grainger, 2013). Therefore, the act of scaling of the eigenvalues of a covariance matrix by a factor of $1+k$ increases its entropy by $n\log_2|1+k|$. An increase in entropy implies more noise and uncertainty. Thus, if $\Delta d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B};k)$ is zero, the d_{FM} between $\boldsymbol{A}$ and $\boldsymbol{B}$ is the same as the d_{FM} between two covariance matrices whose entropy differs by $\log_2|1+k|$ per data vector element.

To harness this fact, rather than using $\Delta d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B};k)$ directly as a measure, we opt to use the value of k which minimises $\Delta d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B};k)$ such that

$$k_{\min}(\boldsymbol{A},\boldsymbol{B}) \equiv \arg\min_{k\in\mathbb{R}}[\Delta d_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B};k)]. \tag{4.17}$$

To summarise, $k_{\min}$ is therefore an approximate measure of the value of k for which $d^2_{\mathrm{FM}}(\boldsymbol{A},\boldsymbol{B}) = n\ln^2(1+k)$. In other words, it is the approximate scaling which would have to be applied to the principal components of a covariance matrix for the Förstner-Moonen distance to change by the same amount as the Förstner-Moonen

distance between covariance matrix $\boldsymbol{A}$ and covariance matrix $\boldsymbol{B}$. Alternatively, $k_{\rm min}$ can be interpreted as a measure of the change in entropy, ΔS, between each covariance matrix, given by

$$\Delta S \equiv S^{'} - S \approx n \log_2 |1 + k_{\rm min}|. \tag{4.18}$$

To quantify any systematic uncertainty in the $k_{\rm min}$ metric defined in Equation (4.17), we explore the distribution of $\Delta d_{\rm FM}$ near the minimum $k_{\rm min}$. As shown in the example in Figure 4.14, we generally find that, for the covariance matrices considered in this work, $\Delta d_{\rm FM}$ increases monotonically with both increasing and decreasing values of k. Although the analytic expression for $\Delta d_{\rm FM}$ is currently unknown and beyond the scope of this work, we choose to fit a Lorentzian distribution to $\Delta d_{\rm FM}(k)$ near $k_{\rm min}$ in order to quantify how quickly $\Delta d_{\rm FM}$ approaches its minimum using the following equation

$$\Delta d_{\rm FM}(k) \approx A - \frac{B}{\pi} \frac{\frac{1}{2}\Gamma}{(k - k_{\rm min})^2 - \left(\frac{1}{2}\Gamma\right)^2}, \tag{4.19}$$

where A is the value of $\Delta d_{\rm FM}$ where the gradient of distribution approaches zero, B is the amplitude of the distribution and Γ is the full-width half maximum (FWHM) of the distribution. This fitting is useful as Γ allows us to quantify the width of the trough around $k_{\rm min}$, and hence measure how degenerate $k_{\rm min}$ is with other values of k nearby. If $\Gamma \to 0$, the gradient of $\Delta d_{\rm FM}$ approaches infinity and the distribution becomes a Dirac delta distribution centred at $k_{\rm min}$. Alternatively, if $\Gamma \to \infty$, the gradient of $\Delta d_{\rm FM}$ goes to 0, so $k_{\rm min}$ is fully degenerate with all other values of k. Any values of Γ in between can be interpreted as an inverse measure of the gradient near $k_{\rm min}$ and hence a measure of how degenerate the value of $\Delta d_{\rm FM}(k_{\rm min})$ is with values of $d_{\rm FM}(k)$ near the minimum.

4.5.2 Tests of the Analytic Covariance

To ascertain the accuracy of the fiducial analytic uncertainty model used in KiDS-1000 described in Section 4.3.1 in the context of KiDS-Legacy, we compare the

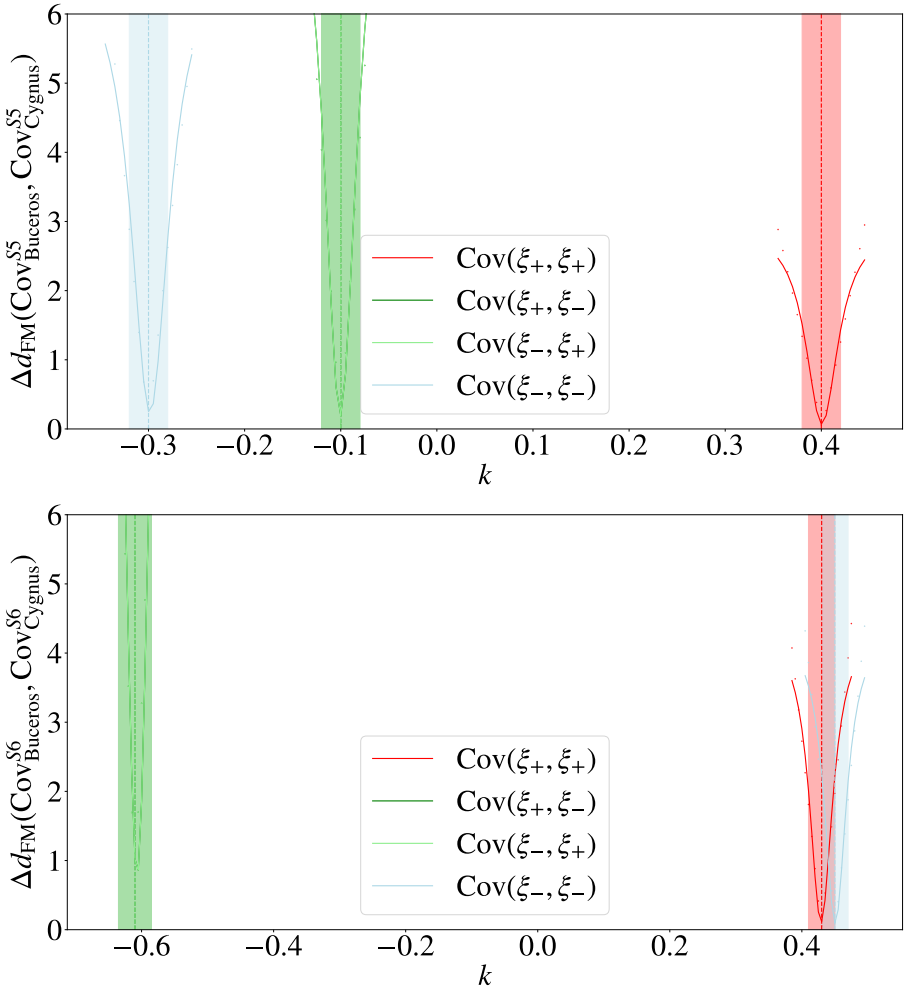

Figure 4.14: Plots of the relative Förstner-Moonen distance, $\Delta d_{\rm FM}$, between two sets of covariance matrices from *Buceros* and *Cygnus* as a function of the fractional change in the reference matrix, k. The top panel shows $\Delta d_{\rm FM}$ between the *Cygnus* and *Buceros* covariance matrices when considering only the first five KiDS-1000 tomographic bins, while the bottom panel shows the same when considering all six KiDS-Legacy-like tomographic bins. The dashed vertical lines show the values of $k_{\rm min}$ for a given covariance matrix pair. The data points show the values of $\Delta d_{\rm FM}(k)$ near $k_{\rm min}$ at intervals in k of 0.005. The solid lines show the Lorentzian fit given by Equation (4.19) of the data points, while the associated shaded region of the same colour shows the FWHM of the Lorentzian around $k_{\rm min}$.

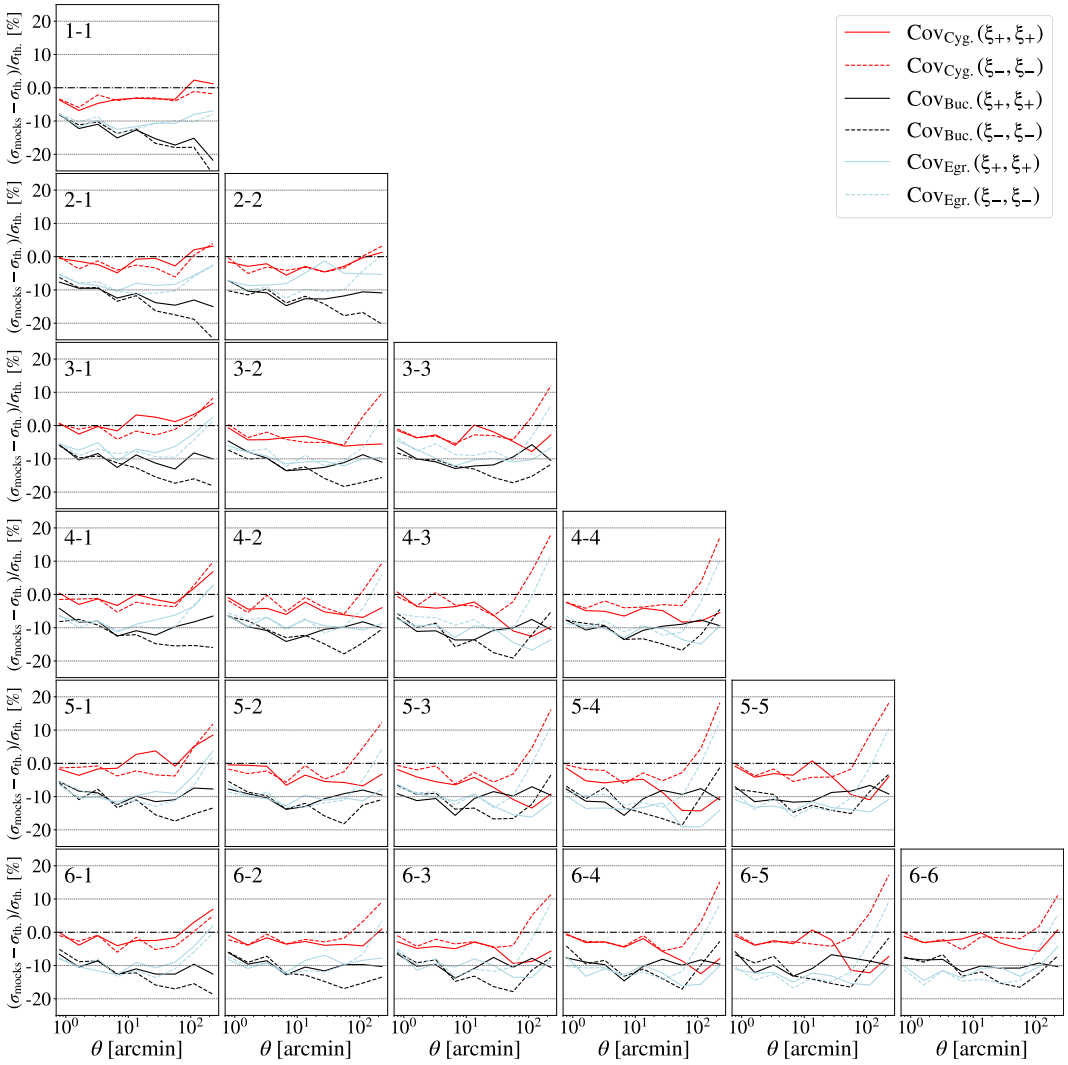

Figure 4.15: Relative difference between the standard deviations calculated from the diagonal elements of the analytic covariance of cosmic shear two-point correlation functions, $\mathrm{Cov}[\xi_\pm(\theta), \xi_\pm(\theta)]$, and the covariance as calculated from 5,000 realisations different mocks (*Buceros* in dashed lines, *Cygnus* in solid lines, and *Egretta* in dotted lines), as a function of angular separation, θ, in arcmin. The orange lines show the relative differences between $\mathrm{Cov}[\xi_+(\theta), \xi_+(\theta)]$, while the blue lines show the same for $\mathrm{Cov}[\xi_-(\theta), \xi_-(\theta)]$. Each panel shows the difference in signal for a given tomographic bin pair of the bins from S1 to S6.

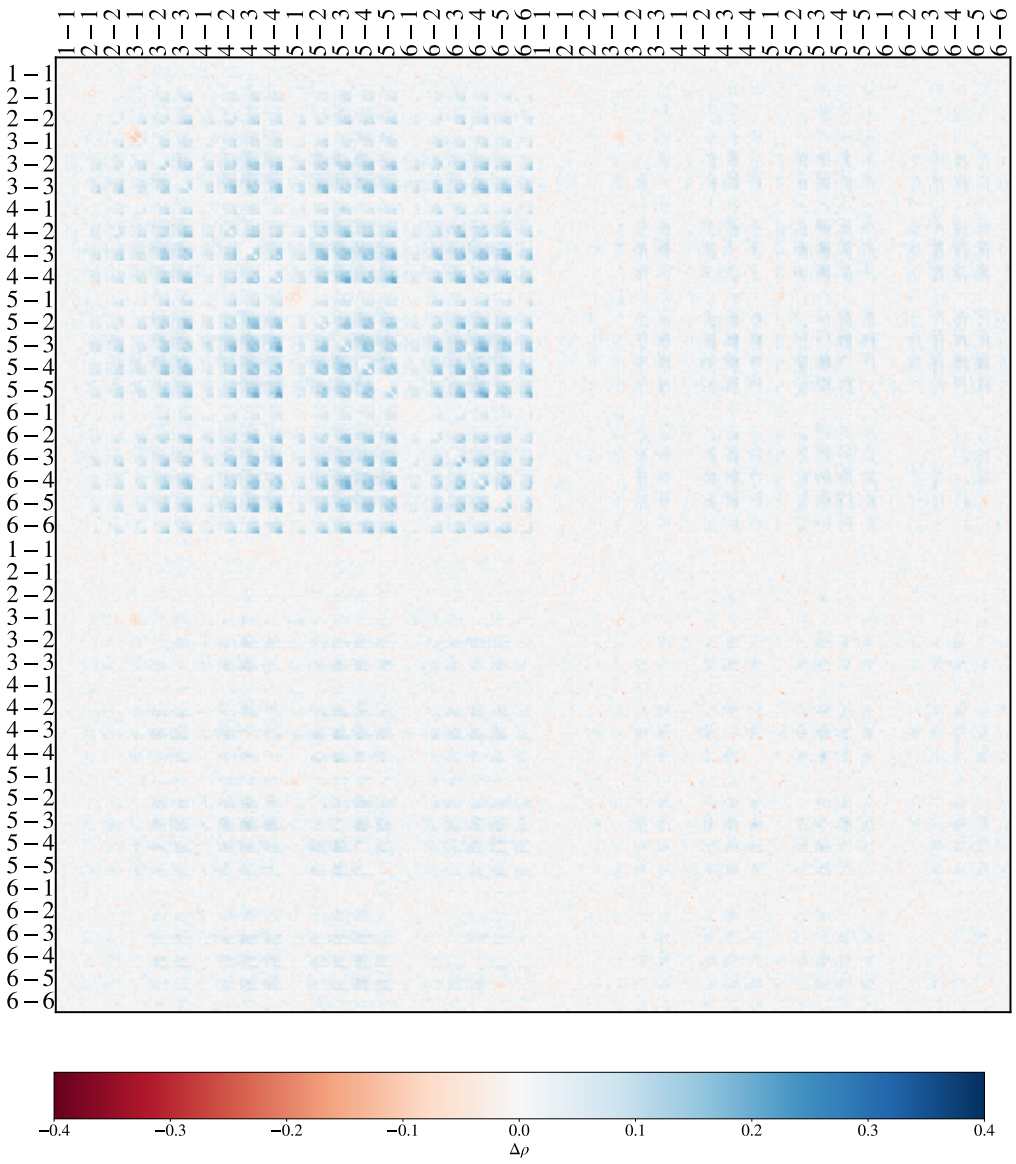

Figure 4.16: Bitmap of the difference in the correlation coefficients, $\Delta\rho$, between the analytic covariance matrix (see Figure 4.8), and the covariance matrix from *Cygnus* (realistic footprint, and homogeneous galaxy selection; shown in Figure 4.10). The upper left panels show $\Delta\rho$ for ξ_+- ξ_+, with each pixel showing the value for a single log-spaced bin in angular separation, θ, between $\theta = 0.1$ arcmin and $\theta = 300$ arcmin. The upper right and lower left panels show the same, but for the change in correlation between ξ_+ and ξ_-, while the lower right panels show the change in correlation of ξ_- with itself.

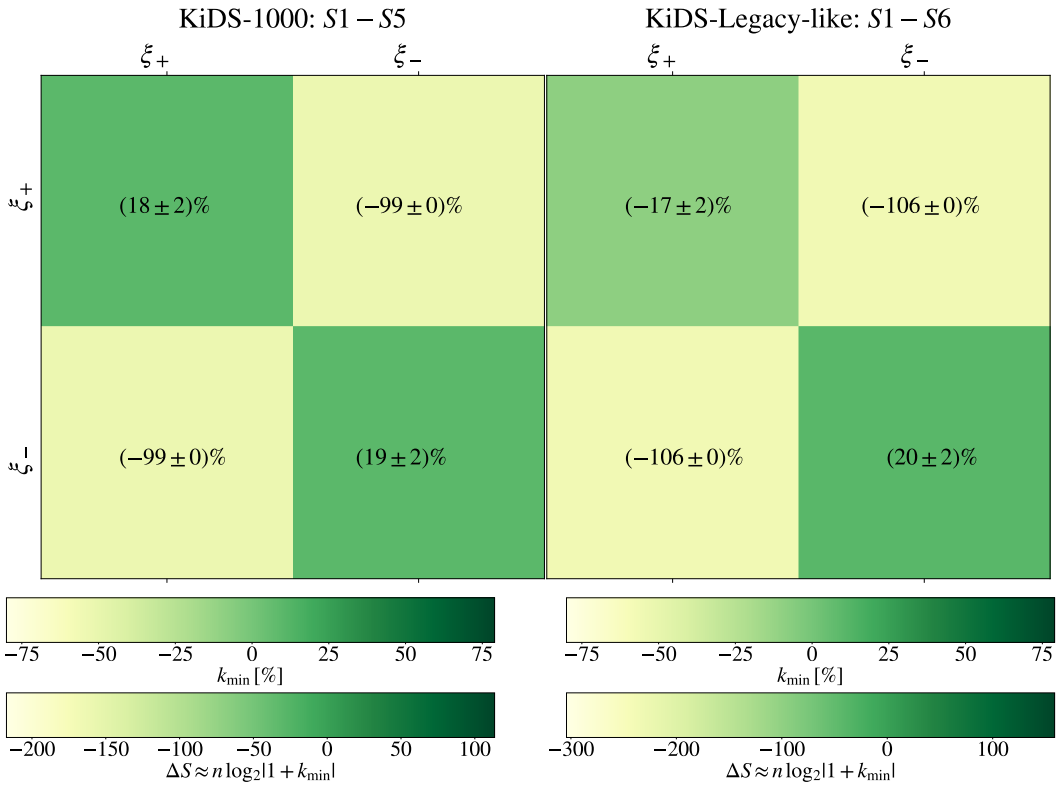

Figure 4.17: Bitmap showing the $k_{\min}$ measure from comparing the analytical covariance matrix, $\mathrm{Cov}(\xi_\pm, \xi_\pm)$, and the $\mathrm{Cov}(\xi_\pm, \xi_\pm)$ as calculated from *Cygnus* as four separate blocks. The uncertainties given for each $k_{\min}$ are a measure of the potential degeneracy of $k_{\min}$ as given by $\Gamma/2$ as shown in Equation (4.19), where $\Gamma \to 0$ implies that $k_{\min}$ is unique and $\Gamma \to \infty$ implies that $k_{\min}$ is not a unique minimum. The left 4×4 bitmap shows $k_{\min}$ when only considering the covariance terms which correlate the uncertainties from tomographic bins S1 to S5 (as in KiDS-1000). The right 4×4 bitmap shows $k_{\min}$ when considering all covariance terms from tomographic bins S1 to S6 (as in KiDS-Legacy). Lighter panels imply that the given block of the analytical covariance matrix is overall less noisy than $\mathrm{Cov}_{\mathrm{Cygnus}}$, darker green panels imply the reverse.

analytical covariance matrix as shown in Figure 4.8 with the numerical covariance matrices obtained from mock observations as described in Sections 4.2 and 4.3.2. Firstly, we compare the analytical covariance, $\mathrm{Cov}_{\mathrm{theory}}$, to the covariance obtained from 5,000 realisations of the *Cygnus* forward simulations shown in Figure 4.10, $\mathrm{Cov}_{\mathrm{Cygnus}}$. These simulations assume a realistic KiDS-1000 footprint, while neglecting any spatial variations in the observational depth. As shown in Figure 4.15, we find that the diagonal elements of both covariance matrices agree at percent-level at all scales and for all tomographic bin combinations. The agreement is ex-

ceptionally good (less than 5%) at small to medium scales. This is not surprising, as the diagonal terms in the analytical model are dominated by shape noise (Joachimi et al., 2021), and the shape noise amplitude is mostly dependent on the survey area, which is the same in both models by construction. We also highlight that the agreement between the uncertainties from *Cygnus* and the analytical model is similar for all bin combinations, including the sixth bin. Therefore, when including the sixth tomographic bin, the analytic uncertainty model is similarly accurate as when only including five tomographic bins.

Figures 4.16 and 4.17 compare the entirety of $\mathrm{Cov}_{\mathrm{theory}}$ and $\mathrm{Cov}_{\mathrm{Cygnus}}$. Figure 4.16 shows that, despite the diagonal covariance terms agreeing exceptionally well between *Cygnus* and the analytic model, the off-diagonal terms in the $\xi_+ - \xi_+$ covariance can have changes in their correlation coefficients of order 10^{-1}. This indicates a slight inconsistency between the numerical simulations and the analytic model in the covariance terms typically dominated by the Gaussian mixed term and the super-sample covariance. Despite the analytic SSC taking into consideration the survey geometry, this may be a consequence of the analytic comic variance model not taking it into consideration which may bias the mixed term and/or change the contributions from the sampling variance. These findings are in line with previous tests of the KiDS-1000 analytic uncertainty model against mocks (Joachimi et al., 2021).

From the generally good agreement between these two models, we also conclude that the underlying assumptions in the modelling of *Cygnus* within KiDS-SBI are broadly consistent with the assumptions made within analytical model. In other words, under the assumption of a Gaussian likelihood, the simulated log-normal random fields from KiDS-SBI recover the signal and the variance of the 2PCF with a similar level of accuracy to current analytical models. This is consistent with the findings in Chapter 3.

As evidenced by Figure 4.17, the inclusion of the sixth tomographic bin does not cause major changes in the agreement between the *Cygnus* and the analytic models. The exception to this is the $\xi_+ - \xi_+$ block of the covariance where $k_{\mathrm{min}} = 18\%$

when considering five tomographic bins, while $k_{\min} = -17\%$ when considering the six KiDS-Legacy-like tomographic bins. This implies that the inclusion of the sixth tomographic bin causes the analytic covariance matrix to be less noisy overall than the one computed from *Cygnus*, while the reverse is the case when excluding the sixth bin. Although this effect seems small, it is of note as it indicates that the analytic uncertainty model becomes less conservative when including the sixth bin, so it may be underestimating the uncertainty of the cosmic shear signal in ξ_+ with respect to the uncertainty estimates from *Cygnus*.

Moreover, we note that the $\xi_\pm - \xi_\mp$ matrix blocks in Figure 4.17 consistently have a $k_{\min}$ near 100%. This is an expected consequence because the measure can only reconcile both matrices by scaling all terms in the *Cygnus* covariance matrix to near zero, as predicted by the analytical model. The terms in the $\xi_\pm - \xi_\mp$ blocks of the covariance matrix from *Cygnus* can be orders of magnitude larger than the equivalent terms in the analytic covariance due to random variations in the numerical realisations.

Repeating the same comparison of the analytic model against the covariance matrix from *Buceros* (see Figure 4.9), we find the agreement between the two models changes as expected. Figure 4.15 shows that the standard deviation in both ξ_+ and ξ_- is consistently approximately 10% smaller across all scales in the *Buceros* model than in the analytic model. The diagonal terms are dominated by the shape noise which in the analytical model considers the realistic KiDS-1000 survey geometry, while in *Buceros* only matches the effective survey area without matching the survey geometry. This underlines the importance of accurately modelling the survey geometry in the analytical shape noise model.

When comparing the entirety of the covariance matrix from *Buceros* to the analytical model in Figure 4.18, we find that the discrepancy in the correlation coefficients is generally not as pronounced as it was when compared to *Cygnus* in Figure 4.16. This illustrates that the off-diagonal terms of the analytical covariance matrix appear to be in better agreement with a model which does not consider a realistic survey geometry. The small discrepancies observed are most likely caused by

Figure 4.18: Bitmap of the difference in the correlation coefficients, $\Delta\rho$, between the analytic covariance matrix (see Figure 4.8), and the covariance matrix from *Buceros* (idealised footprint , and homogeneous galaxy selection; shown in Figure 4.9). The upper left panels show $\Delta\rho$ for ξ_+- ξ_+, with each pixel showing the value for a single log-spaced bin in angular separation, θ, between $\theta = 0.1$ arcmin and $\theta = 300$ arcmin. The upper right and lower left panels show the same, but for the change in correlation between ξ_+ and ξ_-, while the lower right panels show the change in correlation of ξ_- with itself.

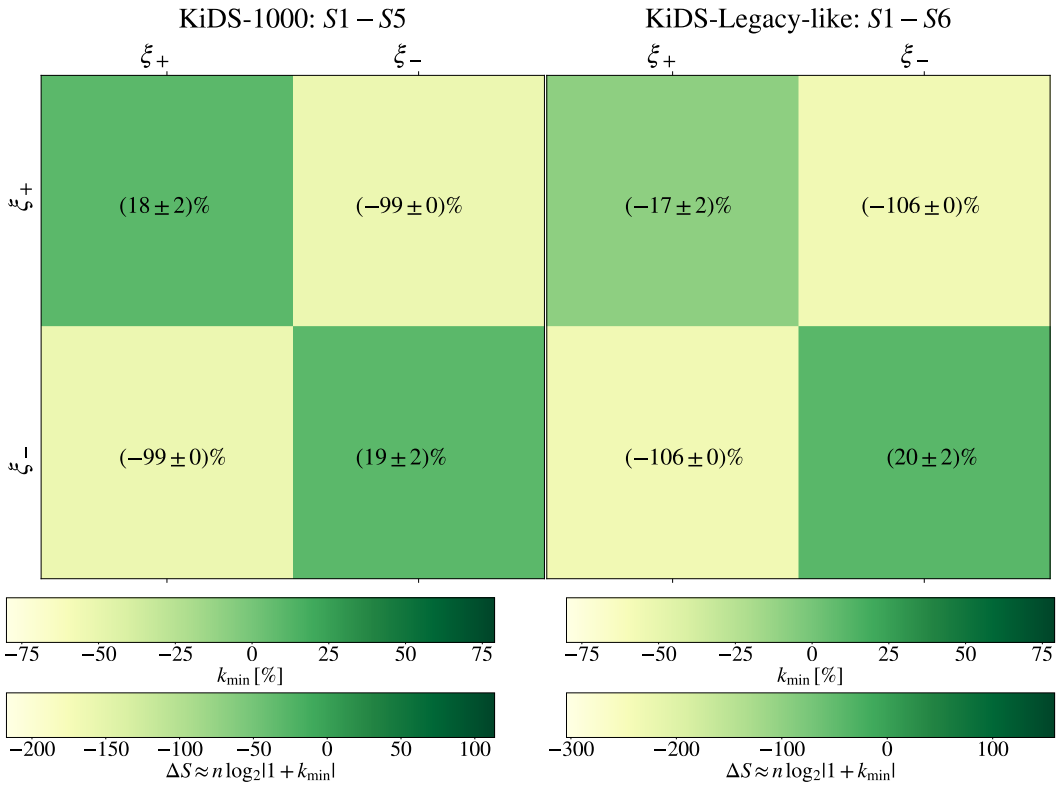

Figure 4.19: Bitmap showing the $k_{\min}$ measure from comparing the analytical covariance matrix, $\mathrm{Cov}(\xi_\pm, \xi_\pm)$, and the $\mathrm{Cov}(\xi_\pm, \xi_\pm)$ as calculated from *Buceros* as four separate blocks. The uncertainties given for each $k_{\min}$ are a measure of the potential degeneracy of $k_{\min}$ as given by $\Gamma/2$ as shown in Equation (4.19), where $\Gamma \to 0$ implies that $k_{\min}$ is unique and $\Gamma \to \infty$ implies that $k_{\min}$ is not a unique minimum. The left 4×4 bitmap shows $k_{\min}$ when only considering the covariance terms which correlate the uncertainties from tomographic bins S1 to S5 (as in KiDS-1000). The right 4×4 bitmap shows $k_{\min}$ when considering all covariance terms from tomographic bins S1 to S6 (as in KiDS-Legacy). Lighter panels imply that the given block of the analytical covariance matrix is overall less noisy than $\mathrm{Cov}_{\mathrm{Buceros}}$, darker green panels imply the reverse.

the fact that the analytical model considers the realistic KiDS-1000 survey footprint when computing the contributions due to the super-sample covariance.

With that said, Figure 4.19 shows that the values of $k_{\min}$ when comparing the analytical covariance matrix versus the one from *Buceros* are virtually the same as the $k_{\min}$ when comparing it against *Cygnus* (see Figure 4.17). Hence, the relative difference in entropy between the analytical and numerical model is unaffected by the consideration of the realistic footprint or lack thereof. The same is the case for the sensitivity of the constraints to the inclusion of the sixth tomographic bin.

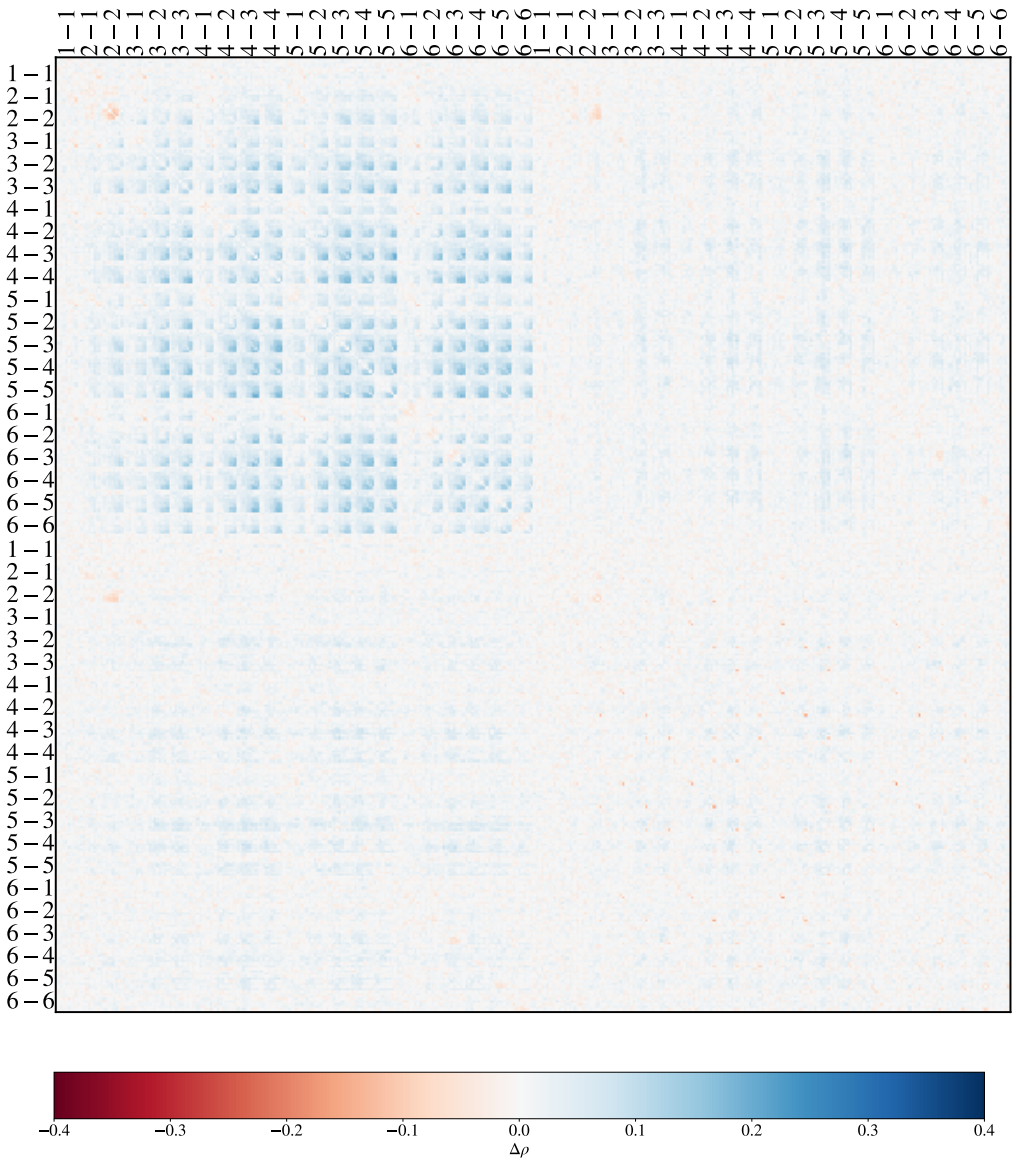

Figure 4.20: Bitmap of the difference in the correlation coefficients, $\Delta\rho$, between the analytic covariance matrix (see Figure 4.8), and the covariance matrix from *Egretta* (realistic footprint, and inhomogeneous and anistropic galaxy selection; shown in Figure 4.11). The upper left panels show $\Delta\rho$ for ξ_+- ξ_+, with each pixel showing the value for a single log-spaced bin in angular separation, θ, between $\theta = 0.1$ arcmin and $\theta = 300$ arcmin. The upper right and lower left panels show the same, but for the change in correlation between ξ_+ and ξ_-, while the lower right panels show the change in correlation of ξ_- with itself.

To finalise our assessment of the accuracy of the analytical covariance matrix, we compare it against the most general numerical model: *Egretta*. *Egretta* includes the realistic KiDS-1000 footprint (as is the case for some terms of the analytical covariance), but also takes into account variable depth. Even though both models consider the shape noise's dependence on the survey footprint, we observe a consistent bias of approximately 10% in the diagonal terms of the covariance matrix due to variable depth (Figure 4.15). This implies that the lack of variable depth in the noise model may similarly bias the shape noise as the lack of consideration for the survey geometry (see Section 4.5.4 for a detailed discussion of this). Nevertheless, this discrepancy is not substantial, while implying that the analytical model is more conservative. When comparing the entirety of the covariance matrix in Figure 4.20, we find that the change in correlation coefficients is slightly larger overall than the one observed between the analytic and *Cygnus* model. Hence, the analytical model still appears to slightly underestimate the noise contributions from the off-diagonals when compared to the *Egretta* model. This again seems to be driven by the analytic model's lack of consideration of the survey footprint in the cosmic variance terms.

Overall, we find that the analytic model for uncertainty on the cosmic shear signal from KiDS-1000 is broadly in agreement with the numerical uncertainty model, particularly, *Cygnus*. Even when removing the realistic survey geometry from or including variable depth in the numerical model, the discrepancies with the analytic model are consistently less than 15%, and insensitive to the inclusion of an additional KiDS-Legacy-like tomographic bin. Hence, even with the improvements of the forward simulations and the inclusion of an additional tomographic bin, these findings are consistent with previous tests of the analytic KiDS-1000 uncertainty model (Joachimi et al., 2021).

4.5.3 Impact of the Spatial Footprint

To test the impact of the assumption an idealised square footprint on the noise modelling when compared to a realistic footprint geometry, we compare the numerical covariances computed from the *Buceros* (based on an idealised square footprint) and *Cygnus* (based on the realistic KiDS-1000 footprint of the same area as in *Buceros*)

following the prescription described in Section 4.3.2. From one realisation to the next, the seed of the sampled matter field and galaxies tracing the overdensities is changed to account for cosmic variance. We use the same 5,000 seeds for both *Buceros* and *Cygnus*, so they both encode the same cosmic variance.

Figure 4.21 shows the difference between standard deviations from *Cygnus* and *Buceros*. This shows that the diagonal elements of the KiDS-Legacy cosmic shear covariance increase by up to around 20% when considering a realistic survey footprint. To be more precise, the covariance of ξ_- increases by about 5% at small scales and by up to 20 to 30% at large scales for most of the tomographic bin combination considered. This increase is consistent with what was found for KiDS-1000 (see Fig. D.2 in Joachimi et al. 2021). When looking at the covariance of ξ_+ instead, we find that, at small scales (within the bounds of the footprint dimensions), the standard deviation increases by about 10% throughout all tomographic bin combinations. However, after $\theta \sim 20$ arcmin, the excess in the standard deviation is consistently smaller than 10% (with the exception of the first tomographic bin). For many tomographic bin pairs, the variance in ξ_+ at scales of $\theta \sim 10^2$ arcmin is in fact smaller in *Cygnus* than in *Buceros*. Although the signal at these scales may be slightly biased due to the choice of *bin_slop* in TREECORR (see Section 4.4.1), this is again consistent with the findings in previous studies for KiDS-1000 (Joachimi et al., 2021). From this, we conclude that at large scales the effect of the survey footprint on the cosmic shear uncertainty is negligible, particularly, in the tomographic bin combinations with the highest signal-to-noise ratio. This is consistent with the fact that both *Buceros* and *Cygnus* probe large scales in a similar fashion as their survey areas are by construction the same. Hence, the difference in the survey geometry (the addition of holes and edges) only drives changes in the uncertainty in the cosmic shear signal at small to medium scales.

Therefore, from Figure 4.21, we can conclude that, despite the updates in the simulation pipeline highlighted in Section 4.2.1, the conclusions from KiDS-1000 about the impact of the footprint on the uncertainty modelling still hold. At the same time, we note that the changes in the standard deviation observed for tomographic

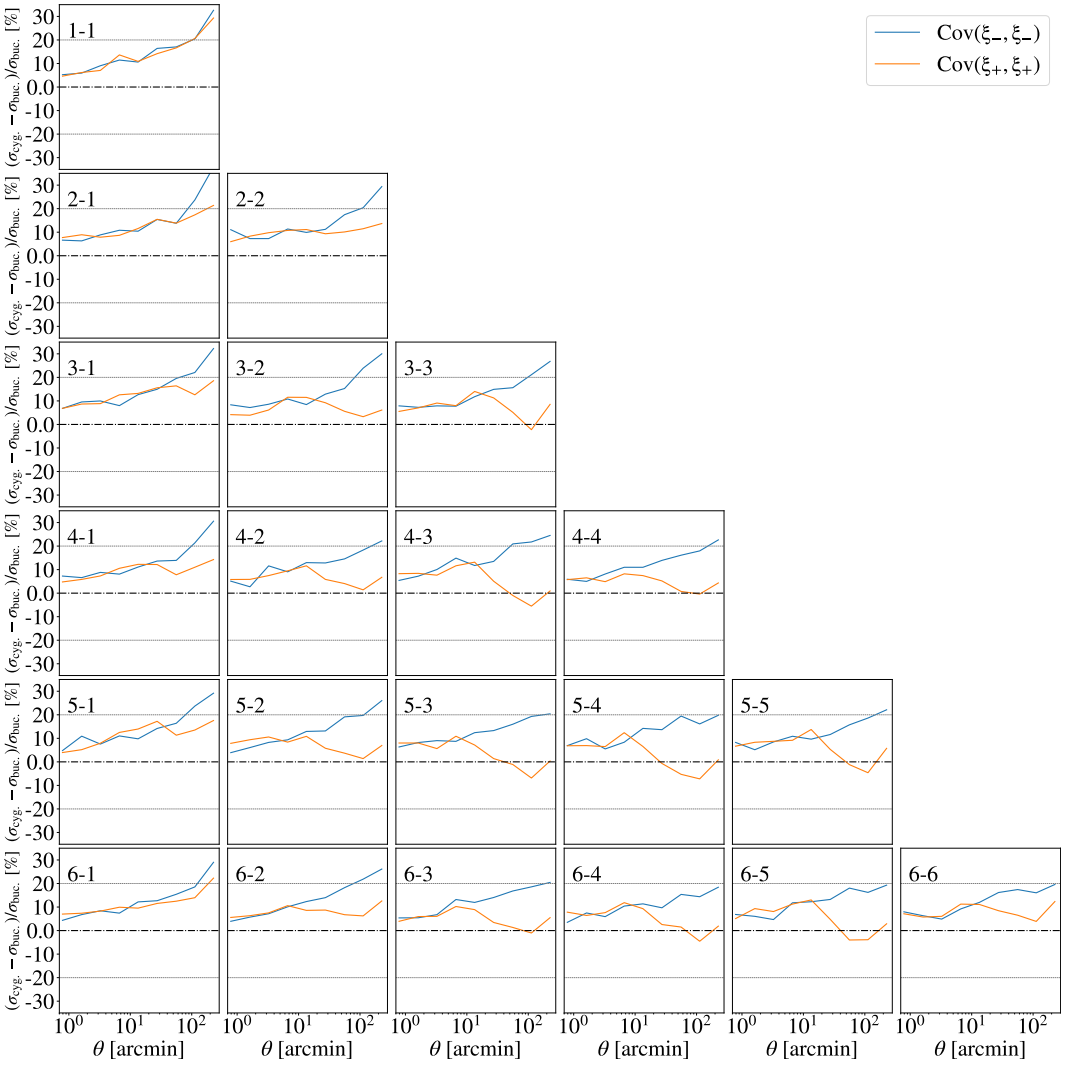

Figure 4.21: Relative difference between the standard deviations calculated from the diagonal elements of the covariance of two-point correlation functions, $\mathrm{Cov}[\xi_{\pm}(\theta), \xi_{\pm}(\theta)]$, as measured from 5,000 realisations of *Cygnus* and from 5,000 realisations of *Buceros*, as a function of angular separation, θ, in arcmin. The orange lines show the relative differences between $\mathrm{Cov}[\xi_{+}(\theta), \xi_{+}(\theta)]$, while the blue lines show the same for $\mathrm{Cov}[\xi_{-}(\theta), \xi_{-}(\theta)]$. Each panel shows the difference in signal for a given tomographic bin pair of the bins from S1 to S6.

bin combinations which include the KiDS-Legacy sixth bin do not significantly differ from the changes observed for any of the other tomographic bins. Nevertheless, as discussed in Section 4.5, the diagonal elements may not necessarily give the full picture.

To assess the impact of the choice of survey footprint on the covariance matrix as a whole, we calculate the difference between the covariance matrices from *Cygnus* and *Buceros* as shown in Figure 4.22. It becomes apparent that despite the terms in most diagonals increasing from *Buceros* to *Cygnus*, in many off-diagonal terms, this is not the case. Particularly, some of the off-diagonal terms which are typically dominated by the Gaussian mixed term or the super-sample covariance decrease due to the assumption of a realistic footprint. To evaluate the aggregate effect on the constraining power of the covariance based on all these terms changing in tandem, we utilize the $k_{\rm min}$ measure presented in Section 4.5.1.

Figure 4.23 shows the $k_{\rm min}(\mathrm{Cov}_{\rm Cygnus}, \mathrm{Cov}_{\rm Buceros})$ calculated individually for each 9×9 block matrix for a given pair of tomographic bins, (ij), within the covariance matrices, $\mathrm{Cov}[\xi_{\pm}^{(ij)}(\theta), \xi_{\pm}^{(ij)}(\theta)]$. We see that $k_{\rm min}[\mathrm{Cov}_{\rm Cygnus}(\xi_{+}^{(ij)}, \xi_{+}^{(ij)}), \mathrm{Cov}_{\rm Buceros}(\xi_{+}^{(ij)}, \xi_{+}^{(ij)})]$ indicates that the constraining power from $\xi_{+}^{(ij)}$ decreases substantially with a realistic footprint when $i, j > 3$, while when $i \leq 3$ and $j \leq 3$ simultaneously, the opposite is the case. This is a good example of how the changes in the diagonals of the covariance matrix may not be representative of the effect the whole covariance has in aggregate: from Figure 4.21, one would expect the constraining power to reduce similarly for all tomographic bin combinations, yet this is not the case when considering Figure 4.23. This becomes even more apparent when considering $k_{\rm min}[\mathrm{Cov}_{\rm Cygnus}(\xi_{-}^{(ij)}, \xi_{-}^{(ij)}), \mathrm{Cov}_{\rm Buceros}(\xi_{-}^{(ij)}, \xi_{-}^{(ij)})]$. Almost for all tomographic bin combinations, the constraining power of ξ_{-} improves with the adoption of a realistic footprint in *Cygnus*. A notable exception to this is the sixth tomographic bin which becomes significantly more noisy in $\xi_{-}^{(ij)}$ for $(ij) \in \{(63), (64), (65), (66)\}$. Although this is consistent with the noise increase seen in $\xi_{-}^{(55)}$, the uncertainty change in $\xi_{-}^{(63)}$, $\xi_{-}^{(64)}$ and $\xi_{-}^{(65)}$ is much larger than for other off-diagonal tomo-

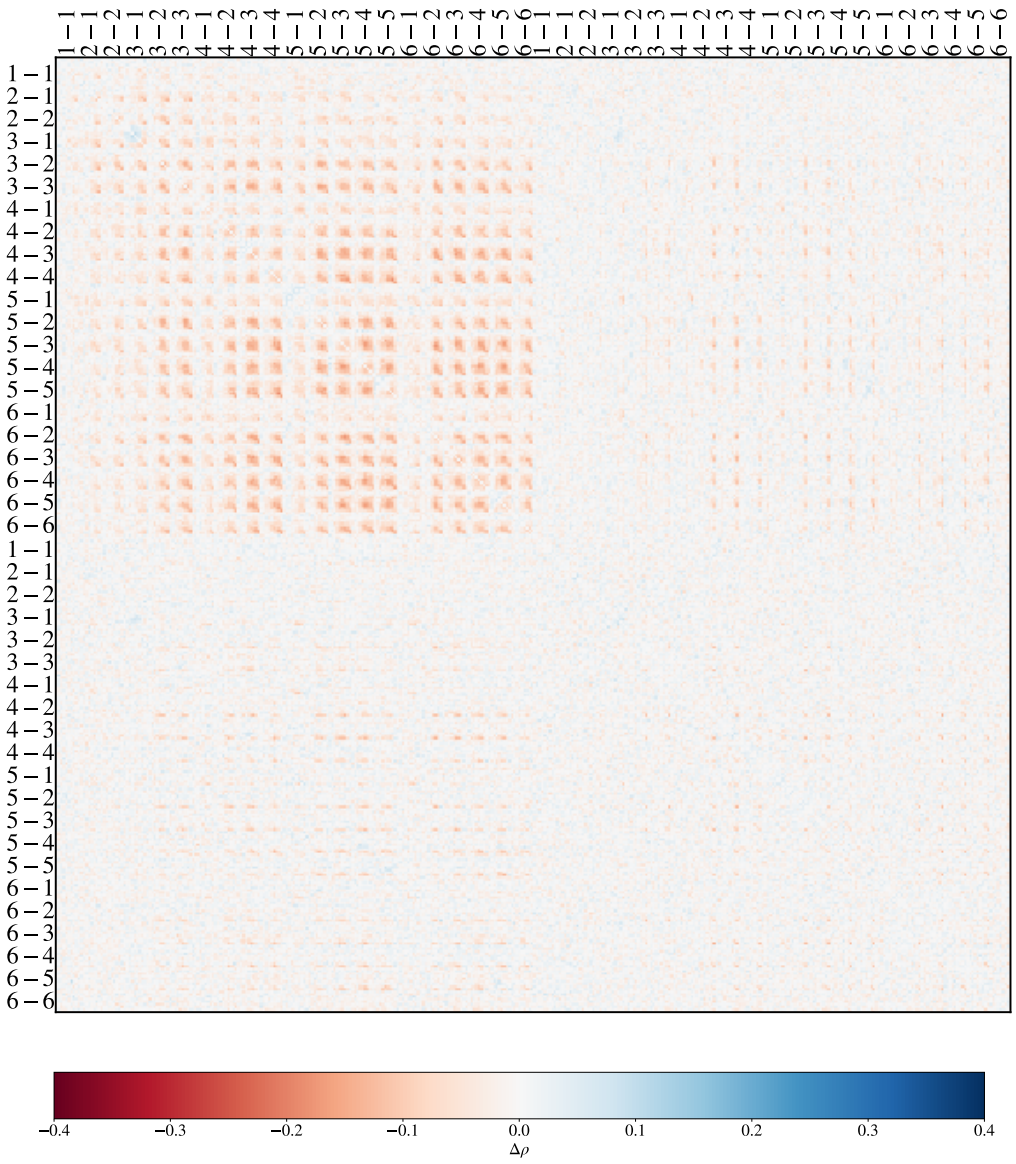

Figure 4.22: Bitmap of the difference in the correlation coefficients, $\Delta\rho$, between *Cygnus* (realistic footprint, and homogeneous galaxy selection; shown in Figure 4.10), and *Buceros* (idealised footprint , and homogeneous galaxy selection; shown in Figure 4.9). The upper left panels show $\Delta\rho$ for ξ_+- ξ_+, with each pixel showing the value for a single log-spaced bin in angular separation, θ, between $\theta = 0.1$ arcmin and $\theta = 300$ arcmin. The upper right and lower left panels show the same, but for the change in correlation between ξ_+ and ξ_-, while the lower right panels show the change in correlation of ξ_- with itself.

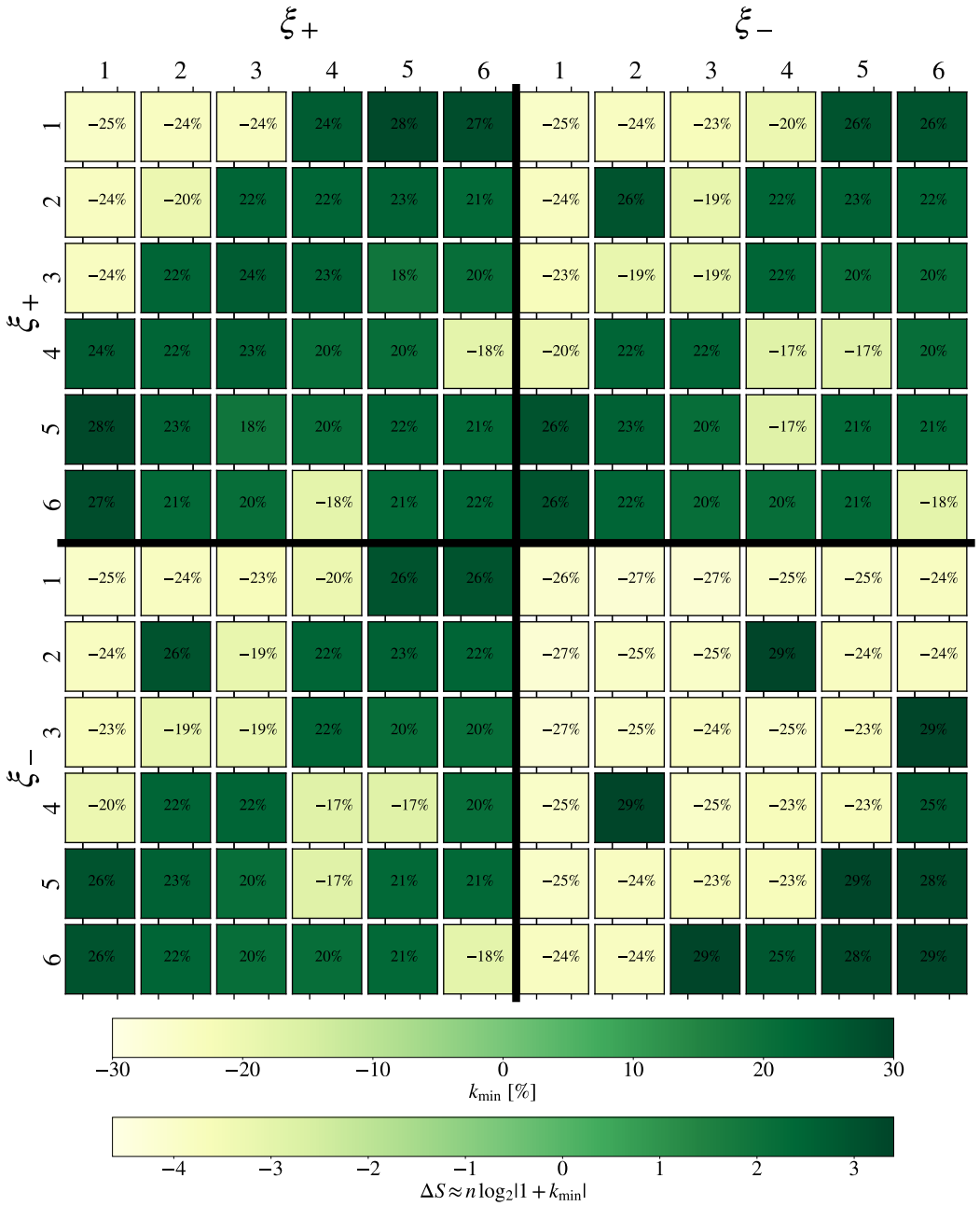

Figure 4.23: Bitmap showing the $k_{\min}$ measure between the diagonal blocks of the covariance matrix calculated from the *Cygnus* simulations, $\mathrm{Cov}_{\mathrm{Cygnus}}$ (see Figure 4.10), and the covariance matrix calculated from the *Buceros* simulations, $\mathrm{Cov}_{\mathrm{Buceros}}$ (see Figure 4.9). Lighter panels imply that the given block of $\mathrm{Cov}_{\mathrm{Cygnus}}$ is overall less noisy than $\mathrm{Cov}_{\mathrm{Buceros}}$, darker green panels imply the reverse. The second colourbar shows the difference in the entropy, ΔS, associated with a given $k_{\min}$ for $n = 9$ (i.e. each block matrix in a given panel being made up of 9×9 elements). Each panel gives $k_{\min}$ for a given tomographic bin combination of the bins from S1 to S6.

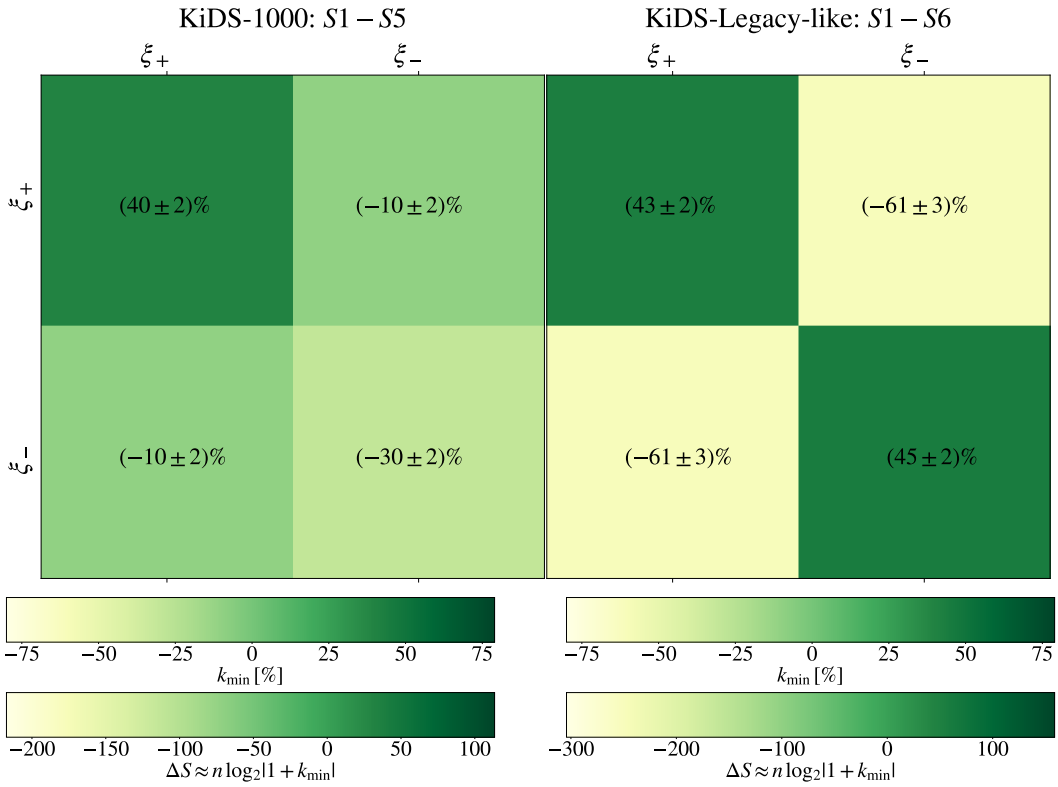

Figure 4.24: Bitmap showing the $k_{\min}$ measure from comparing the $\mathrm{Cov}(\boldsymbol{\xi}_{\pm}, \boldsymbol{\xi}_{\pm})$ as calculated from *Cygnus* and the $\mathrm{Cov}(\boldsymbol{\xi}_{\pm}, \boldsymbol{\xi}_{\pm})$ as calculated from *Buceros* as four separate blocks. The uncertainties given for each $k_{\min}$ are a measure of the potential degeneracy of $k_{\min}$ as given by $\Gamma/2$ as shown in Equation (4.19), where $\Gamma \to 0$ implies that $k_{\min}$ is unique and $\Gamma \to \infty$ implies that $k_{\min}$ is not a unique minimum. The fits from which Γ is estimated are shown in Figure 4.14. The left 4×4 bitmap shows $k_{\min}$ when only considering the covariance terms which correlate the uncertainties from tomographic bins S1 to S5 (as in KiDS-1000). The right 4×4 bitmap shows $k_{\min}$ when considering all covariance terms from tomographic bins S1 to S6 (as in KiDS-Legacy). Lighter panels imply that the given block of $\mathrm{Cov}_{\mathrm{Cygnus}}$ is overall less noisy than $\mathrm{Cov}_{\mathrm{Buceros}}$, darker green panels imply the reverse.

graphic bin combinations in ξ_-. This points at the fact that the combination of a broad redshift distribution in S6 combined with a relatively low galaxy density and high shape dispersion may cause a more substantial response in the uncertainty of $\xi_-^{(6\,j)}$. Having said this, the overall loss of constraining power in $\xi_-^{(6\,j)}$ for $j > 2$ is not significantly larger than the increase observed in the covariance of ξ_+.

To quantify more precisely how the sensitivity to the survey footprint geometry is altered when including S6, we calculate $k_{\rm min}$ between the covariances from *Cygnus* and from *Buceros*, once for a KiDS-1000 set of covariances which only include bins S1 to S5, and once again for a KiDS-Legacy-like set of covariances including six tomographic bins. Figure 4.24 shows that the $k_{\rm min}(\mathrm{Cov}_{\rm Cygnus}(\xi_+,\xi_+),\mathrm{Cov}_{\rm Buceros}(\xi_+,\xi_+))$ goes from 40% to 43% when including S6 in the uncertainty model. This points to the sensitivity to the survey footprint in the uncertainty of ξ_+ not substantially increasing in KiDS-Legacy. This is less so the case for the $k_{\rm min}$ of $\mathrm{Cov}(\xi_-,\xi_-)$ and $\mathrm{Cov}(\xi_+,\xi_-)$. Its sensitivity to the footprint appears to be significantly altered by the inclusion of the sixth bin. However, as can be seen from Figure 4.22, for these terms, the changes in the correlation coefficients are not significant. Thus, the substantial changes in $k_{\rm min}$ are caused by large relative changes in the covariance matrix terms which are insignificant in absolute terms, as the values of most terms in question are near zero.

4.5.4 Impact of Spatial Variability

To test the impact of spatial variability of the galaxy selection or variable depth on the uncertainty model, we repeat the tests discussed in Section 4.5.3, but instead compare the covariance matrix obtained from the realisation of *Egretta* to the other simulation suites (*Cygnus* and *Buceros*). *Egretta* not only takes into consideration a realistic survey footprint geometry based on KiDS-1000 like *Cygnus*, it also considers that within this footprint the selection function of galaxies is inhomogenous and anisotropic, i.e. the sensitivity to galaxies varies as a function of angular position on the sky. To isolate this effect on its own, we first compare *Egretta* to *Cygnus* as is shown in Figures 4.27, 4.28 and 4.29. The first thing to note here is the counter-intuitive apparent reduction in the diagonal elements and many of the

off-diagonal elements of the covariance when including variable depth. A priori, one might assume that the additional spatial variation in the galaxy density and in the shape dispersion would increase the covariance terms. Upon closer inspection, one realises that this need not necessarily be the case.

In Figure 4.28, we see that variable depth does not seem to have a significant impact on the covariance matrix, even including S6. Nevertheless, the diagonal terms shown in Figure 4.27 are suppressed throughout due variable depth. These terms are typically dominated by shape noise which is precisely why we observe a reduction in these terms due to variable depth. Figure 4.25 shows the distribution of the parameter used to calibrate variable depth across the KiDS footprint, $\sigma_{\rm rms}$, based on the map shown in Figure 3.8. From Figure 4.25, we see that the distribution of $\sigma_{\rm rms}$ is heavily skewed with a tail going into the high $\sigma_{\rm rms}$ values. Consequently, the median $\sigma_{\rm rms}$ is below the mean $\sigma_{\rm rms}$ by approx. 7×10^{-4}. This skew implies that, despite the pixels in the sky in *Egretta* having the same mean value of $\sigma_{\rm rms}$ as the patches in *Cygnus*, there are considerably more pixels which have low values of $\sigma_{\rm rms}$. From Figures 4.5 and 4.6, it becomes apparent that the majority of pixels in the footprint have a shape dispersion, σ_ε, below average, a galaxy density, $n_{\rm gal}$, above average, and a redshift distribution with a mean above average. Considering that the covariance due to shape noise scales to first order approximately with $\sigma_\varepsilon^2/n_{\rm gal} \propto \sigma_{\rm rms}^3$ (Kaiser, 1992, 1998), one can estimate the approximate reduction in the shape noise due to a shift of -7×10^{-4} in $\sigma_{\rm rms}$ as shown in Figure 4.26. These estimates imply a shift of up to -7% in the standard deviation across all scales. Additionally, the covariance terms do not vary linearly with $n_{\rm gal}$ or $\sigma_{\rm rms}$ which induces higher-order contributions that may reduce the covariance terms further. We remark that this is an inherent feature of the KiDS data, as the selection presented here (including the values shown Figures 3.8 and 4.25) is entirely calibrated from the KiDS DR4 data.

Again, Figure 4.29 gives a slightly different picture. It appears that, despite the decrease in many covariance terms due to variable depth, the constraining power is decreasing of the blocks in the covariance matrix along the diagonals, with

$k_{\rm min}[{\rm Cov}_{\rm Egretta}(\xi_{\pm}^{(ij)},\xi_{\pm}^{(ij)}),{\rm Cov}_{\rm Cygnus}(\xi_{\pm}^{(ij)},\xi_{\pm}^{(ij)})]$ increased for most tomographic bin combinations (with only a few exceptions). Thus, this indicates that the additional spatial variations in the galaxy density, shape dispersion and redshift distributions due to variable depth end up increasing the overall uncertainties on the cosmic shear two-point correlation function measurements.

Importantly, we note that the uncertainty on the measurements which include the sixth tomographic bin in KiDS-Legacy, S6, does not appear to be disproportionately affected by variable depth. The terms in ${\rm Cov}(\xi_{+}^{(6j)},\xi_{+}^{(6j)})$ seems to be similarly affected by variable depth as the terms in ${\rm Cov}(\xi_{+}^{(5j)},\xi_{+}^{(5j)})$. At same time, there appears to be a trade-off between ${\rm Cov}(\xi_{-}^{(6j)},\xi_{-}^{(6j)})$ being more noisy than in *Egretta*. The inverse appears to be the case for tomographic bin combinations with the fifth bin.

Similarly, as for the survey footprint, in Figure 4.30, we find that $k_{\rm min}({\rm Cov}_{\rm Egretta}(\xi_{+},\xi_{+}),{\rm Cov}_{\rm Cygnus}(\xi_{+},\xi_{+}))$ goes from -21% to -23% when including S6 in the uncertainty model. Similarly, for ${\rm Cov}(\xi_{-},\xi_{-}))$ terms, $k_{\rm min}$ changes from 27% to 31%. This points to the sensitivity to the depth variability in the uncertainty of ξ_{+} and ξ_{-} not varying substantially when including the sixth tomographic bin. The variation in the other $k_{\rm min}$ terms changes more substantially, but from Figure 4.28, this can be attributed to large relative variations in the covariance matrix blocks where the values of most terms are near zero, namely, ${\rm Cov}(\xi_{+},\xi_{-}))$ and ${\rm Cov}(\xi_{-},\xi_{+}))$, as shown in Figures 4.10 and 4.11.

Finally, we compare the covariances obtained from *Egretta* and *Buceros* directly in order to determine whether compounding the effects of a realistic footprint and variable depth at the same time affects the uncertainty model differently. Figures 4.31, 4.32 and 4.33 show the results of this comparison using the previously discussed metrics.

Figures 4.31 and 4.32 align well with the previous comparisons of *Cygnus* against *Buceros*, and *Egretta* against *Cygnus*. The differences between *Egretta* and *Buceros* are dominated by the contributions due to the footprint geometry and match well the changes seen in Section 4.5.3. Hence, when comparing *Egretta* and

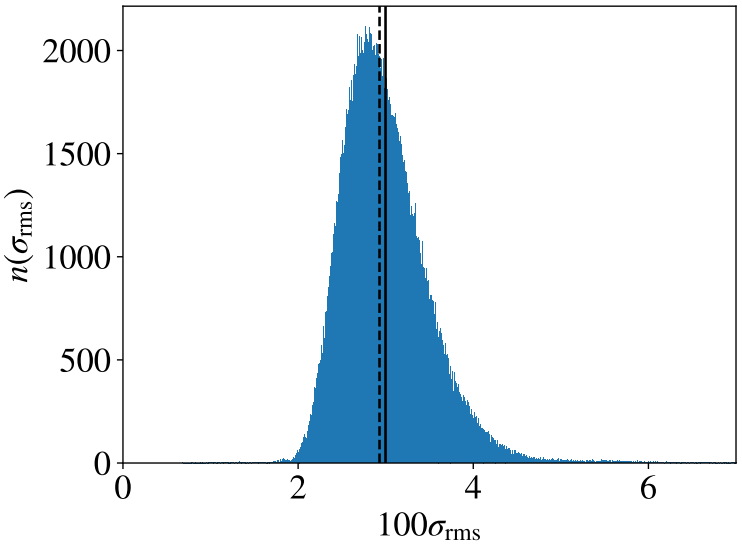

Figure 4.25: Histogram of the number of pixels on the KiDS-1000 footprint (see Figure 3.8) as a function of the root-mean square of the background noise, $\sigma_{\rm rms}$. The solid vertical line marks the mean $\sigma_{\rm rms}$ of the distribution, while the dashed vertical line marks the median $\sigma_{\rm rms}$.

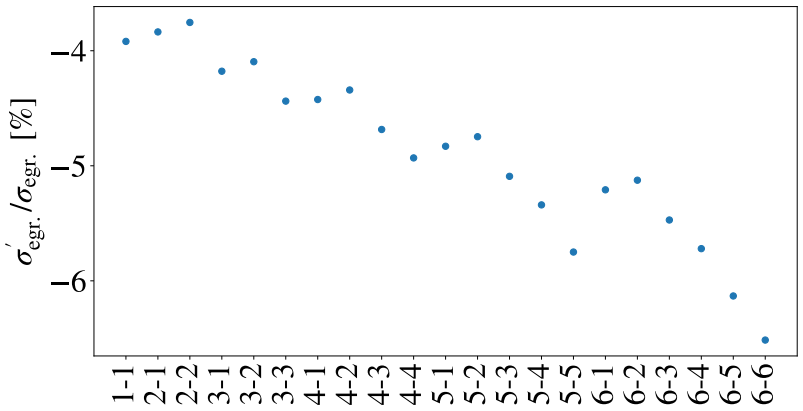

Figure 4.26: Plot of the ratio of $\sigma'_{\rm egr.}$, the approximate shape noise contribution in *Egretta* after a shift of -7×10^{-4} in root-mean square of the background noise, $\sigma_{\rm rms}$, over $\sigma_{\rm egr.}$, the approximate shape noise contribution in *Egretta* before the shift, as a function of the tomographic bin pair. The bin pairs are all 21 unique combinations of the six KiDS-Legacy tomographic bins.

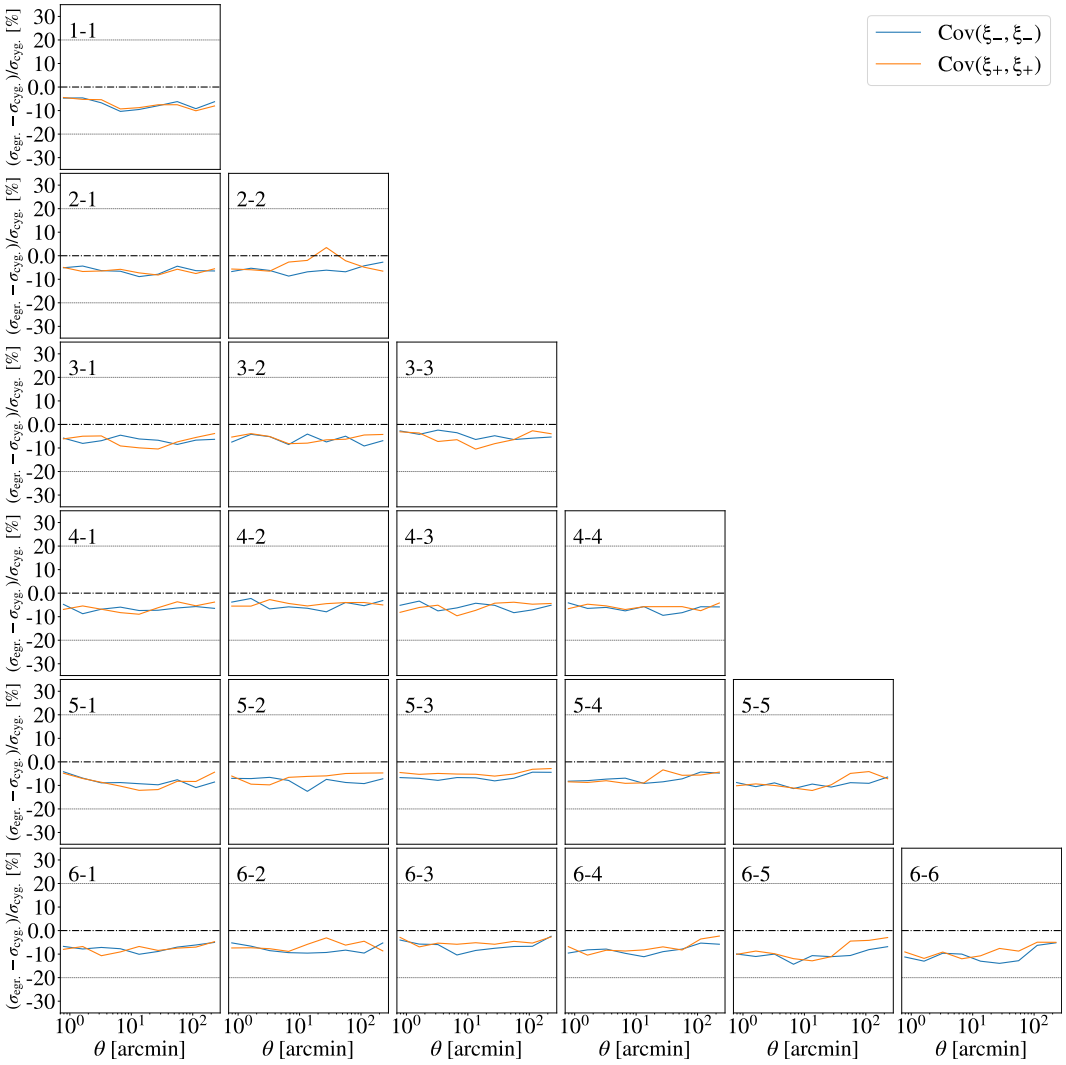

Figure 4.27: Relative difference between the standard deviations calculated from the diagonal elements of the covariance of two-point correlation functions, $\mathrm{Cov}[\xi_{\pm}(\theta), \xi_{\pm}(\theta)]$, as measured from 5,000 realisations of *Egretta* and from 5,000 realisations of *Cygnus*, as a function of angular separation, θ, in arcmin. The orange lines show the relative differences between $\mathrm{Cov}[\xi_{+}(\theta), \xi_{+}(\theta)]$, while the blue lines show the same for $\mathrm{Cov}[\xi_{-}(\theta), \xi_{-}(\theta)]$. Each panel shows the difference in signal for a given tomographic bin pair of the bins from S1 to S6.

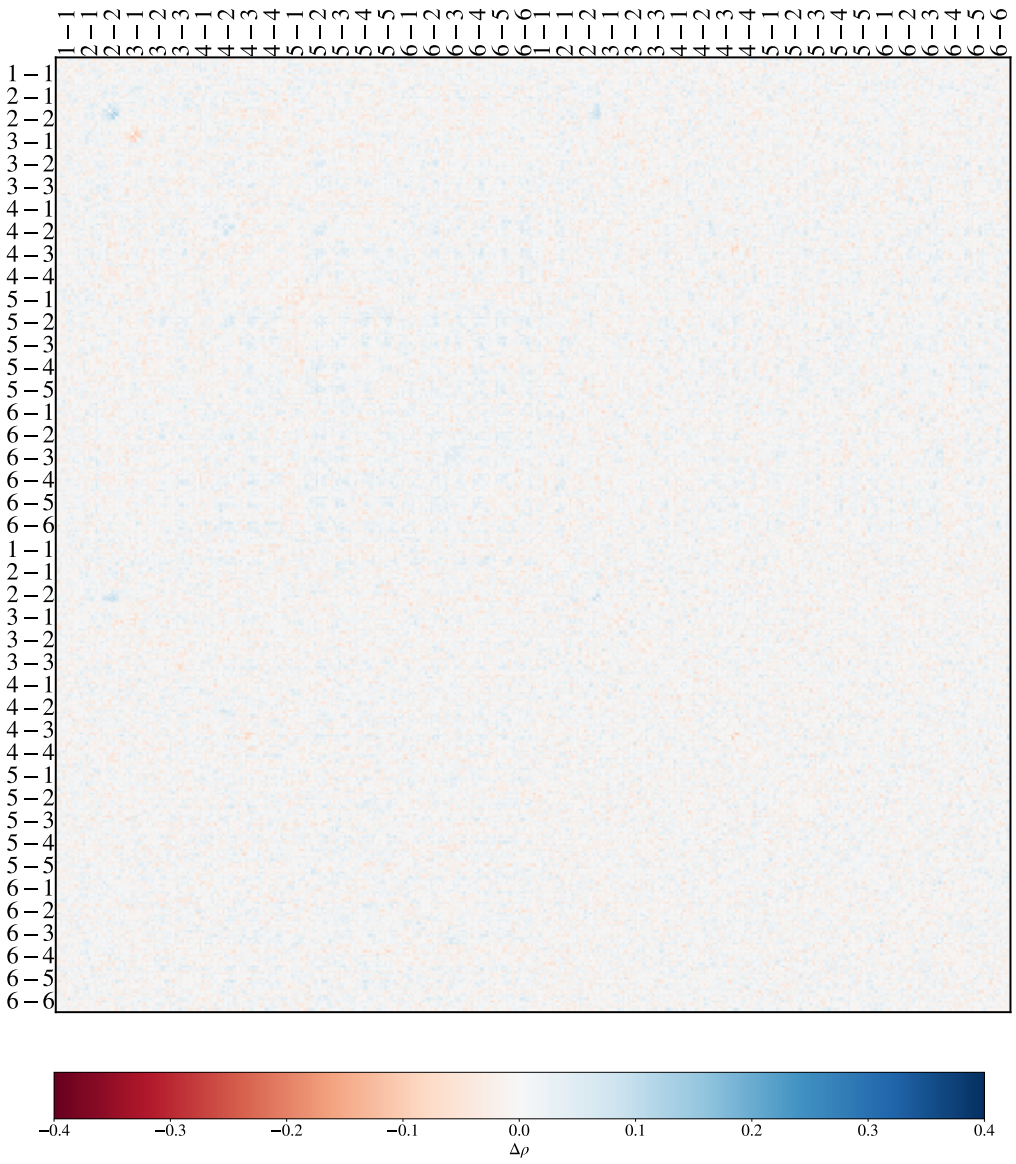

Figure 4.28: Bitmap of the difference in the correlation coefficients, $\Delta\rho$, between *Egretta* (realistic footprint, and inhomogeneous and anisotropic galaxy selection; shown in Figure 4.11), and *Cygnus* (realistic footprint , and homogeneous galaxy selection; shown in Figure 4.10). The upper left panels show $\Delta\rho$ for ξ_+- ξ_+, with each pixel showing the value for a single log-spaced bin in angular separation, θ, between $\theta = 0.1$ arcmin and $\theta = 300$ arcmin. The upper right and lower left panels show the same, but for the correlation between ξ_+ and ξ_-, while the lower right panels show the covariance of ξ_- with itself.

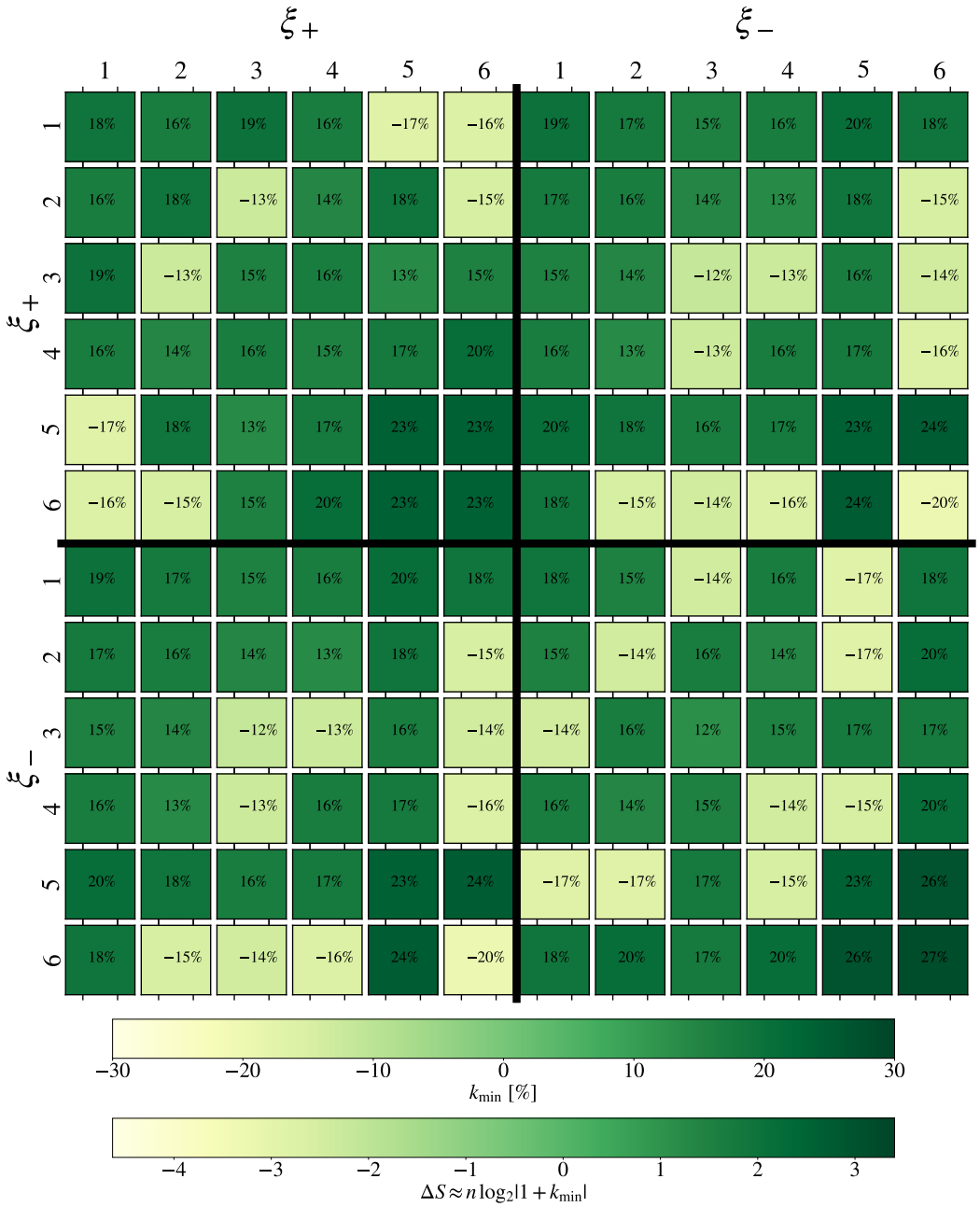

Figure 4.29: Bitmap showing the $k_{\min}$ measure between the diagonal blocks of the covariance matrix calculated from the *Egretta* simulations, $\mathrm{Cov}_{\mathrm{Egretta}}$ (see Figure 4.11), and the covariance matrix calculated from the *Cygnus* simulations, $\mathrm{Cov}_{\mathrm{Cygnus}}$ (see Figure 4.10). Lighter panels imply that the given block of $\mathrm{Cov}_{\mathrm{Egretta}}$ is overall less noisy than $\mathrm{Cov}_{\mathrm{Cygnus}}$, darker green panels imply the reverse. The second colourbar shows the difference in the entropy, ΔS, associated with a given $k_{\min}$ for $n = 9$ (i.e. each block matrix in a given panel being made up of 9×9 elements). Each panel gives $k_{\min}$ for a given tomographic bin combination of the bins from S1 to S6.

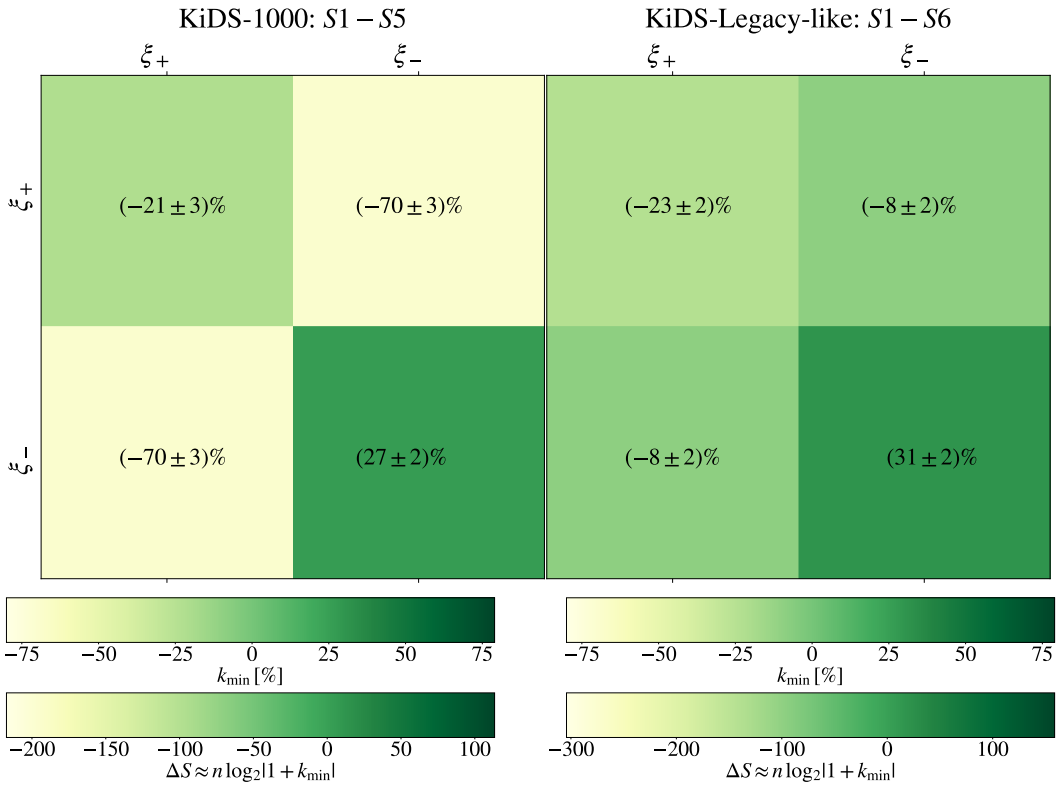

Figure 4.30: Bitmap showing the $k_{\min}$ measure from comparing the $\mathrm{Cov}(\xi_\pm, \xi_\pm)$ as calculated from *Egretta* and the $\mathrm{Cov}(\xi_\pm, \xi_\pm)$ as calculated from *Cygnus* as four separate blocks. The uncertainties given for each $k_{\min}$ are a measure of the potential degeneracy of $k_{\min}$ as given by $\Gamma/2$ as shown in Equation (4.19), where $\Gamma \to 0$ implies that $k_{\min}$ is unique and $\Gamma \to \infty$ implies that $k_{\min}$ is not a unique minimum. The left 4×4 bitmap shows $k_{\min}$ when only considering the covariance terms which correlate the uncertainties from tomographic bins S1 to S5 (as in KiDS-1000). The right 4×4 bitmap shows $k_{\min}$ when considering all covariance terms from tomographic bins S1 to S6 (as in KiDS-Legacy). Lighter panels imply that the given block of $\mathrm{Cov}_{\mathrm{Egretta}}$ is overall less noisy than $\mathrm{Cov}_{\mathrm{Cygnus}}$, darker green panels imply the reverse.

Buceros, it is unsurprising to find that the correlation coefficients in the covariance have mostly increased as expected from having assumed a realistic footprint. All this is consistent with tests done with mocks for KiDS-1000 (Joachimi et al., 2021).

Figure 4.33 is consistent with this picture, and shows that the contributions to the uncertainty due to variable depth and due to the survey geometry cancel each other out to an extent that the constraining power of the covariance matrix from *Buceros* and *Egretta* is relatively similar (at least when compared to *Cygnus*). Note that this does not imply that these two matrices ought to have similar constraints

on cosmology, as despite the change in total information being similar, the degeneracies in the likelihood of the cosmological parameters may still be considerably affected.

Notably, when comparing the covariance matrices in their entirety in Figure 4.34, we find that the inclusion of the sixth tomographic bin does not significantly affect the conjoint sensitivity to both the survey footprint geometry and variable depth at the same time.

Lastly, in this instance, the sixth tomographic bin from KiDS-Legacy again does not stand out from the other tomographic bins. Therefore, based on the expected redshift distribution, galaxy density and shape dispersion for S6, we do not expect that its inclusion in KiDS-Legacy would significantly alter the sensitivity to the assumptions in the analytic uncertainty model with respect to the measurements from KiDS-1000.

4.6 Conclusions

In this chapter, we presented a novel pipeline of forward simulations of cosmic shear two-point statistics focused on testing the assumptions made in the signal and uncertainty modelling for the upcoming KiDS data release, KiDS DR5 or KiDS-Legacy. This pipeline makes some key improvements over previous works which improve the overall accuracy, while including an additional sixth tomographic bin to extend the depth along the line-of-sight further than in KiDS-1000. Moreover, we show that the calibration of the model for variable depth used in KiDS-1000 is not solely dependent on survey characteristics and can bias cosmology. Alternatively, we suggest, implement and test a new unbiased calibration procedure to model variable depth.

From these forward simulations, we confirm that the two-point correlation functions are unbiased estimators of the cosmic shear signal when varying the geometry of the survey footprint. However, we note that this is not always the case, as numerical uncertainties may become relevant at large scales if the numerical precision is inadequate. We also determine that the bias in the cosmic shear signal

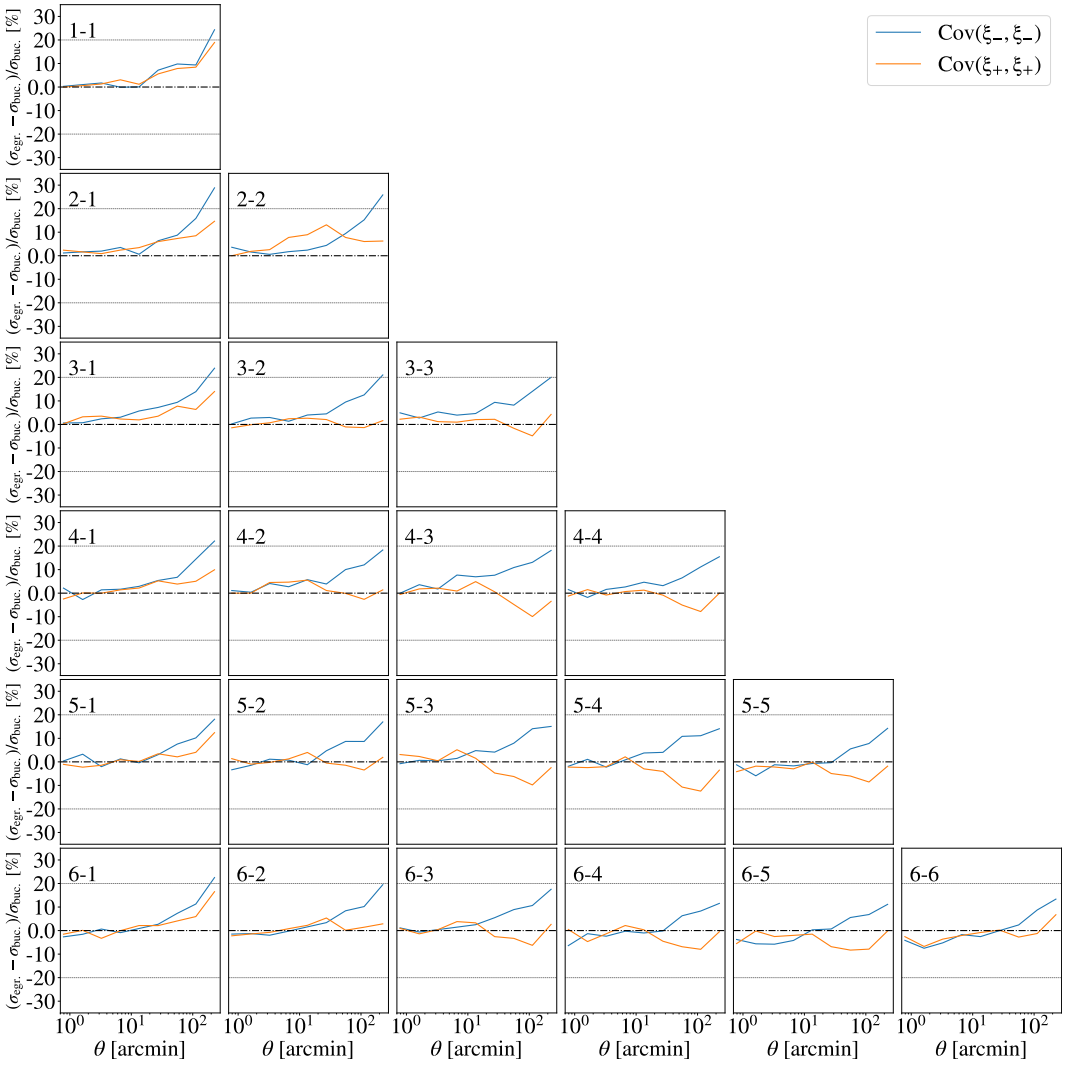

Figure 4.31: Relative difference between the standard deviations calculated from the diagonal elements of the covariance of two-point correlation functions, $\mathrm{Cov}[\xi_{\pm}(\theta), \xi_{\pm}(\theta)]$, as measured from 5,000 realisations of *Egretta* and from 5,000 realisations of *Buceros*, as a function of angular separation, θ, in arcmin. The orange lines show the relative differences between $\mathrm{Cov}[\xi_{+}(\theta), \xi_{+}(\theta)]$, while the blue lines show the same for $\mathrm{Cov}[\xi_{-}(\theta), \xi_{-}(\theta)]$. Each panel shows the difference in signal for a given tomographic bin pair of the bins from S1 to S6.

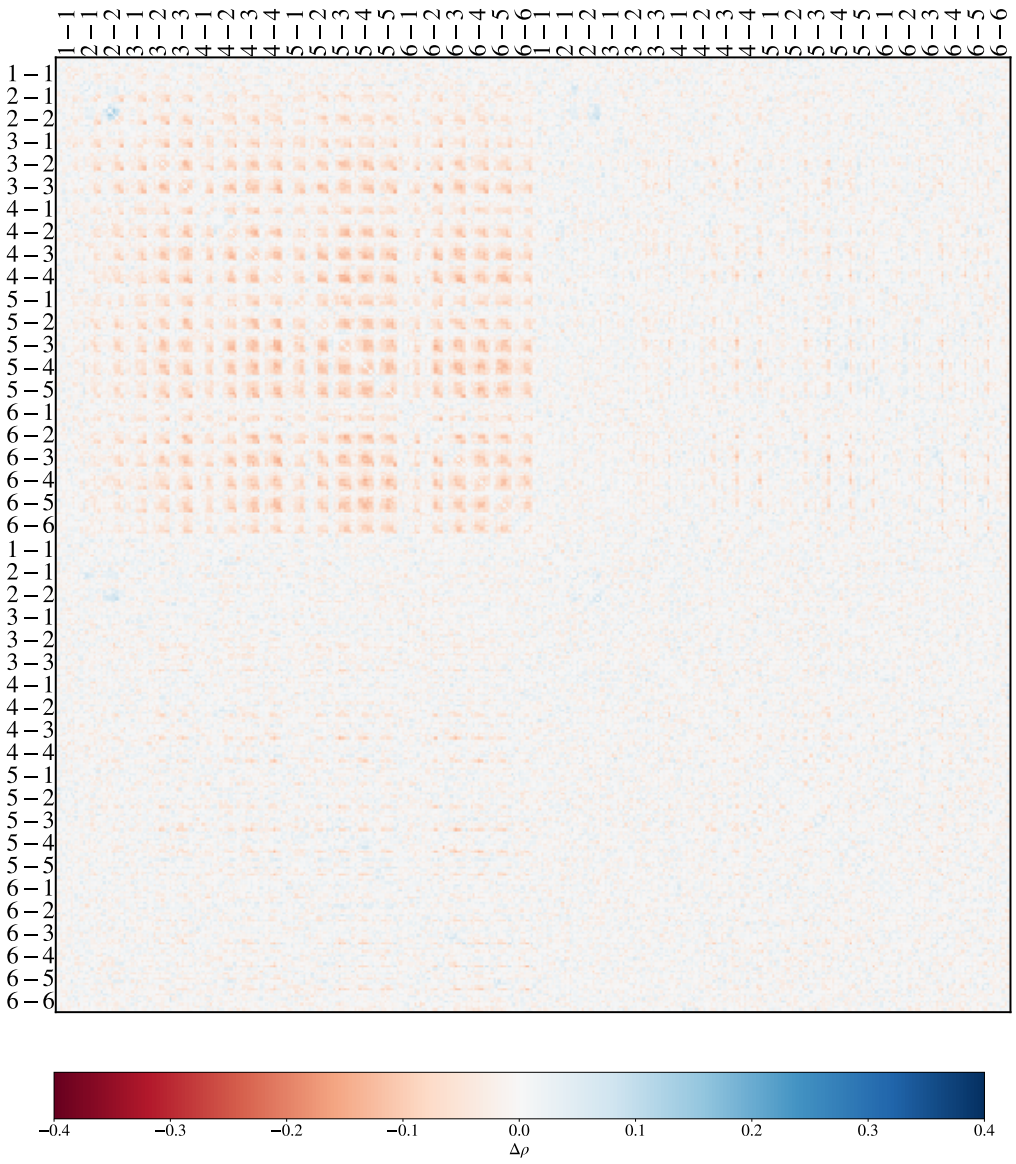

Figure 4.32: Bitmap of the difference in the correlation coefficients, $\Delta\rho$, between *Egretta* (realistic footprint, and inhomogeneous and anisotropic galaxy selection ; shown in Figure 4.11), and *Buceros* (idealised footprint , and homogeneous galaxy selection; shown in Figure 4.9). The upper left panels show $\Delta\rho$ for ξ_+-ξ_+, with each pixel showing the value for a single log-spaced bin in angular separation, θ, between $\theta = 0.1$ arcmin and $\theta = 300$ arcmin. The upper right and lower left panels show the same, but for the correlation between ξ_+ and ξ_-, while the lower right panels show the covariance of ξ_- with itself.

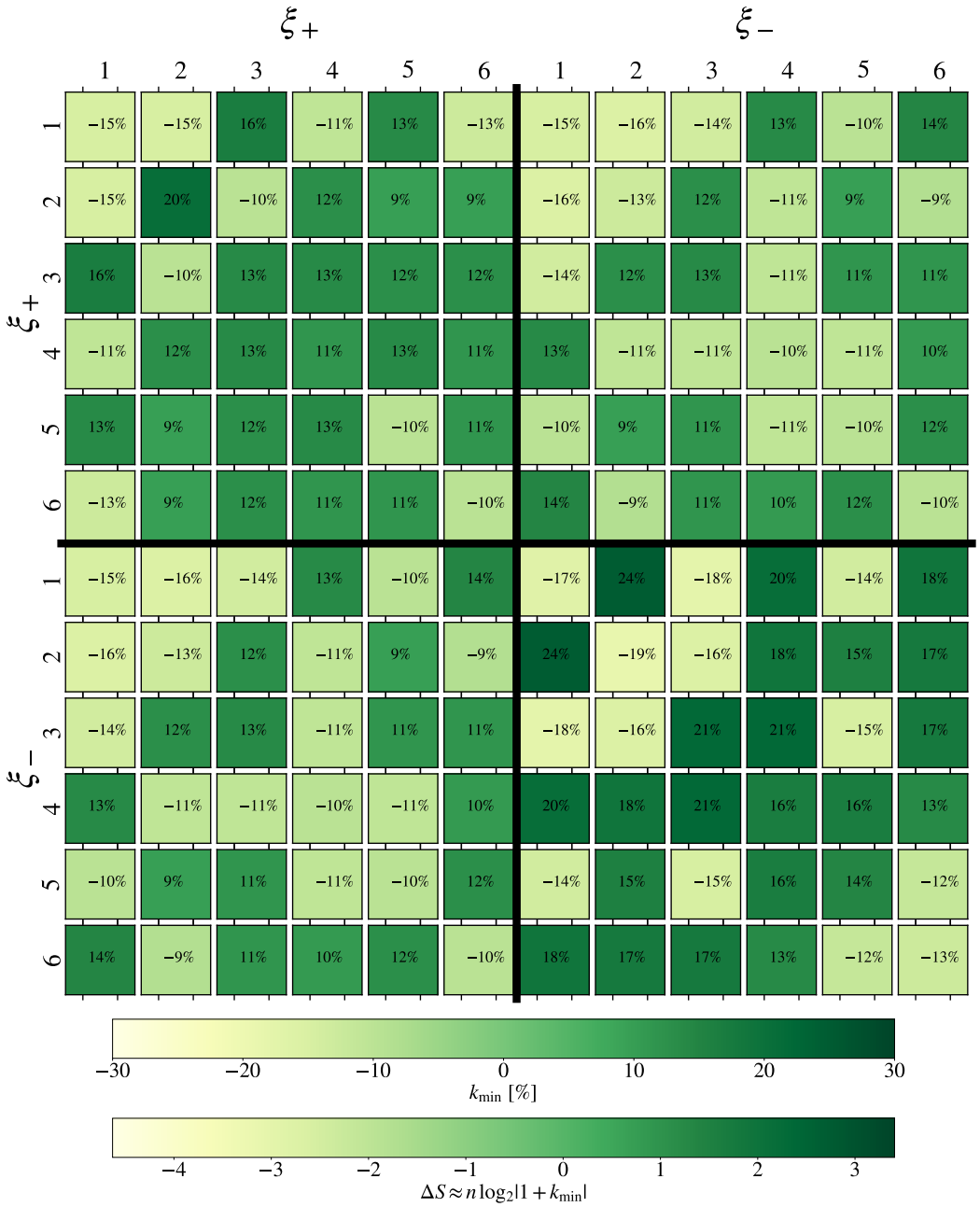

Figure 4.33: Bitmap showing the $k_{\min}$ measure between the diagonal blocks of the covariance matrix calculated from the *Egretta* simulations, $\mathrm{Cov}_{\mathrm{Egretta}}$ (see Figure 4.11), and the covariance matrix calculated from the *Buceros* simulations, $\mathrm{Cov}_{\mathrm{Buceros}}$ (see Figure 4.9). Lighter panels imply that the given block of $\mathrm{Cov}_{\mathrm{Egretta}}$ is overall less noisy than $\mathrm{Cov}_{\mathrm{Buceros}}$, darker green panels imply the reverse. The second colourbar shows the difference in the entropy, ΔS, associated with a given $k_{\min}$ for $n = 9$ (i.e. each block matrix in a given panel being made up of 9×9 elements). Each panel gives $k_{\min}$ for a given tomographic bin combination of the bins from S1 to S6.

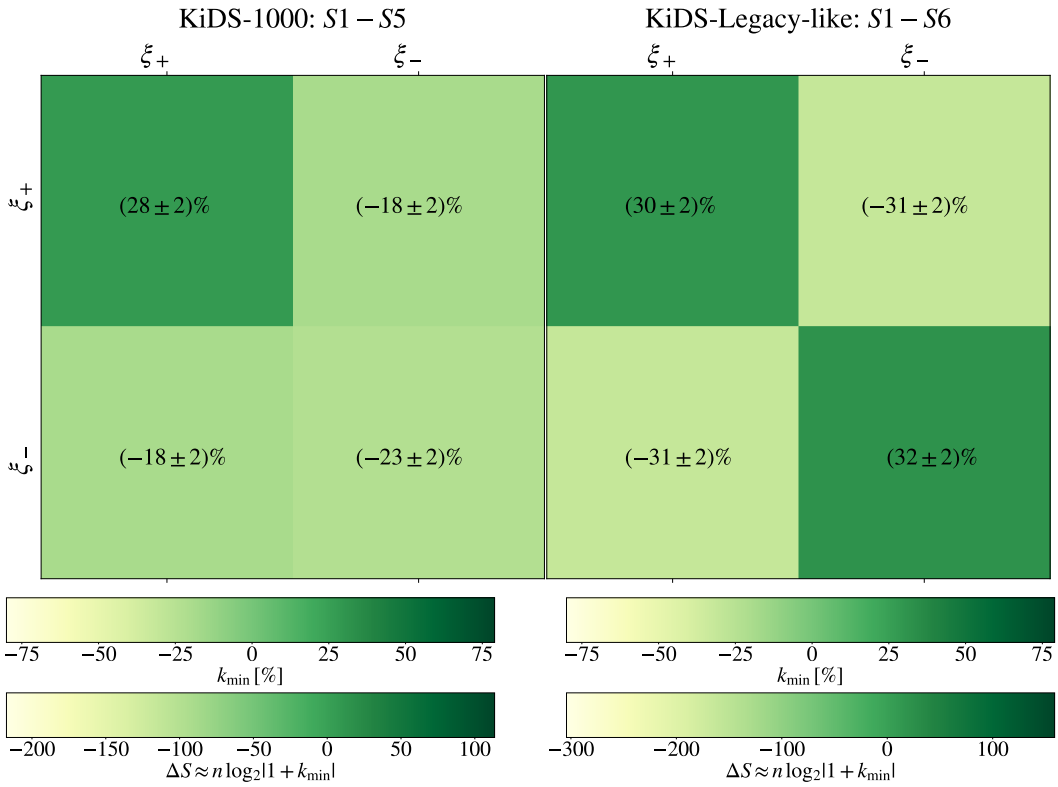

Figure 4.34: Bitmap showing the $k_{\min}$ measure from comparing the $\mathrm{Cov}(\xi_\pm, \xi_\pm)$ as calculated from *Egretta* and the $\mathrm{Cov}(\xi_\pm, \xi_\pm)$ as calculated from *Buceros* as four separate blocks. The uncertainties given for each $k_{\min}$ are a measure of the potential degeneracy of $k_{\min}$ as given by $\Gamma/2$ as shown in Equation (4.19), where $\Gamma \to 0$ implies that $k_{\min}$ is unique and $\Gamma \to \infty$ implies that $k_{\min}$ is not a unique minimum. The left 4×4 bitmap shows $k_{\min}$ when only considering the covariance terms which correlate the uncertainties from tomographic bins S1 to S5 (as in KiDS-1000). The right 4×4 bitmap shows $k_{\min}$ when considering all covariance terms from tomographic bins S1 to S6 (as in KiDS-Legacy). Lighter panels imply that the given block of $\mathrm{Cov}_{\mathrm{Egretta}}$ is overall less noisy than $\mathrm{Cov}_{\mathrm{Buceros}}$, darker green panels imply the reverse.

in KiDS-Legacy caused by variable depth is $< 1\%$ and consistent with the semi-analytical predictions from Heydenreich et al. (2020).

To test the uncertainty modelling in KiDS-Legacy, we propose and derive a novel measure to compare two covariance matrices: $k_{\min}$. It is an indicator of the difference in the overall entropy between the probability density distributions defined by two different covariance matrices. We then apply this measure to test the analytic uncertainty model, while also ascertaining the effect that the geometry of the survey footprint and variable depth have on the uncertainty.

Firstly, we determine that the KiDS-1000 analytic uncertainty model is overall in good agreement with the numerical model when similar assumptions about the survey geometry are made in both, even when including a sixth KiDS-Legacy-like tomographic bin. At the same time, we conclude that the exclusion of variable depth in the uncertainty model only causes minor biases in the analytical model which can be neglected at the precision level of stage III surveys such as KiDS. However, we note that neglecting the survey geometry in the off-diagonal cosmic variance terms of the cosmic shear covariance matrix may bias the analytic uncertainty model when compared to numerical modelling. We note that the impact on the constraining power due to this bias can vary strongly depending on the number of tomographic bins considered in the cosmic shear analysis. To determine the impact this would have on the constraints of cosmological parameters, further work is needed to forecast the impact of these choices in the covariance model on the posterior distributions.

When isolating different modelling choices in the forward-simulations, we find that, although the choice of footprint geometry can have a significant effects on the uncertainty model, it does not exceed previous results for KiDS-1000, even when including the sixth tomographic bin. This is also the case for variable depth where its impact on the overall uncertainty does not largely vary when including an additional tomographic bin. Hence, we conclude that the sensitivity of the uncertainty to survey geometry and variable depth is largely unaffected by the inclusion of an additional tomographic bin between redshifts of 1.2 and 2.0.

With this analysis, we also conclude that the inclusion of variable depth in the modelling reduces the diagonal terms of the covariance of the cosmic shear signal, despite variable depth causing additional spatial variance in the galaxy positions, redshifts and shapes. This previously unobserved effect arises due to the shape of the frequency distribution of the root-mean square of the background noise, $\sigma_{\rm rms}$, i.e. a measure of local observational depth. Depending on the symmetry of this distribution, the measured mean galaxy density and galaxy shape dispersion do not fully characterise the distribution. Consequently, as the shape noise varies with $\sigma^3_{\rm rms}$

and other noise terms respond non-linearly to changes in the galaxy density, any small skewness in the distribution can produce large changes in the uncertainty. We find that in the KiDS-1000 data, this distribution is asymmetric such that the median and the mode are below the mean. In consequence, the inclusion of variable depth in the uncertainty model unbiases the estimation of the galaxy density which reduces the shape noise and other near-diagonal terms in the cosmic shear covariance by approximately 10%. Other current and upcoming galaxy surveys which have different relations between observational depth and galaxy density/galaxy shape dispersion may encounter smaller/larger effects on the estimated uncertainty. In any case, this highlights the importance of forecasting the effects of variable depth on the uncertainty of the cosmic shear signal.

Having said that, the overall constraining power is still reduced by variable depth as the spatial variation introduces stronger correlations between the off-diagonal terms in the covariance of cosmic shear 2PCF. Lastly, we also find that the effects on the noise model due to the inclusion of variable depth and due to the adoption of a realistic footprint at least partially counteract each other. In consequence, the obtained constraints when including both systematics in the uncertainty model can appear to be better than the constraints from an uncertainty model which only includes one of the two effects. Throughout, we find the results are consistent with the tests conducted for KiDS-1000 (Joachimi et al., 2021), even when including the sixth tomographic bin.

Hence, we reach the conclusion that the analytical uncertainty modelling established for KiDS-1000 is sufficient for the sensitivity of KiDS-Legacy. However, this analysis is based on preliminary redshifts distributions, galaxy densities and galaxy shape dispersions which are subject to change when KiDS DR5 is finalised. In addition, the forward simulations assume a KiDS-1000-like footprint which is smaller than the KiDS-Legacy footprint, while having a more complex geometry. Accordingly, it will be necessary to reevaluate the analysis shown in this chapter once KiDS DR5 and the KiDS-Legacy mask are available.

Chapter 5

Conclusions

In this thesis, I show different ways in which forward models can be used to better understand large-scale structure observables to test cosmological models, and improve our understanding of gravity, dark matter and dark energy.

Firstly, I present a novel method, MAGBET, to estimate the magnitude of the magnification bias observed in the galaxy-galaxy lensing and galaxy clustering signals from galaxy surveys which are not flux-limited. The method involves calibrating the observed luminosity function slope from N-body simulations which incorporate the complex selection function of the galaxy survey in question. I apply this method to estimating the magnification bias in the BOSS DR12 data, and these are implemented in the KiDS-1000 3x2pt analysis (Heymans et al., 2021) and the KiDS-1000 LRG sample analysis (Fortuna et al., 2021). I also forecast the impacts which the magnification bias may have on the galaxy-galaxy lensing measurements from combining KiDS-1000, HSC Wide and *Euclid* sources with BOSS or DESI lenses. I find that next-generation galaxy-galaxy lensing analysis may be able to detect magnification directly, and use it as an additional probe for testing astrophysical and cosmological models.

Secondly, I outline a new suite of efficient and accurate forward simulations of cosmic shear observables, KIDS-SBI, which include all relevant systematics considered in a standard KiDS cosmic shear analysis. The KIDS-SBI forward simulations include a novel module to compute non-Limber projections of angular correlation functions (LEVIN), implement an efficient code to generate accurate log-normal random matter fields consistent with a given input cosmology (GLASS),

and incorporate the modelling of systematics such as intrinsic alignments, baryonic feedback, shear biases, variable depth, etc. These simulations are efficient enough to freely compute as many evaluations as needed to cover a large parameter space within computational runtimes comparable to running standard Markov-Chain Monte Carlo (MCMC) methods. This then allows me to use these simulations to conduct the first full simulation-based inference analysis of large-scale structure observables at the same level of complexity as current state-of-the-art cosmic shear analyses.

When propagating the uncertainties from the data to the posteriors in a non-Gaussian manner using SBI, I find that the likelihood of cosmic shear two-point statistics appears to have non-negligible deviations from Gaussianity which are not typically considered in the standard cosmic shear analyses. The observed deviations from standard Gaussian likelihood likelihood analyses are driven by a cosmology-dependence in the uncertainties. I find that the reduction in the cosmological constraints is consistent with cosmic variance increasing with S_8. At the same time, I expect at least part of the non-Gaussian contributions to the uncertainty in the cosmic shear being caused by the full uncertainty propagation of systematic effects, such as the intrinsic alignments, shear biases, the survey footprint or variable depth, whose contributions to the noise may or may not be Gaussian. This cosmology-dependence in the uncertainty appears to increase the relative uncertainties on S_8 from 7% to 9% when compared to previous KiDS-1000 cosmic shear analyses. Consequently, if these constraints persist when applied to the real KiDS-1000 data, the σ_8 "tension" observed between late-Universe large-scale structure probes, such as KiDS-1000 cosmic shear, and early-Universe probes, such as *Planck* 2018 CMB measurements, would decrease from 2.8σ to approximately 2.2σ. This implies that the constraining power of current cosmic shear surveys may not be enough to confidently determine the consistency/inconsistency between early- and late-Universe observations, even when the assumption of a Gaussian likelihood is dropped. To help resolve or confirm this "tension", the methods outlined in this thesis can be used in future work to facilitate the scientific exploitation of the upcoming data

from stage-IV galaxy surveys and other cosmological probes.

Besides providing a full non-Gaussian Bayesian uncertainty propagation, the SBI analysis in this thesis also highlights many other advantages of the method which may help address many of the challenges facing cosmology and astrophysics in the future. The next generation of galaxy surveys, such as *Euclid*, *Rubin* and *Roman*, will drastically improve the signal-to-noise ratio for cosmic shear, galaxy clustering and galaxy-galaxy lensing. Consequently, many systematic effects (such as the magnification bias, variable depth, shear biases, etc.) will no longer be obscured by random uncertainties. Modelling their contribution to the observed shear signal as well as its uncertainty will become a necessity. Since current analytical models of such effects can be inaccurate or intractable, SBI offers the opportunity to incorporate numerical models of any systematic effect into inference pipelines. The same applies for non-standard summary statistics (e.g. higher-order correlation functions, fields, etc.) whose likelihood functions may be non-Gaussian and/or analytically intractable, even when disregarding systematic effects. Additionally, the uncertainty propagation facilitated by SBI is not only conducive to parameter inference, but it also allows to perform rigorous Bayesian model testing to identify physical models which are preferred by observations. Therefore, SBI has the potential to become a key method in order to learn as much as possible about dark matter, dark energy, gravity and astrophysics from the plethora of galaxy survey data that is going to become available over the coming years.

Finally, I describe a separate suite of KiDS-SBI simulations which is designed to test the signal and uncertainty models for the upcoming KiDS-Legacy cosmic shear analysis. To formalise this testing, I propose a new statistical measure, k_{min}, which quantifies the relative distance between two matrices along the manifold of all real, symmetric positive-definite matrices of the same dimensions. I conclude that the signal modelling is largely unaffected by the consideration of a realistic survey footprint or by the modelling of variable depth, even when considering the additional sensitivity which KiDS-Legacy is going to have thanks to a potential sixth tomographic bin. Moreover, I find that the analytical uncertainty model used

within the KiDS-1000 analysis is consistent with the numerical modelling based on the KiDS-SBI forward simulations, while the analytical model is also sufficiently accurate for KiDS-Legacy despite not including variable depth in its modelling.

During this analysis, I discover that, despite variable depth adding additional angular fluctuations in galaxy density and shape to galaxy observations, when accurately forward-modelling the effect for KiDS-Legacy, the overall uncertainty of the 2PCF cosmic shear measurement decreases when compared to a model which assumes isotropic depth. This unexpected result is driven by the survey strategy of KiDS. It causes anisotropies in the observational depth that are distributed with positive skewness, such that the true shape noise across the majority of the sky is below the measured mean value. In the case of KiDS, modelling the variable depth in the uncertainty model reduces the estimated shape noise by up to 10%.

In future galaxy surveys, it will become even more important to consider variable depth. It is predicted to be a sub-percent level systematic in the cosmic shear, galaxy clustering and galaxy-galaxy lensing signal, so it may become detectable in the signal of stage-IV surveys. Simultaneously, as the statistical power of these measurements improves, the importance of variable depth in the noise model will become further exacerbated. In *Euclid*, for example, angular variable depth will not necessarily be as strong of an effect as in ground-based surveys, because of the absence of atmospheric seeing. Nonetheless, variable depth may still significantly contaminate measurements in *Euclid* due to its sensitivity to zodiacal background light, galactic extinction, camera degradation over time, overlaps between pointings, etc. I also show in this thesis that the effects of variable depth on the redshift selection of galaxy is an important driver of bias and uncertainty in the cosmic shear measurement. All upcoming photometric galaxy surveys, including the space-based *Euclid* survey, depend on some degree of redshift calibration from ground-based observations where variations in atmospheric seeing can be substantial. Hence, future work should also aim to assess the impact of variable depth on photometric redshifts from stage-IV galaxy surveys.

The work showcased in this thesis underlines the importance of accurate and

true-to-observations modelling of observables in cosmology. With stage-IV galaxy surveys already beginning to take observations with unprecedented precision, scale and depth, it will become necessary to drastically increase the accuracy of models in order to take full advantage of the large breadth of data that will come to be available. To address this, the work outlined here could be expanded to include a wider scope of large-scale structure probes, while also adding additional realism.

Appendix A

Magnification Bias: Flux-Limited Case

As discussed in Section 2.2.2, we conduct a sanity check of our method by comparing the effective $\overline{\alpha}_\kappa$ to $\overline{\alpha}_{\rm obs}$ for each redshift bin given a simulated magnitude-limited ($i < 20.2$) galaxy population spanning the whole sky. This sample is also based on MICE2 simulations. We estimate $\overline{\alpha}_\kappa$ from the known matter convergence κ and the relative difference between the lensed and unlensed cumulative galaxy number counts finding that $\overline{\alpha}_\kappa^{\rm zlow} = 0.97 \pm 0.13$ in the zlow bin ($0.2 < z \leq 0.5$) and that $\overline{\alpha}_\kappa^{\rm zhigh} = 3.15 \pm 0.10$ in the zhigh bin ($0.5 < z \leq 0.75$).

The $\overline{\alpha}_\kappa$ values are compared to $\overline{\alpha}_{\rm obs}$ in Figure A.1. We find that in zlow, $\overline{\alpha}_{\rm obs}$ optimally overlaps with the $\overline{\alpha}_\kappa$ estimate from the convergence when taking the weighted mean of $\alpha_{\rm obs}(m)$ over a magnitude range of $\Delta i = 0.67$ below the effective magnitude limit; giving $\overline{\alpha}_{\rm obs}^{\rm zlow} = 0.96 \pm 0.06$. This agrees well with the $\overline{\alpha}_\kappa$ of galaxies in zlow ($\overline{\alpha}_\kappa^{\rm zlow} = 0.97 \pm 0.13$). The agreement is similarly good in the zhigh bin where the optimal $\overline{\alpha}_{\rm obs}$ is computed over a $\Delta i = 0.48$ and found to be $\overline{\alpha}_{\rm obs}^{\rm zhigh} = 3.12 \pm 0.20$. The excellent agreement between $\overline{\alpha}_\kappa$ and when evaluating near the faint end of the sample ($\Delta i < 0.7$) reinforces that Equation 2.5 and Equation 2.6 indeed describe the same α; at least for the flux-limited. Such a good agreement is not really surprising, since the underlying assumptions leading to Equation 2.5 ($|\kappa| \ll 1; |\gamma| \ll 1$) are ingrained in the way the MICE2 simulations determine the magnified magnitude and position of galaxy (Fosalba et al., 2015a). Nonetheless, it still provides a check which allows us to understand how the method

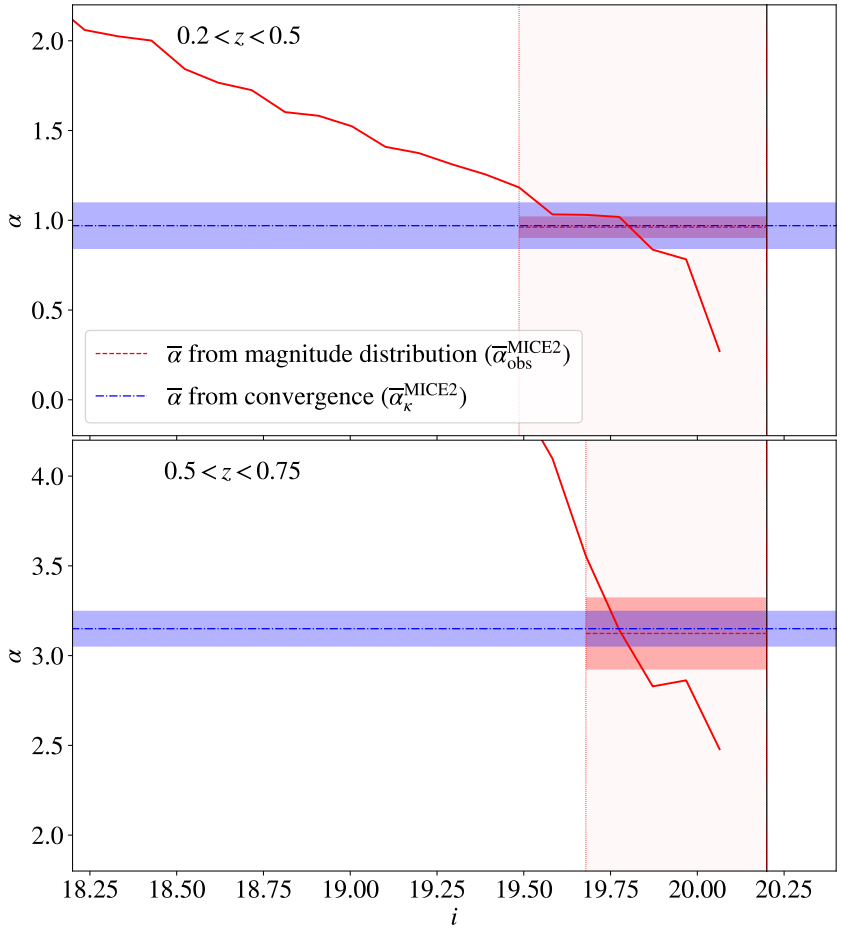

Figure A.1: The slope of the luminosity function, α, as a function of the i-band magnitude, i, for the magnitude-limited case. Two redshift bins are considered: $0.2 < z \leq 0.5$ (top) and $0.5 < z \leq 0.75$ (bottom). The vertical black line marks the magnitude limit at $i = 20.2$ and the dashed red vertical lines mark the upper and lower bounds of the highlighted magnitude range which was used to determine $\overline{\alpha}^{\rm MICE2}_{\rm obs}$. The dashed red horizontal line marks the $\overline{\alpha}^{\rm MICE2}_{\rm obs}$ estimate and the blue dot-dashed horizontal line marks the effective $\overline{\alpha}^{\rm MICE2}_{\kappa}$ determined from the weak lensing convergence with Equation 2.5 and used to calibrate $\overline{\alpha}^{\rm MICE2}_{\rm obs}$.

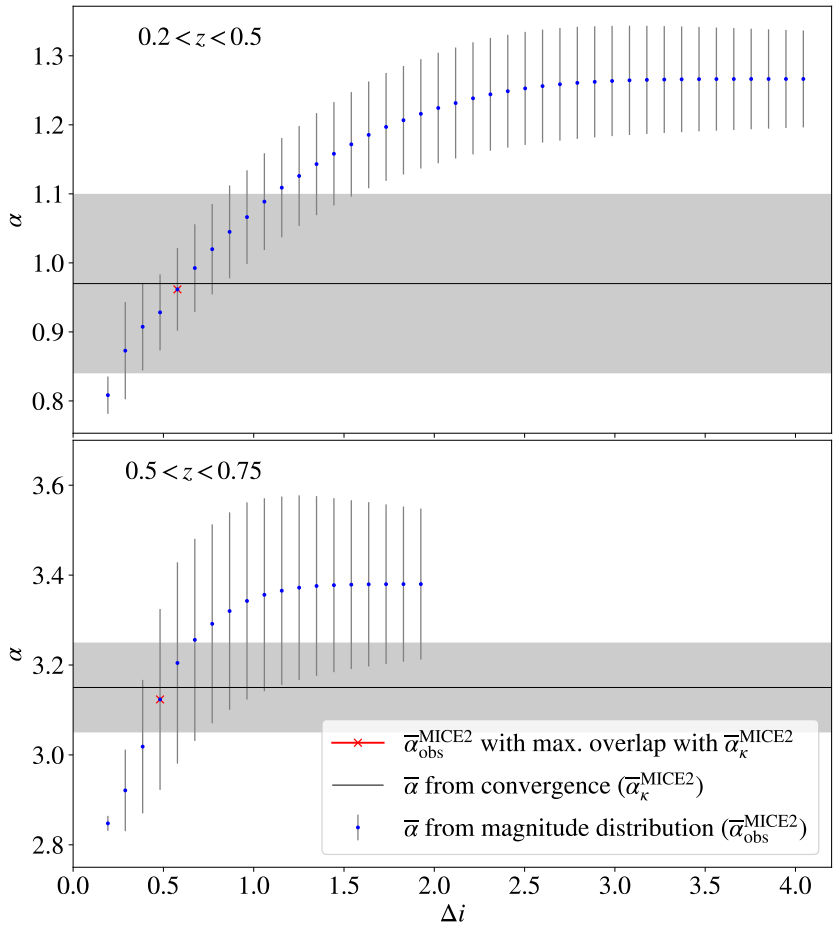

Figure A.2: $\overline{\alpha}_{\rm obs}$ estimates from the MICE2 simulations for the magnitude-limited case ($i < 20.2$) over i-band magnitude ranges below the turn-off magnitude (Δi) considered to calculate the weighted average. Two redshift bins are considered: $0.2 < z \leq 0.5$ (top) and $0.5 < z \leq 0.75$ (bottom). The red cross marks the $\overline{\alpha}_{\rm obs}$ estimate which overlaps the most with the $\overline{\alpha}_{\kappa}$ estimate from the weak lensing convergence (black line).

estimates α in the absence of a complex selection function .

In addition, when looking at Figure A.2, one finds that for zlow the $\overline{\alpha}_{\rm obs}$ estimates are accurate over a large domain of magnitude ranges (being less than 1σ apart when considering Δi anywhere between 0 and ~ 2). This confirms that a power law is a good approximation for the luminosity function over a large magnitude range near the faint end of the distribution which implies that, in a magnitude-limited survey, $\overline{\alpha}_{\rm obs}$ estimates are robust and accurate even after substantial changes in the magnitude range considered. We find similarly good agreement between the $\overline{\alpha}_{\kappa}$ and $\overline{\alpha}_{\rm obs}$ in the zhigh bin where $\overline{\alpha}_{\rm obs}^{\rm zhigh} = 3.12 \pm 0.2$, while Figure A.2 shows

that this estimate is robust at high redshifts.

Despite the consistency between $\overline{\alpha}_\kappa$ and $\overline{\alpha}_{\rm obs}$ and the robustness of the estimate to small changes in the calibration magnitude range Δi, it is surprising to see such a drastic increase in $\overline{\alpha}_{\rm obs}$ between zlow and zhigh. This seems to be a consequence of the magnitude limit at $i = 20.2$ being low enough to exclude a substantial fraction of faint galaxies at high redshifts, such that the power law in flux assumed in Equation 2.2 no longer applies. If we consider the luminosity function of the galaxies as a Schechter function (Schechter, 1976), such a selection of bright galaxies would lead to a dominant exponential term in the Schechter function which leads to overestimates of α. In general, this is not of much concern, since most magnitude limited surveys operate within a regime where the power law approximation holds.

Appendix B

KiDS-SBI: Theoretical Signal Modelling

To compute the derivatives needed for score compression, and to provide consistency checks, we also theoretically model the expected signal for a given mock. For this, we calculate the matter power spectrum using CAMB and proceed to make a Limber projection to obtain the following cosmic shear angular power spectrum

$$C_{\varepsilon\varepsilon}^{(pq)}(\ell) = \frac{(\ell+2)!}{4(\ell-2)!}\int_0^\infty \frac{d\chi}{f_k^2(\chi)} W_\varepsilon^{(p)}(\chi) W_\varepsilon^{(q)}(\chi) P_\delta\left(\frac{\ell+1/2}{f_k(\chi)}, \chi\right), \tag{B.1}$$

where $W_{\rm G}$ is the weak lensing kernel, given by

$$W_\varepsilon^{(p)}(\chi) = \frac{3H_0^2\Omega_{\rm m}}{2}\frac{f_{\rm K}(\chi)}{a(\chi)}\int_\chi^{\chi_{\rm hor}} {\rm d}\chi' n_{\rm S}^{(p)}(\chi')\frac{f_{\rm K}(\chi'-\chi)}{f_{\rm K}(\chi')}, \tag{B.2}$$

where $n_{\rm S}^{(p)}$ is the redshift distribution of the objects in jth tomographic bin as shown in Figure 3.2.

Knowing the full-sky angular power spectra, $C_{\varepsilon\varepsilon,\nu}^{(pq)}(\ell';\mathbf{\Theta})$, we can define the pseudo-Cl, $\tilde{C}_{\varepsilon\varepsilon,\mu}^{(pq)}(\ell;\mathbf{\Theta})$, as follows (Peebles, 1973; Brown et al., 2005; Hikage et al., 2011),

$$\tilde{C}_{\varepsilon\varepsilon,\mu}^{(pq)}(\ell;\mathbf{\Theta}) = \sum_{\ell'=0}^{\ell'_{\rm max}}\sum_{\ell''=0}^{\ell''_{\rm max}}\sum_{\nu=1}^{3}\sum_{\nu'=1}^{3} M_{\mu\nu',\ell\ell''} M'^{(pq)}_{\nu'\nu,\ell''\ell'} C_{\varepsilon\varepsilon,\nu}^{(pq)}(\ell';\mathbf{\Theta}), \tag{B.3}$$

where $\mu,\nu,\nu' \in \{1,2,3\}$ such that 1 stands for the EE component, 2 for the BB

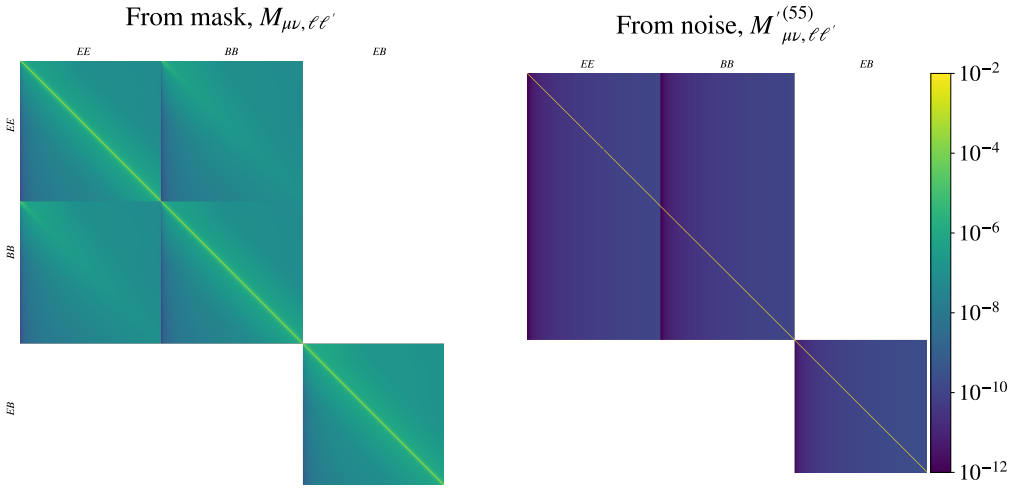

Figure B.1: Bitmap of the mixing matrices to model the theoretical signal for the pseudo-Cls as seen by KiDS-1000. On the right panel is the mixing matrix derived from the KiDS-1000 mask, $M_{\mu\nu',\ell\ell''}$, (the input mask is at a resolution of $N_{\rm side} = 1024$). On the left panel is the mixing matrix caused by the selection due to randomly sampling galaxies for the auto-correlation of the fifth tomographic bin, $M'^{(55)}_{\nu'\nu,\ell''\ell'}$. Both matrices are decomposed into block matrices separating the $EE \to EE$, $EE \to BB$ and $EB \to EB$ mixing. Each block matrix has dimensions of 8193×8193 with $\ell \in \{\ell \in \mathbb{Z}^{0+} | \ell \leq 8192\}$.

component and 3 for the EB component, $M_{\ell\ell''}$ is the mixing matrix of the survey mask, $\mathrm{W}(\theta)$, which does not vary between tomographic bins, and $M'^{(pq)}_{\ell''\ell'}$ is the mixing matrix of the effective mask imposed by the average random variations in the observed galaxies in a given tomographic bin over many realisations.

The first mixing matrix from the left in (B.3), $M_{\ell\ell''}$, is the standard block mixing matrix to account for the partial sky coverage in a survey (Hivon et al., 2002; Bennett et al., 2003; Brown et al., 2005). The second mixing matrix, $M'^{(pq)}_{\ell''\ell'}$, accounts for an additional effect: the mode mixing induced by a mask of uncorrelated noise originating from the random sampling described in Sect. 3.2.6. In other words, when randomly Poisson sampling galaxies within a discrete pixel k given a probability of $P^{(m)}(N)$, there will be a non-zero probability of observing zero counts in that given pixel. In essence, all the pixels which are sampled to have zero galaxy counts impose an additional masking on our sample. This can become significant at the resolutions and galaxy densities which we probe in this work. For

example, for an HEALPIX $N_{\rm side} = 2048$ and a galaxy density of $n = 1\,arcmin^{-2}$ for a given tomographic bin k (similar to KiDS-1000), $P^{(m)}(N = 0) \approx 0.07$. This means that nearly a tenth of all pixels could be "masked" in a given realisation of simulations at a higher resolution. To allow for these cases, we decide to also model the mode mixing induced by the sampling variance. However, for the settings discussed in this work for the main pipeline of forwards simulations where $N_{\rm side} = 1024$, this effect is not so relevant. Even for the sparsest source bin, S1, with an $n_{\rm gal}^{(1)} = 0.62\,{\rm arcmin}^{-2}$, $P^{(m)}(N = 0) \approx 0.0005$, so only a negligible fraction of pixels are masked due to random sample variance.

To provide a general overview, the distinguishing factor between the two mixing matrices is how we define the mask $\mathrm{W}(\theta)$ which characterises them. For $M_{\mu\nu,\ell\ell''}$, $\mathrm{W}(\theta)$ is defined as the mask of the survey footprint, $\Omega_{\rm survey}$, such that

$$\mathrm{W}(\theta) = \begin{cases} 1, & \theta \in \Omega_{\rm survey} \\ 0, & \theta \notin \Omega_{\rm survey} \end{cases}. \tag{B.4}$$

For $M'^{(pq)}_{\mu\nu,\ell\ell''}$, rather than defining it using an explicit spatial mask from the pixels masked by each random realisation, $\mathrm{W}'^{(p)}(\theta)$, we can define the expectation value of the mixing matrix due to the random sampling of galaxies directly from the angular power spectrum of $\mathrm{W}'^{(p)}(\theta)$ as

$$\frac{\mathscr{W}'^{(pq)}(\ell)}{2\ell+1} = \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell} w_{\ell m}^{(p)} w_{\ell m}^{(q)*} = \frac{4\pi\sigma^{(pq)}}{N_{\rm pix}-1} + 4\pi\left[\mu^{(pq)} - \frac{\sigma^{(pq)}}{N_{\rm pix}-1}\right]\delta_{\ell 0}, \tag{B.5}$$

where $\mu^{(pq)} = P^{(p)}(N>0)P^{(q)}(N>0)$, i.e. the product of the mean probabilities of success within each tomographic bin (the probability that a given pixel will be populated by at least one galaxy), while $2\sigma^{(pq)2} = P^{(p)2}(N>0)(1-P^{(p)}(N>0))^2 + P^{(q)2}(N>0)(1-P^{(q)}(N>0))^2$, i.e. the associated standard deviation of the

mean, and $N_{\rm pix}$ is the number of pixels.

As described in Sect. 3.2.6, the galaxies are Poisson sampled according to the probability of success is given in (3.22). As we are interested in the the mean over many iterations, we can assume that the matter overdensities are small, $1 + b^{(p)}\delta^{(p)} \approx 1$, so we can rewrite the probability as

$$P^{(p)}(N>0) = 1 - P^{(p)}(N=0) = 1 - e^{-\langle N_m^{(p)}\rangle(\mathbf{\Theta})} \approx 1 - e^{-n_{\rm gal}^{(p)} A_{\rm pix}}, \tag{B.6}$$

where $n_{\rm gal}^{(p)}$ is the mean galaxy density per tomographic bin p and $A_{\rm pix}$ is the mean pixel size for a given HEALPIX resolution. From this, we can calculate both mixing matrices as shown in Fig. B.1. In practice, we can see that from Fig. B.1 that the mode mixing $M'^{(pq)}_{\mu\nu,\ell''\ell'}$ is negligible when compared to $M_{\mu\nu,\ell\ell''}$ at an $N_{\rm side} = 1024$. However, we still include the term in the signal modelling to allow the flexibility to choose higher resolutions within KIDS-SBI.

With (B.4) and (B.5) defining the selections which define $M_{\mu\nu,\ell\ell''}$ and $M'^{(pq)}_{\mu\nu,\ell''\ell'}$, respectively, we can compute the matrices as follows (Brown et al., 2005; Hikage et al., 2011):

$$M_{\mu\nu,\ell\ell''} = \begin{pmatrix} W_{\ell\ell''}^{++} & (W_{\ell\ell''}^{-+} + W_{\ell\ell''}^{+-}) & W_{\ell\ell''}^{--} \\ -W_{\ell\ell''}^{+-} & (W_{\ell\ell''}^{++} - W_{\ell\ell''}^{--}) & W_{\ell\ell''}^{-+} \\ W_{\ell\ell''}^{--} & -(W_{\ell\ell''}^{-+} + W_{\ell\ell''}^{+-}) & W_{\ell\ell''}^{++} \end{pmatrix}, \tag{B.7}$$

$$M'^{(pq)}_{\mu\nu,\ell''\ell'} = \begin{pmatrix} W'^{++}_{\ell''\ell'} & (W'^{-+}_{\ell''\ell'} + W'^{+-}_{\ell''\ell'}) & W'^{--}_{\ell''\ell'} \\ -W'^{+-}_{\ell''\ell'} & (W'^{++}_{\ell''\ell'} - W'^{--}_{\ell''\ell'}) & W'^{-+}_{\ell''\ell'} \\ W'^{--}_{\ell''\ell'} & -(W'^{-+}_{\ell''\ell'} + W'^{+-}_{\ell''\ell'}) & W'^{++}_{\ell''\ell'} \end{pmatrix}^{(pq)}, \tag{B.8}$$

where each of the Wigner symbols is given by

$$W_{\ell\ell'}^{\pm\pm} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \sum_{m'=-\ell'}^{\ell'} W_{\ell\ell' mm'}^{\pm} (W_{\ell'\ell m'm}^{\pm})^*. \tag{B.9}$$

Since the shear field is a spin 2 field

$$W^{+}_{\ell\ell' mm'} = \frac{1}{2}({}_2W^{mm'}_{\ell\ell'} + {}_{-2}W^{mm'}_{\ell\ell'}), \tag{B.10}$$

$$W^{-}_{\ell\ell' mm'} = \frac{i}{2}({}_2W^{mm'}_{\ell\ell'} - {}_{-2}W^{mm'}_{\ell\ell'}), \tag{B.11}$$

where

$${}_sW^{+}_{\ell\ell' mm'} = \int \mathrm{d}\theta\, {}_sY_{\ell' m'}(\theta)\mathrm{W}(\theta)\, {}_sY^{*}_{\ell m}(\theta), \tag{B.12}$$

The latter is difficult to compute, so it is more convenient to compute in phase space such that

$$\sum_{m=-\ell}^{\ell}\sum_{m'=-\ell'}^{\ell'} {}_sW^{mm'}_{\ell\ell'}({}_{s'}W^{m'm}_{\ell'\ell})^{*} = \frac{(2\ell+1)(2\ell'+1)}{4\pi}\sum_{\ell''=0}^{\ell''_{\max}} \mathscr{W}_{\ell''}\begin{pmatrix}\ell & \ell' & \ell''\\ -s & s & 0\end{pmatrix}\begin{pmatrix}\ell & \ell' & \ell''\\ -s' & s' & 0\end{pmatrix}, \tag{B.13}$$

where $s \in \{-2, 2\}$ and

$$\mathscr{W}_{\ell} = \sum_{m=-\ell}^{\ell} w_{\ell m} w^{*}_{\ell m}, \tag{B.14}$$

where $w_{\ell m}$ are the coefficients of the spherical harmonic transform of the mask, $\mathrm{W}(\theta)$, such that

$$w_{\ell m} = \int \mathrm{d}\theta\, \mathrm{W}(\theta) Y^{*}_{\ell m}(\theta). \tag{B.15}$$

Bibliography

Abazajian, K. N., Adshead, P., Ahmed, Z., et al. 2016, CMB-S4 Science Book, First Edition, arXiv e-prints, arXiv:1610.02743

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, Observation of Gravitational Waves from a Binary Black Hole Merger, Phys. Rev. Lett., 116, 061102

Abbott, T. M. C., Alarcon, A., Allam, S., et al. 2019a, Cosmological Constraints from Multiple Probes in the Dark Energy Survey, Phys. Rev. Lett., 122, 171301

Abbott, T. M. C., Abdalla, F. B., Alarcon, A., et al. 2019b, Dark Energy Survey year 1 results: Joint analysis of galaxy clustering, galaxy lensing, and CMB lensing two-point functions, Phys. Rev. D, 100, 023541

Abbott, T. M. C., Aguena, M., Alarcon, A., et al. 2022, Dark Energy Survey Year 3 results: Cosmological constraints from galaxy clustering and weak lensing, Phys. Rev. D, 105, 023520

Abramo, L. R., Dinarte Ferri, J. V., Tashiro, I. L., & Loureiro, A. 2022, Fisher matrix for the angular power spectrum of multi-tracer galaxy surveys, J. Cosm.&Astrop. Phys., 2022, 073

Abramo, L. R., Secco, L. F., & Loureiro, A. 2016, Fourier analysis of multitracer cosmological surveys, MNRAS, 455, 3871

Ade, P., Aguirre, J., Ahmed, Z., et al. 2019, The Simons Observatory: science goals and forecasts, J. Cosm.&Astrop. Phys., 2019, 056

Aihara, H., Armstrong, R., Bickerton, S., et al. 2018, First data release of the Hyper Suprime-Cam Subaru Strategic Program, PASJ, 70, S8

Aihara, H., AlSayyad, Y., Ando, M., et al. 2022, Third data release of the Hyper Suprime-Cam Subaru Strategic Program, PASJ, 74, 247

Aitchison, J., & Brown, J. A. C. 1969, THE LOGNORMAL DISTRIBUTION, WITH SPECIAL REFERENCE TO ITS USES IN ECONOMICS,.

Akita, K., & Yamaguchi, M. 2020, A precision calculation of relic neutrino decoupling, J. Cosm.&Astrop. Phys., 2020, 012

Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III, ApJS, 219, 12

Alam, S., Ata, M., Bailey, S., et al. 2017, The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample, MNRAS, 470, 2617

Alam, S., Aubert, M., Avila, S., et al. 2021, Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological implications from two decades of spectroscopic surveys at the Apache Point Observatory, Phys. Rev. D, 103, 083533

Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, Report of the Dark Energy Task Force, arXiv e-prints, astro

ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, et al. 2006, Precision electroweak measurements on the Z resonance, Phys. Rep., 427, 257

Alonso, D., Sanchez, J., Slosar, A., & LSST Dark Energy Science Collaboration. 2019, A unified pseudo-C_ℓ framework, MNRAS, 484, 4127

Alsing, J., Charnock, T., Feeney, S., & Wandelt, B. 2019, Fast likelihood-free cosmology with neural density estimators and active learning, MNRAS, 488, 4440

Alsing, J., & Wandelt, B. 2018, Generalized massive optimal data compression, MNRAS, 476, L60

Alsing, J., Wandelt, B., & Feeney, S. 2018, Massive optimal data compression and density estimation for scalable, likelihood-free inference in cosmology, Monthly Notices of the Royal Astronomical Society, 477, 2874

Amon, A., & Efstathiou, G. 2022, A non-linear solution to the S_8 tension?, MNRAS, 516, 5355

Amon, A., Gruen, D., Troxel, M. A., et al. 2022, Dark Energy Survey Year 3 results: Cosmology from cosmic shear and robustness to data calibration, Phys. Rev. D, 105, 023514

Amon, A., Robertson, N. C., Miyatake, H., et al. 2023, Consistent lensing and clustering in a low-S_8 Universe with BOSS, DES Year 3, HSC Year 1, and KiDS-1000, MNRAS, 518, 477

Anderson, L., Aubourg, E., Bailey, S., et al. 2012, The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Release 9 spectroscopic galaxy sample, MNRAS, 427, 3435

Asgari, M., Lin, C.-A., Joachimi, B., et al. 2021, KiDS-1000 cosmology: Cosmic shear constraints and comparison between two point statistics, A&A, 645, A104

Bacon, D. J., Refregier, A. R., & Ellis, R. S. 2000, Detection of weak gravitational lensing by large-scale structure, MNRAS, 318, 625

Baldauf, T., Smith, R. E., Seljak, U., & Mandelbaum, R. 2010, Algorithm for the direct reconstruction of the dark matter correlation function from weak lensing and galaxy clustering, Phys. Rev. D, 81, 063531

Baleato Lizancos, A., & White, M. 2023, The impact of anisotropic redshift distributions on angular clustering, J. Cosm.&Astrop. Phys., 2023, 044

Bardeen, J. M. 1980, Gauge-invariant cosmological perturbations, Phys. Rev. D, 22, 1882

Bardeen, J. M., Steinhardt, P. J., & Turner, M. S. 1983, Spontaneous creation of almost scale-free density perturbations in an inflationary universe, Phys. Rev. D, 28, 679

Barrow, J. D., & Coles, P. 1990, The statistics of primordial density fluctuations, MNRAS, 244, 188

Bartelmann, M. 1995, Cluster mass estimates from weak lensing., A&A, 303, 643

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

Beaumont, M. A. 2019, Approximate Bayesian Computation, Annual Review of Statistics and Its Application, 6, 379

Becker, M. R., Troxel, M. A., MacCrann, N., et al. 2016, Cosmic shear measurements with Dark Energy Survey Science Verification data, Phys. Rev. D, 94, 022002

Benítez, N. 2000, Bayesian Photometric Redshift Estimation, ApJ, 536, 571

Bennett, C. L., Halpern, M., Hinshaw, G., et al. 2003, First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results, ApJS, 148, 1

Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results, ApJS, 208, 20

Beutler, F., Blake, C., Colless, M., et al. 2011, The 6dF Galaxy Survey: baryon acoustic oscillations and the local Hubble constant, MNRAS, 416, 3017

Beutler, F., Seo, H.-J., Saito, S., et al. 2017, The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: anisotropic galaxy clustering in Fourier space, MNRAS, 466, 2242

Bhattacharya, S., Heitmann, K., White, M., et al. 2011, Mass Function Predictions Beyond ΛCDM, ApJ, 732, 122

Bianchi, L. 1902, Sui simboli a quattro indici e sulla curvatura di Riemann: nota (Tip. della R. Accademia dei Lincei)

Binggeli, B., Sandage, A., & Tammann, G. A. 1988, The luminosity function of galaxies., ARA&A, 26, 509

Bird, S., Viel, M., & Haehnelt, M. G. 2012, Massive neutrinos and the non-linear matter power spectrum, MNRAS, 420, 2551

Bishop, C. M. 1994, Mixture density networks (Aston University)

Biswas, A., & Mani, K. R. S. 2008, Relativistic perihelion precession of orbits of Venus and the Earth, Central European Journal of Physics, 6, 754

Blake, C., Kazin, E. A., Beutler, F., et al. 2011, The WiggleZ Dark Energy Survey: mapping the distance-redshift relation with baryon acoustic oscillations, MNRAS, 418, 1707

Blazek, J., McQuinn, M., & Seljak, U. 2011, Testing the tidal alignment model of galaxy intrinsic alignment, J. Cosm.&Astrop. Phys., 2011, 010

Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. 1984, Formation of galaxies and large-scale structure with cold dark matter., Nature, 311, 517

Böhm, V., Hilbert, S., Greiner, M., & Enßlin, T. A. 2017, Bayesian weak lensing tomography: Reconstructing the 3D large-scale distribution of matter with a lognormal prior, Phys. Rev. D, 96, 123510

Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, Weak Gravitational Lensing by Galaxies, ApJ, 466, 623

Brandenberger, R. H. 1985, Quantum field theory methods and inflationary universe models, Reviews of Modern Physics, 57, 1

Bridle, S., & King, L. 2007, Dark energy constraints from cosmic shear power spectra: impact of intrinsic alignments on photometric redshift requirements, New Journal of Physics, 9, 444

Broadhurst, T., & Lehar, J. 1995, A Gravitational Lens Solution for the IRAS Galaxy FSC 10214+4724, ApJ, 450, L41

Brown, M. L., Castro, P. G., & Taylor, A. N. 2005, Cosmic microwave background temperature and polarization pseudo-C_l estimators and covariances, MNRAS, 360, 1262

Brown, M. L., Taylor, A. N., Hambly, N. C., & Dye, S. 2002, Measurement of intrinsic alignments in galaxy ellipticities, MNRAS, 333, 501

Castro, P. G., Heavens, A. F., & Kitching, T. D. 2005, Weak lensing analysis in three dimensions, Phys. Rev. D, 72, 023516

Catelan, P., Kamionkowski, M., & Blandford, R. D. 2001, Intrinsic and extrinsic galaxy alignment, MNRAS, 320, L7

Chen, A., Harness, A., & Melchior, P. 2022, Lightweight starshade position sensing with convolutional neural networks and simulation-based inference, arXiv e-prints, arXiv:2204.03853

Chiu, I., Dietrich, J. P., Mohr, J., et al. 2016, Detection of enhancement in number densities of background galaxies due to magnification by massive galaxy clusters, MNRAS, 457, 3050

Clampitt, J., Sánchez, C., Kwan, J., et al. 2017, Galaxy-galaxy lensing in the Dark Energy Survey Science Verification data, MNRAS, 465, 4204

Clemence, G. M. 1947, The Relativity Effect in Planetary Motions, Reviews of Modern Physics, 19, 361

Clerkin, L., Kirk, D., Manera, M., et al. 2017, Testing the lognormality of the galaxy and weak lensing convergence distributions from Dark Energy Survey maps, MNRAS, 466, 1444

Clowe, D., Bradač, M., Gonzalez, A. H., et al. 2006, A Direct Empirical Proof of the Existence of Dark Matter, ApJ, 648, L109

Cole, S., & Kaiser, N. 1989, Biased clustering in the cold dark matter cosmogony., MNRAS, 237, 1127

Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications, MNRAS, 362, 505

Coles, P., & Jones, B. 1991, A lognormal model for the cosmological mass distribution., MNRAS, 248, 1

Crocce, M., Castander, F. J., Gaztañaga, E., Fosalba, P., & Carretero, J. 2015, The MICE Grand Challenge lightcone simulation - II. Halo and galaxy catalogues, MNRAS, 453, 1513

Dark Energy Survey Collaboration, Abbott, T., Abdalla, F. B., et al. 2016, The Dark Energy Survey: more than dark energy - an overview, MNRAS, 460, 1270

Das, S., & Ostriker, J. P. 2006, Testing a New Analytic Model for Gravitational Lensing Probabilities, ApJ, 645, 1

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, The evolution of large-scale structure in a universe dominated by cold dark matter, ApJ, 292, 371

Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, The Baryon Oscillation Spectroscopic Survey of SDSS-III, AJ, 145, 10

de Salas, P. F., & Pastor, S. 2016, Relic neutrino decoupling with flavour oscillations revisited, J. Cosm.&Astrop. Phys., 2016, 051

Deshpande, A. C., Kitching, T. D., Cardone, V. F., et al. 2020, Euclid: The reduced shear approximation and magnification bias for Stage IV cosmic shear experiments, A&A, 636, A95

DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, The DESI Experiment Part I: Science,Targeting, and Survey Design, arXiv e-prints, arXiv:1611.00036

Dolag, K., Borgani, S., Schindler, S., Diaferio, A., & Bykov, A. M. 2008, Simulation Techniques for Cosmological Simulations, Space Sci. Rev., 134, 229

Duffy, A. R., Schaye, J., Kay, S. T., et al. 2010, Impact of baryon physics on dark matter structures: a detailed simulation study of halo density profiles, MNRAS, 405, 2161

Duncan, C. A. J., Harnois-Déraps, J., Miller, L., & Langedijk, A. 2022, On cosmological bias due to the magnification of shear and position samples in modern weak lensing analyses, MNRAS, 515, 1130

Duncan, C. A. J., Joachimi, B., Heavens, A. F., Heymans, C., & Hildebrandt, H. 2014, On the complementarity of galaxy clustering with cosmic shear and flux magnification, MNRAS, 437, 2471

Dyson, F. W., Eddington, A. S., & Davidson, C. 1920, A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations Made at the Total Eclipse of May 29, 1919, Philosophical Transactions of the Royal Society of London Series A, 220, 291

Eckart, A., & Genzel, R. 1996, Observations of stellar proper motions near the Galactic Centre, Nature, 383, 415

Edge, A., Sutherland, W., Kuijken, K., et al. 2013, The VISTA Kilo-degree Infrared Galaxy (VIKING) Survey: Bridging the Gap between Low and High Redshift, The Messenger, 154, 32

Eifler, T., Schneider, P., & Hartlap, J. 2009, Dependence of cosmic shear covariances on cosmology. Impact on parameter estimation, A&A, 502, 721

Einstein, A. 1905, Zur Elektrodynamik bewegter Körper, Annalen der Physik, 322, 891

—. 1915, Die Feldgleichungen der Gravitation, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, 844

—. 1917, Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, 142

Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies, ApJ, 633, 560

Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems, AJ, 142, 72

Ellis, G. F. R., & Bruni, M. 1989, Covariant and gauge-invariant approach to cosmological density fluctuations, Phys. Rev. D, 40, 1804

Ellis, G. F. R., Hwang, J., & Bruni, M. 1989, Covariant and gauge-independent perfect-fluid Robertson-Walker perturbations, Phys. Rev. D, 40, 1819

Elvin-Poole, J., Crocce, M., Ross, A. J., et al. 2018, Dark Energy Survey year 1 results: Galaxy clustering for combined probes, Phys. Rev. D, 98, 042006

Erben, T., Schirmer, M., Dietrich, J. P., et al. 2005, GaBoDS: The Garching-Bonn Deep Survey. IV. Methods for the image reduction of multi-chip cameras demonstrated on data from the ESO Wide-Field Imager, Astronomische Nachrichten, 326, 432

Euclid Collaboration, Blanchard, A., Camera, S., et al. 2020, Euclid preparation. VII. Forecast validation for Euclid cosmological probes, A&A, 642, A191

Euler, L. 1757, Principes généraux du mouvement des fluides, Mémoires de l'académie des sciences de Berlin, 274

Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019, First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole, ApJ, 875, L1

Fenech Conti, I., Herbonnet, R., Hoekstra, H., et al. 2017, Calibration of weak-lensing shear in the Kilo-Degree Survey, MNRAS, 467, 1627

Fischer, P., McKay, T. A., Sheldon, E., et al. 2000, Weak Lensing with Sloan Digital Sky Survey Commissioning Data: The Galaxy-Mass Correlation Function to 1 H^{-1} Mpc, AJ, 120, 1198

Flaugher, B., & DES Collaboration. 2013, in American Astronomical Society Meeting Abstracts, Vol. 221, American Astronomical Society Meeting Abstracts #221, 335.02

Fluri, J., Kacprzak, T., Lucchi, A., et al. 2022, Full w CDM analysis of KiDS-1000 weak lensing maps using deep learning, Phys. Rev. D, 105, 083518

Förstner, W., & Moonen, B. 2003, A metric for covariance matrices, Geodesy-the Challenge of the 3rd Millennium, 299

Fortuna, M. C., Hoekstra, H., Johnston, H., et al. 2021, KiDS-1000: Constraints on the intrinsic alignment of luminous red galaxies, A&A, 654, A76

Fosalba, P., Crocce, M., Gaztañaga, E., & Castander, F. J. 2015a, The MICE grand challenge lightcone simulation - I. Dark matter clustering, MNRAS, 448, 2987

Fosalba, P., Gaztañaga, E., Castander, F. J., & Crocce, M. 2015b, The MICE Grand Challenge light-cone simulation - III. Galaxy lensing mocks from all-sky lensing maps, MNRAS, 447, 1319

Friedmann, A. 1924, Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes, Zeitschrift fur Physik, 21, 326

Friedrich, O., Andrade-Oliveira, F., Camacho, H., et al. 2020, Dark Energy Survey Year 3 Results: Covariance Modelling and its Impact on Parameter Estimation and Quality of Fit, arXiv e-prints, arXiv:2012.08568

Garcia-Fernandez, M., Sanchez, E., Sevilla-Noarbe, I., et al. 2018, Weak lensing magnification in the Dark Energy Survey Science Verification data, MNRAS, 476, 1071

Gatti, M., Chang, C., Friedrich, O., et al. 2020, Dark Energy Survey Year 3 results: cosmology with moments of weak lensing mass maps - validation on simulations, MNRAS, 498, 4060

Gatti, M., Sheldon, E., Amon, A., et al. 2021, Dark energy survey year 3 results: weak lensing shape catalogue, MNRAS, 504, 4312

Gerardi, F., Feeney, S. M., & Alsing, J. 2021, Unbiased likelihood-free inference of the Hubble constant from light standard sirens, Phys. Rev. D, 104, 083531

Germain, M., Gregor, K., Murray, I., & Larochelle, H. 2015, in Proceedings of Machine Learning Research, Vol. 37, Proceedings of the 32nd International Conference on Machine Learning, ed. F. Bach & D. Blei (Lille, France: PMLR), 881–889

Giblin, B., Heymans, C., Asgari, M., et al. 2021, KiDS-1000 catalogue: Weak gravitational lensing shear measurements, A&A, 645, A105

Giocoli, C., Di Meo, S., Meneghetti, M., et al. 2017, Fast weak-lensing simulations with halo model, MNRAS, 470, 3574

Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere, ApJ, 622, 759

Grainger, R. 2013, An Atmospheric Radiative Transfer Primer (CRC Press London)

Guth, A. H. 1981, Inflationary universe: A possible solution to the horizon and flatness problems, Phys. Rev. D, 23, 347

Guth, A. H., & Pi, S. Y. 1982, Fluctuations in the New Inflationary Universe, Phys. Rev. Lett., 49, 1110

Guzik, J., & Bernstein, G. 2005, Inhomogeneous systematic signals in cosmic shear observations, Phys. Rev. D, 72, 043503

Hall, A., & Taylor, A. 2022, Non-Gaussian likelihood of weak lensing power spectra, Phys. Rev. D, 105, 123527

Hambly, N. C., MacGillivray, H. T., Read, M. A., et al. 2001, The SuperCOSMOS Sky Survey - I. Introduction and description, MNRAS, 326, 1279

Hamimeche, S., & Lewis, A. 2008, Likelihood analysis of CMB temperature and polarization power spectra, Phys. Rev. D, 77, 103013

Harms, R. J., Ford, H. C., Tsvetanov, Z. I., et al. 1994, HST FOS Spectroscopy of M87: Evidence for a Disk of Ionized Gas around a Massive Black Hole, ApJ, 435, L35

Heavens, A. 2009, in Data Analysis in Cosmology, ed. V. J. Martínez, E. Saar, E. Martínez-González, & M. J. Pons-Bordería, Vol. 665 (Springer), 585–600

Heavens, A., Refregier, A., & Heymans, C. 2000a, Intrinsic correlation of galaxy shapes: implications for weak lensing measurements, MNRAS, 319, 649

Heavens, A. F., Jimenez, R., & Lahav, O. 2000b, Massive lossless data compression and multiple parameter estimation from galaxy spectra, MNRAS, 317, 965

Heydenreich, S., Schneider, P., Hildebrandt, H., et al. 2020, The effects of varying depth in cosmic shear surveys, A&A, 634, A104

Heymans, C., & Heavens, A. 2003, Weak gravitational lensing: reducing the contamination by intrinsic alignments, MNRAS, 339, 711

Heymans, C., Van Waerbeke, L., Bacon, D., et al. 2006, The Shear Testing Programme - I. Weak lensing analysis of simulated ground-based observations, MNRAS, 368, 1323

Heymans, C., Grocutt, E., Heavens, A., et al. 2013, CFHTLenS tomographic weak lensing cosmological parameter constraints: Mitigating the impact of intrinsic galaxy alignments, MNRAS, 432, 2433

Heymans, C., Tröster, T., Asgari, M., et al. 2021, KiDS-1000 Cosmology: Multi-probe weak gravitational lensing and spectroscopic galaxy clustering constraints, A&A, 646, A140

Hikage, C., Takada, M., Hamana, T., & Spergel, D. 2011, Shear power spectrum reconstruction using the pseudo-spectrum method, MNRAS, 412, 65

Hikage, C., Oguri, M., Hamana, T., et al. 2019, Cosmology from cosmic shear power spectra with Subaru Hyper Suprime-Cam first-year data, PASJ, 71, 43

Hilbert, D. 1915, Die Grundlagen der Physik . (Erste Mitteilung.), Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, 1915, 395

Hilbert, S., Hartlap, J., & Schneider, P. 2011, Cosmic shear covariance: the log-normal approximation, A&A, 536, A85

Hilbert, S., Xu, D., Schneider, P., et al. 2017, Intrinsic alignments of galaxies in the Illustris simulation, MNRAS, 468, 790

Hildebrandt, H. 2016, Observational biases in flux magnification measurements, MNRAS, 455, 3943

Hildebrandt, H., van Waerbeke, L., & Erben, T. 2009, CARS: The CFHTLS-Archive-Research Survey. III. First detection of cosmic magnification in samples of normal high-z galaxies, A&A, 507, 683

Hildebrandt, H., Viola, M., Heymans, C., et al. 2017, KiDS-450: cosmological parameter constraints from tomographic weak gravitational lensing, MNRAS, 465, 1454

Hildebrandt, H., van den Busch, J. L., Wright, A. H., et al. 2021, KiDS-1000 catalogue: Redshift distributions and their calibration, A&A, 647, A124

Hirata, C. M., & Seljak, U. 2004, Intrinsic alignment-lensing interference as a contaminant of cosmic shear, Phys. Rev. D, 70, 063526

Hirata, C. M., Mandelbaum, R., Seljak, U., et al. 2004, Galaxy-galaxy weak lensing in the Sloan Digital Sky Survey: intrinsic alignments and shear calibration errors, MNRAS, 353, 529

Hivon, E., Górski, K. M., Netterfield, C. B., et al. 2002, MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets, ApJ, 567, 2

Hoffmann, K., Bel, J., Gaztañaga, E., et al. 2015, Measuring the growth of matter fluctuations with third-order galaxy correlations, MNRAS, 447, 1724

Howlett, C., Lewis, A., Hall, A., & Challinor, A. 2012, CMB power spectrum parameter degeneracies in the era of precision cosmology, J. Cosm.&Astrop. Phys., 2012, 027

Hu, T., Khaire, V., Hennawi, J. F., et al. 2022, Measuring the thermal and ionization state of the low-z IGM using likelihood free inference, MNRAS, 515, 2188

Hu, W. 2000, Weak lensing of the CMB: A harmonic approach, Phys. Rev. D, 62, 043007

Hu, W., Barkana, R., & Gruzinov, A. 2000, Fuzzy Cold Dark Matter: The Wave Properties of Ultralight Particles, Phys. Rev. Lett., 85, 1158

Hubble, E. 1929, A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae, Contributions from the Mount Wilson Observatory, 3, 23

—. 1934, The Distribution of Extra-Galactic Nebulae, ApJ, 79, 8

Huff, E. M., & Graves, G. J. 2014, Magnificent Magnification: Exploiting the Other Half of the Lensing Signal, ApJ, 780, L16

Hui, L., Gaztañaga, E., & Loverde, M. 2007, Anisotropic magnification distortion of the 3D galaxy correlation. I. Real space, Phys. Rev. D, 76, 103502

Jarvis, M., Bernstein, G., & Jain, B. 2004, The skewness of the aperture mass statistic, MNRAS, 352, 338

Jaynes, E. T. 2003, Probability theory: The logic of science (Cambridge university press)

Jeans, J. H. 1902, The Stability of a Spherical Nebula, Philosophical Transactions of the Royal Society of London Series A, 199, 1

Jeffrey, N., Alsing, J., & Lanusse, F. 2021, Likelihood-free inference with neural compression of DES SV weak lensing map statistics, MNRAS, 501, 954

Jenkins, A., Frenk, C. S., Pearce, F. R., et al. 1998, Evolution of Structure in Cold Dark Matter Universes, ApJ, 499, 20

Jing, Y. P. 1998, Accurate Fitting Formula for the Two-Point Correlation Function of Dark Matter Halos, ApJ, 503, L9

Joachimi, B., & Bridle, S. L. 2010, Simultaneous measurement of cosmology and intrinsic alignments using joint cosmic shear and galaxy number density correlations, A&A, 523, A1

Joachimi, B., Mandelbaum, R., Abdalla, F. B., & Bridle, S. L. 2011a, Constraints on intrinsic alignment contamination of weak lensing surveys using the MegaZ-LRG sample, A&A, 527, A26

Joachimi, B., Taylor, A. N., & Kiessling, A. 2011b, Cosmological information in Gaussianized weak lensing signals, MNRAS, 418, 145

Joachimi, B., Lin, C. A., Asgari, M., et al. 2021, KiDS-1000 methodology: Modelling and inference for joint weak gravitational lensing and spectroscopic galaxy clustering analysis, A&A, 646, A129

Johnston, H., Wright, A. H., Joachimi, B., et al. 2021, Organised randoms: Learning and correcting for systematic galaxy clustering patterns in KiDS using self-organising maps, A&A, 648, A98

Joudaki, S., Blake, C., Johnson, A., et al. 2018, KiDS-450 + 2dFLenS: Cosmological parameter constraints from weak gravitational lensing tomography and overlapping redshift-space galaxy clustering, MNRAS, 474, 4894

Kaaret, P., Feng, H., & Roberts, T. P. 2017, Ultraluminous X-Ray Sources, ARA&A, 55, 303

Kaiser, N. 1992, Weak Gravitational Lensing of Distant Galaxies, ApJ, 388, 272

—. 1998, Weak Lensing and Cosmology, ApJ, 498, 26

Kaiser, N., Wilson, G., & Luppino, G. A. 2000, Large-Scale Cosmic Shear Measurements, arXiv e-prints, astro

Kayo, I., Taruya, A., & Suto, Y. 2001, Probability Distribution Function of Cosmological Density Fluctuations from a Gaussian Initial Condition: Comparison of One-Point and Two-Point Lognormal Model Predictions with N-Body Simulations, ApJ, 561, 22

Kiessling, A., Cacciato, M., Joachimi, B., et al. 2015, Galaxy Alignments: Theory, Modelling & Simulations, Space Sci. Rev., 193, 67

Kilbinger, M. 2015, Cosmology with cosmic shear observations: a review, Reports on Progress in Physics, 78, 086901

Kilbinger, M., Heymans, C., Asgari, M., et al. 2017, Precision calculations of the cosmic shear power spectrum projection, MNRAS, 472, 2126

King, L., & Schneider, P. 2002, Suppressing the contribution of intrinsic galaxy alignments to the shear two-point correlation function, A&A, 396, 411

Kirk, D., Brown, M. L., Hoekstra, H., et al. 2015, Galaxy Alignments: Observations and Impact on Cosmology, Space Sci. Rev., 193, 139

Kitching, T., Balan, S., Bernstein, G., et al. 2010, Gravitational Lensing Accuracy Testing 2010 (GREAT10) Challenge Handbook, arXiv e-prints, arXiv:1009.0779

Kitching, T. D., Alsing, J., Heavens, A. F., et al. 2017, The limits of cosmic shear, MNRAS, 469, 2737

Kitching, T. D., & Heavens, A. F. 2017, Unequal-time correlators for cosmology, Phys. Rev. D, 95, 063522

Kitching, T. D., Paykari, P., Hoekstra, H., & Cropper, M. 2019, Propagating Residual Biases in Cosmic Shear Power Spectra, The Open Journal of Astrophysics, 2, 5

Kodama, H., & Sasaki, M. 1984, Cosmological Perturbation Theory, Progress of Theoretical Physics Supplement, 78, 1

Kohonen, T. 1990, The self-organizing map, Proceedings of the IEEE, 78, 1464

Krause, E., Eifler, T., & Blazek, J. 2016, The impact of intrinsic alignment on current and future cosmic shear surveys, MNRAS, 456, 207

Kuijken, K. 2011, OmegaCAM: ESO's Newest Imager, The Messenger, 146, 8

Kuijken, K., Heymans, C., Hildebrandt, H., et al. 2015, Gravitational lensing analysis of the Kilo-Degree Survey, MNRAS, 454, 3500

Kuijken, K., Heymans, C., Dvornik, A., et al. 2019, The fourth data release of the Kilo-Degree Survey: ugri imaging and nine-band optical-IR photometry over 1000 square degrees, A&A, 625, A2

Landy, S. D., & Szalay, A. S. 1993, Bias and Variance of Angular Correlation Functions, ApJ, 412, 64

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, Euclid Definition Study Report, arXiv e-prints, arXiv:1110.3193

Le Brun, A. M. C., McCarthy, I. G., Schaye, J., & Ponman, T. J. 2014, Towards a realistic population of simulated galaxy groups and clusters, MNRAS, 441, 1270

Leclercq, F. 2018, Bayesian optimization for likelihood-free cosmological inference, Phys. Rev. D, 98, 063511

Leclercq, F., & Heavens, A. 2021, On the accuracy and precision of correlation functions and field-level inference in cosmology, MNRAS, 506, L85

Legin, R., Hezaveh, Y., Perreault Levasseur, L., & Wandelt, B. 2021, Simulation-Based Inference of Strong Gravitational Lensing Parameters, arXiv e-prints, arXiv:2112.05278

Leistedt, B., Peiris, H. V., Mortlock, D. J., Benoit-Lévy, A., & Pontzen, A. 2013, Estimating the large-scale angular power spectrum in the presence of systematics: a case study of Sloan Digital Sky Survey quasars, MNRAS, 435, 1857

Lemaître, G. 1927, Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques, Annales de la Société Scientifique de Bruxelles, 47, 49

Lemos, P., Challinor, A., & Efstathiou, G. 2017, The effect of Limber and flat-sky approximations on galaxy weak lensing, J. Cosm.&Astrop. Phys., 2017, 014

Lemos, P., Coogan, A., Hezaveh, Y., & Perreault-Levasseur, L. 2023a, Sampling-Based Accuracy Testing of Posterior Estimators for General Inference, arXiv e-prints, arXiv:2302.03026

Lemos, P., Cranmer, M., Abidi, M., et al. 2023b, Robust simulation-based inference in cosmology with Bayesian neural networks, Machine Learning: Science and Technology, 4, 01LT01

Lemos, P., Jeffrey, N., Whiteway, L., et al. 2021, Sum of the masses of the Milky Way and M31: A likelihood-free inference approach, Phys. Rev. D, 103, 023009

Lesgourgues, J. 2011, The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview, arXiv e-prints, arXiv:1104.2932

Lesgourgues, J., Mangano, G., Miele, G., & Pastor, S. 2013, Neutrino cosmology (Cambridge University Press)

Levi, M., Bebek, C., Beers, T., et al. 2013, The DESI Experiment, a whitepaper for Snowmass 2013, arXiv e-prints, arXiv:1308.0847

Levin, D. 1996, Fast integration of rapidly oscillatory functions, Journal of Computational and Applied Mathematics, 67, 95

Lewis, A., & Bridle, S. 2002, Cosmological parameters from CMB and other data: A Monte Carlo approach, Phys. Rev. D, 66, 103511

Lewis, A., & Challinor, A. 2002, Evolution of cosmological dark matter perturbations, Phys. Rev. D, 66, 023531

Lewis, A., Challinor, A., & Lasenby, A. 2000, Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models, ApJ, 538, 473

Li, S.-S., Kuijken, K., Hoekstra, H., et al. 2022, KiDS-Legacy calibration: unifying shear and redshift calibration with the SKiLLS multi-band image simulations, arXiv e-prints, arXiv:2210.07163

Li, X., Zhang, T., Sugiyama, S., et al. 2023, Hyper Suprime-Cam Year 3 Results: Cosmology from Cosmic Shear Two-point Correlation Functions, arXiv e-prints, arXiv:2304.00702

Liddle, A. R., & Lyth, D. H. 2000, Cosmological Inflation and Large-Scale Structure (Cambridge university press)

Lifshitz, E. M. 1946, On the gravitational stability of the expanding universe, Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki, 16, 587

Lifshitz, E. M., & Khalatnikov, I. M. 1963, Investigations in relativistic cosmology†, Advances in Physics, 12, 185

Lima, M., Cunha, C. E., Oyaizu, H., et al. 2008, Estimating the redshift distribution of photometric galaxy samples, MNRAS, 390, 118

Limber, D. N. 1953, The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field., ApJ, 117, 134

Lin, C.-A., & Kilbinger, M. 2015, A new model to predict weak-lensing peak counts. II. Parameter constraint strategies, A&A, 583, A70

Lin, C.-A., Kilbinger, M., & Pires, S. 2016, A new model to predict weak-lensing peak counts. III. Filtering technique comparisons, A&A, 593, A88

Lin, C.-H., Harnois-Déraps, J., Eifler, T., et al. 2020, Non-Gaussianity in the weak lensing correlation function likelihood - implications for cosmological parameter biases, MNRAS, 499, 2977

Lin, K., von Wietersheim-Kramsta, M., Joachimi, B., & Feeney, S. 2023, A simulation-based inference pipeline for cosmic shear with the Kilo-Degree Survey, MNRAS, 524, 6167

Linde, A. D. 1982, A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems, Physics Letters B, 108, 389

Linke, L., Burger, P. A., Heydenreich, S., Porth, L., & Schneider, P. 2023, What is the super-sample covariance? A fresh perspective for second-order shear statistics, arXiv e-prints, arXiv:2302.12277

Longair, M. S. 1989, in Evolution of Galaxies: Astronomical Observations, ed. I. Appenzeller, H. J. Habing, & P. Lena, Vol. 333 (Springer Berlin Heidelberg), 1

Loureiro, A., Whittaker, L., Spurio Mancini, A., et al. 2021, KiDS & Euclid: Cosmological implications of a pseudo angular power spectrum analysis of KiDS-1000 cosmic shear tomography, arXiv e-prints, arXiv:2110.06947

LoVerde, M., & Afshordi, N. 2008, Extended Limber approximation, Phys. Rev. D, 78, 123506

LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, LSST Science Book, Version 2.0, arXiv e-prints, arXiv:0912.0201

Ma, C.-P., & Fry, J. N. 2000, Deriving the Nonlinear Cosmological Power Spectrum and Bispectrum from Analytic Dark Matter Halo Profiles and Mass Functions, ApJ, 543, 503

Madhavacheril, M. S., Qu, F. J., Sherwin, B. D., et al. 2023, The Atacama Cosmology Telescope: DR6 Gravitational Lensing Map and Cosmological Parameters, arXiv e-prints, arXiv:2304.05203

Mahony, C., Fortuna, M. C., Joachimi, B., et al. 2022, Forecasting the potential of weak lensing magnification to enhance LSST large-scale structure analyses, MNRAS, 513, 1210

Mancini, A. S., Docherty, M., Price, M., & McEwen, J. 2022, Bayesian model comparison for simulation-based inference, arXiv preprint arXiv:2207.04037

Mandelbaum, R. 2018, Weak Lensing for Precision Cosmology, ARA&A, 56, 393

Mann, R. 2010, An introduction to particle physics and the standard model (Taylor & Francis)

Martizzi, D., Mohammed, I., Teyssier, R., & Moore, B. 2014, The biasing of baryons on the cluster mass function and cosmological parameter estimation, MNRAS, 440, 2290

Massey, R., Heymans, C., Bergé, J., et al. 2007, The Shear Testing Programme 2: Factors affecting high-precision weak-lensing analyses, MNRAS, 376, 13

Mather, J. C., Cheng, E. S., Eplee, R. E., J., et al. 1990, A Preliminary Measurement of the Cosmic Microwave Background Spectrum by the Cosmic Background Explorer (COBE) Satellite, ApJ, 354, L37

Matsumura, T., Akiba, Y., Borrill, J., et al. 2014, Mission Design of LiteBIRD, Journal of Low Temperature Physics, 176, 733

McCarthy, I. G., Schaye, J., Bird, S., & Le Brun, A. M. C. 2017, The BAHAMAS project: calibrated hydrodynamical simulations for large-scale structure cosmology, MNRAS, 465, 2936

McFarland, J. P., Verdoes-Kleijn, G., Sikkema, G., et al. 2013, The Astro-WISE optical image pipeline. Development and implementation, Experimental Astronomy, 35, 45

Mead, A. J., Brieden, S., Tröster, T., & Heymans, C. 2021, HMCODE-2020: improved modelling of non-linear cosmological power spectra with baryonic feedback, MNRAS, 502, 1401

Mead, A. J., Heymans, C., Lombriser, L., et al. 2016, Accurate halo-model matter power spectra with dark energy, massive neutrinos and modified gravitational forces, MNRAS, 459, 1468

Mead, A. J., Peacock, J. A., Heymans, C., Joudaki, S., & Heavens, A. F. 2015, An accurate halo model for fitting non-linear cosmological power spectra and baryonic feedback models, MNRAS, 454, 1958

Miller, L., Kitching, T. D., Heymans, C., Heavens, A. F., & van Waerbeke, L. 2007, Bayesian galaxy shape measurement for weak lensing surveys - I. Methodology and a fast-fitting algorithm, MNRAS, 382, 315

Miller, L., Heymans, C., Kitching, T. D., et al. 2013, Bayesian galaxy shape measurement for weak lensing surveys - III. Application to the Canada-France-Hawaii Telescope Lensing Survey, MNRAS, 429, 2858

Mishra-Sharma, S., & Cranmer, K. 2022, Neural simulation-based inference approach for characterizing the Galactic Center γ -ray excess, Phys. Rev. D, 105, 063017

Miyatake, H., Sugiyama, S., Takada, M., et al. 2023, Hyper Suprime-Cam Year 3 Results: Cosmology from Galaxy Clustering and Weak Lensing with HSC and SDSS using the Emulator Based Halo Model, arXiv e-prints, arXiv:2304.00704

Mo, H. J., & White, S. D. M. 1996, An analytic model for the spatial clustering of dark matter haloes, MNRAS, 282, 347

Mukhanov, V. F., Feldman, H. A., & Brandenberger, R. H. 1992, Theory of cosmological perturbations, Phys. Rep., 215, 203

Nagamine, K. 2018, The Encyclopedia of Cosmology. Volume 2: Numerical Simulations in Cosmology (World Scientific Publishing Co. Pte. Ltd), doi:10.1142/9496-vol2

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, A Universal Density Profile from Hierarchical Clustering, ApJ, 490, 493

Nicola, A., García-García, C., Alonso, D., et al. 2021, Cosmic shear power spectra in practice, J. Cosm.&Astrop. Phys., 2021, 067

Papamakarios, G., & Murray, I. 2016, Fast ε-free Inference of Simulation Models with Bayesian Conditional Density Estimation, arXiv e-prints, arXiv:1605.06376

Papamakarios, G., Pavlakou, T., & Murray, I. 2017, in Advances in Neural Information Processing Systems, ed. I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett, Vol. 30 (Curran Associates, Inc.)

Particle Data Group, Zyla, P. A., Barnett, R. M., et al. 2020, Review of Particle Physics, Progress of Theoretical and Experimental Physics, 2020, 083C01

Peacock, J. A., & Smith, R. E. 2000, Halo occupation numbers and galaxy bias, MNRAS, 318, 1144

Pearson, K. 1901, LIII. On lines and planes of closest fit to systems of points in space, The London, Edinburgh, and Dublin philosophical magazine and journal of science, 2, 559

Peebles, P. J. E. 1973, Statistical Analysis of Catalogs of Extragalactic Objects. I. Theory, ApJ, 185, 413

—. 1982, Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations, ApJ, 263, L1

Pen, U.-L. 1998, Reconstructing Nonlinear Stochastic Bias from Velocity Space Distortions, ApJ, 504, 601

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, Measurements of Ω and Λ from 42 High-Redshift Supernovae, ApJ, 517, 565

Petri, A., Haiman, Z., & May, M. 2017, Validity of the Born approximation for beyond Gaussian weak lensing observables, Phys. Rev. D, 95, 123503

Pillepich, A., Porciani, C., & Hahn, O. 2010, Halo mass function and scale-dependent bias from N-body simulations with non-Gaussian initial conditions, MNRAS, 402, 191

Piras, D., Joachimi, B., & Villaescusa-Navarro, F. 2023, Fast and realistic large-scale structure from machine-learning-augmented random field simulations, MNRAS, 520, 668

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, Planck 2018 results. VI. Cosmological parameters, A&A, 641, A6

Porqueres, N., Heavens, A., Mortlock, D., & Lavaux, G. 2022, Lifting weak lensing degeneracies with a field-based likelihood, MNRAS, 509, 3194

Prat, J., Sánchez, C., Fang, Y., et al. 2018, Dark Energy Survey year 1 results: Galaxy-galaxy lensing, Phys. Rev. D, 98, 042005

Pritchard, J. K., Seielstad, M. T., Perez-Lezaun, A., & Feldman, M. W. 1999, Population growth of human Y chromosomes: a study of Y chromosome microsatellites., Molecular biology and evolution, 16, 1791

Qu, F. J., Sherwin, B. D., Madhavacheril, M. S., et al. 2023, The Atacama Cosmology Telescope: A Measurement of the DR6 CMB Lensing Power Spectrum and its Implications for Structure Growth, arXiv e-prints, arXiv:2304.05202

Reid, B., Ho, S., Padmanabhan, N., et al. 2016, SDSS-III Baryon Oscillation Spectroscopic Survey Data Release 12: galaxy target selection and large-scale structure catalogues, MNRAS, 455, 1553

Reinecke, M. 2011, Libpsht - algorithms for efficient spherical harmonic transforms, A&A, 526, A108

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant, AJ, 116, 1009

Riess, A. G., Yuan, W., Macri, L. M., et al. 2022, A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km s^{-1} Mpc^{-1} Uncertainty from the Hubble Space Telescope and the SH0ES Team, ApJ, 934, L7

Rivinius, T., Baade, D., Hadrava, P., Heida, M., & Klement, R. 2020, A naked-eye triple system with a nonaccreting black hole in the inner binary, A&A, 637, L3

Robert, C. P., et al. 2007, The Bayesian choice: from decision-theoretic foundations to computational implementation, Vol. 2 (Springer)

Robertson, H. P. 1936, Kinematics and World-Structure III., ApJ, 83, 257

Rozo, E., Rykoff, E. S., Abate, A., et al. 2016, redMaGiC: selecting luminous red galaxies from the DES Science Verification data, MNRAS, 461, 1431

Rubin, D. B. 1984, Bayesianly justifiable and relevant frequency calculations for the applied statistician, The Annals of Statistics, 1151

Rubin, V. C., & Ford, W. Kent, J. 1970, Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions, ApJ, 159, 379

Sahni, V., & Coles, P. 1995, Approximation methods for non-linear gravitational clustering, Phys. Rep., 262, 1

Sánchez, A. G., Scoccimarro, R., Crocce, M., et al. 2017, The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmological implications of the configuration-space clustering wedges, MNRAS, 464, 1640

Scharf, C. A., & Lahav, O. 1993, Spherical Harmonic Analysis of the 2-JANSKY IRAS Galaxy Redshift Survey, MNRAS, 264, 439

Schaye, J., Dalla Vecchia, C., Booth, C. M., et al. 2010, The physics driving the cosmic star formation history, MNRAS, 402, 1536

Schechter, P. 1976, An analytic expression for the luminosity function for galaxies., ApJ, 203, 297

Schmidt, F., Leauthaud, A., Massey, R., et al. 2012, A Detection of Weak-lensing Magnification Using Galaxy Sizes and Magnitudes, ApJ, 744, L22

Schneider, A., Giri, S. K., Amodeo, S., & Refregier, A. 2022, Constraining baryonic feedback and cosmology with weak-lensing, X-ray, and kinematic Sunyaev-Zeldovich observations, MNRAS, 514, 3802

Schneider, A., & Teyssier, R. 2015, A new method to quantify the effects of baryons on the matter power spectrum, J. Cosm.&Astrop. Phys., 2015, 049

Schneider, A., Teyssier, R., Stadel, J., et al. 2019, Quantifying baryon effects on the matter power spectrum and the weak lensing shear correlation, J. Cosm.&Astrop. Phys., 2019, 020

Schneider, P. 2003, in 14th Canary Islands Winter School of Astrophysics: Dark Matter and Dark Energy in the Universe

Schneider, P. 2005, Weak Gravitational Lensing, arXiv e-prints, astro

Schneider, P. 2006, in Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro, ed. G. Meylan, P. Jetzer, P. North, P. Schneider, C. S. Kochanek, & J. Wambsganss, 269–451

—. 2016, Generalized shear-ratio tests: A new relation between cosmological distances, and a diagnostic for a redshift-dependent multiplicative bias in shear measurements, A&A, 592, L6

Schneider, P., & Hartlap, J. 2009, Constrained correlation functions, A&A, 504, 705

Schneider, P., van Waerbeke, L., Kilbinger, M., & Mellier, Y. 2002, Analysis of two-point statistics of cosmic shear. I. Estimators and covariances, A&A, 396, 1

Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample, ApJ, 859, 101

Scranton, R., Ménard, B., Richards, G. T., et al. 2005, Detection of Cosmic Magnification with the Sloan Digital Sky Survey, ApJ, 633, 589

Secco, L. F., Samuroff, S., Krause, E., et al. 2022, Dark Energy Survey Year 3 results: Cosmology from cosmic shear and robustness to modeling uncertainty, Phys. Rev. D, 105, 023515

Seitz, S., & Schneider, P. 1996, Cluster lens reconstruction using only observed local data: an improved finite-field inversion technique., A&A, 305, 383

Seljak, U. 2000, Analytic model for galaxy and dark matter clustering, MNRAS, 318, 203

Sellentin, E., & Heavens, A. F. 2018, On the insufficiency of arbitrarily precise covariance matrices: non-Gaussian weak-lensing likelihoods, MNRAS, 473, 2355

Sellentin, E., Heymans, C., & Harnois-Déraps, J. 2018, The skewed weak lensing likelihood: why biases arise, despite data and theory being sound, MNRAS, 477, 4879

Sheth, R. K., Mo, H. J., & Tormen, G. 2001, Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes, MNRAS, 323, 1

Shirasaki, M., Hamana, T., Takada, M., Takahashi, R., & Miyatake, H. 2019, Mock galaxy shape catalogs in the Subaru Hyper Suprime-Cam Survey, arXiv e-prints, arXiv:1901.09488

Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, Stable clustering, the halo model and non-linear cosmological power spectra, MNRAS, 341, 1311

Speagle, J. S., Leauthaud, A., Huang, S., et al. 2019, Galaxy-Galaxy lensing in HSC: Validation tests and the impact of heterogeneous spectroscopic training sets, MNRAS, 490, 5658

Spergel, D., Gehrels, N., Baltay, C., et al. 2015, Wide-Field InfrarRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report, arXiv e-prints, arXiv:1503.03757

Spurio Mancini, A., Reischke, R., Pettorino, V., Schäfer, B. M., & Zumalacárregui, M. 2018a, Testing (modified) gravity with 3D and tomographic cosmic shear, MNRAS, 480, 3725

Spurio Mancini, A., Taylor, P. L., Reischke, R., et al. 2018b, 3D cosmic shear: Numerical challenges, 3D lensing random fields generation, and Minkowski functionals for cosmological inference, Phys. Rev. D, 98, 103507

Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, Sloan Digital Sky Survey: Early Data Release, AJ, 123, 485

Sugiyama, N. 1989, Gauge Invariant Cosmological Perturbations with Cold Dark Matter, Progress of Theoretical Physics, 81, 1021

Sugiyama, S., Takada, M., Miyatake, H., et al. 2022, HSC Year 1 cosmology results with the minimal bias method: HSC ×BOSS galaxy-galaxy weak lensing and BOSS galaxy clustering, Phys. Rev. D, 105, 123537

Takada, M., & Hu, W. 2013, Power spectrum super-sample covariance, Phys. Rev. D, 87, 123504

Takahashi, R., Oguri, M., Sato, M., & Hamana, T. 2011, Probability Distribution Functions of Cosmological Lensing: Convergence, Shear, and Magnification, ApJ, 742, 15

Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. 2012, Revising the Halofit Model for the Nonlinear Matter Power Spectrum, ApJ, 761, 152

Taruya, A., Takada, M., Hamana, T., Kayo, I., & Futamase, T. 2002, Lognormal Property of Weak-Lensing Fields, ApJ, 571, 638

Taylor, P. L., Kitching, T. D., Alsing, J., et al. 2019, Cosmic shear: Inference from forward models, Phys. Rev. D, 100, 023519

Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, Karhunen-Loève Eigenvalue Problems in Cosmology: How Should We Tackle Large Data Sets?, ApJ, 480, 22

Tessore, N., Loureiro, A., Joachimi, B., von Wietersheim-Kramsta, M., & Jeffrey, N. 2023, GLASS: Generator for Large Scale Structure, The Open Journal of Astrophysics, 6, 11

Thiele, L., Duncan, C. A. J., & Alonso, D. 2020, Disentangling magnification in combined shear-clustering analyses, MNRAS, 491, 1746

Thornton, R. J., Ade, P. A. R., Aiola, S., et al. 2016, The Atacama Cosmology Telescope: The Polarization-sensitive ACTPol Instrument, ApJS, 227, 21

Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, The Large-scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests, ApJ, 724, 878

Tinker, J. L., Weinberg, D. H., Zheng, Z., & Zehavi, I. 2005, On the Mass-to-Light Ratio of Large-Scale Structure, ApJ, 631, 41

Tröster, T., Sánchez, A. G., Asgari, M., et al. 2020, Cosmology from large-scale structure. Constraining ΛCDM with BOSS, A&A, 633, L10

Tröster, T., Asgari, M., Blake, C., et al. 2021, KiDS-1000 Cosmology: Constraints beyond flat ΛCDM, A&A, 649, A88

Trotta, R. 2008, Bayes in the sky: Bayesian inference and model selection in cosmology, Contemporary Physics, 49, 71

Troxel, M. A., & Ishak, M. 2015, The intrinsic alignment of galaxies and its impact on weak gravitational lensing in an era of precision cosmology, Phys. Rep., 558, 1

Tyson, J. A., Valdes, F., Jarvis, J. F., & Mills, A. P., J. 1984, Galaxy mass distribution from gravitational light deflection, ApJ, 281, L59

Unruh, S., Schneider, P., Hilbert, S., et al. 2020, The importance of magnification effects in galaxy-galaxy lensing, A&A, 638, A96

Upham, R. E., Brown, M. L., & Whittaker, L. 2021, Sufficiency of a Gaussian power spectrum likelihood for accurate cosmology from upcoming weak lensing surveys, MNRAS, 503, 1999

Vakili, M., Hoekstra, H., Bilicki, M., et al. 2020, Clustering of red-sequence galaxies in the fourth data release ofthe Kilo-Degree Survey, arXiv e-prints, arXiv:2008.13154

Valageas, P., & Nishimichi, T. 2011, Combining perturbation theories with halo models, A&A, 527, A87

van Daalen, M. P., & Schaye, J. 2015, The contributions of matter inside and outside of haloes to the matter power spectrum, MNRAS, 452, 2247

van Daalen, M. P., Schaye, J., Booth, C. M., & Dalla Vecchia, C. 2011, The effects of galaxy formation on the matter power spectrum: a challenge for precision cosmology, MNRAS, 415, 3649

van den Busch, J. L., Hildebrandt, H., Wright, A. H., et al. 2020, Testing KiDS cross-correlation redshifts with simulations, A&A, 642, A200

van den Busch, J. L., Wright, A. H., Hildebrandt, H., et al. 2022, KiDS-1000: Cosmic shear with enhanced redshift calibration, A&A, 664, A170

Van Waerbeke, L., & Mellier, Y. 2003, Gravitational Lensing by Large Scale Structures: A Review, arXiv e-prints, astro

Van Waerbeke, L., Mellier, Y., Erben, T., et al. 2000, Detection of correlated galaxy ellipticities from CFHT data: first evidence for gravitational lensing by large-scale structures, A&A, 358, 30

Viel, M., Becker, G. D., Bolton, J. S., & Haehnelt, M. G. 2013, Warm dark matter as a solution to the small scale crisis: New constraints from high redshift Lyman-α forest data, Phys. Rev. D, 88, 043502

Voss, A. 1880, Zur Theorie der Transformation quadratischer Differentialausdrücke und der Krümmung höherer Mannigfaltigkeiten, Mathematische Annalen, 16, 129

Wadekar, D., & Scoccimarro, R. 2020, Galaxy power spectrum multipoles covariance in perturbation theory, Phys. Rev. D, 102, 123517

Walker, A. G. 1937, On Milne's Theory of World-Structure, Proceedings of the London Mathematical Society, 42, 90

Weinberg, S. 1972, Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity (Wiley)

—. 2008, Cosmology (Oxford University Press)

White, M., & Hu, W. 2000, A New Algorithm for Computing Statistics of Weak Lensing by Large-Scale Structure, ApJ, 537, 1

Wild, V., Peacock, J. A., Lahav, O., et al. 2005, The 2dF Galaxy Redshift Survey: stochastic relative biasing between galaxy populations, MNRAS, 356, 247

Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I., & Bernstein, G. 2000, Detection of weak gravitational lensing distortions of distant galaxies by cosmic dark matter at large scales, Nature, 405, 143

Wright, A. H., Hildebrandt, H., van den Busch, J. L., & Heymans, C. 2020a, Photometric redshift calibration with self-organising maps, A&A, 637, A100

Wright, A. H., Hildebrandt, H., van den Busch, J. L., et al. 2020b, KiDS+VIKING-450: Improved cosmological parameter constraints from redshift calibration with self-organising maps, A&A, 640, L14

Wright, A. H., Hildebrandt, H., Kuijken, K., et al. 2019, KiDS+VIKING-450: A new combined optical and near-infrared dataset for cosmology and astrophysics, A&A, 632, A34

Xavier, H. S., Abdalla, F. B., & Joachimi, B. 2016, Improving lognormal models for cosmological fields, MNRAS, 459, 3693

Zablocki, A., & Dodelson, S. 2016, Extreme data compression for the CMB, Phys. Rev. D, 93, 083525

Zaldarriaga, M., & Seljak, U. 1997, All-sky analysis of polarization in the microwave background, Phys. Rev. D, 55, 1830

Zel'dovich, Y. B. 1970, Gravitational instability: An approximate theory for large density perturbations., A&A, 5, 84

Zieser, B., & Merkel, P. M. 2016, The cross-correlation between 3D cosmic shear and the integrated Sachs-Wolfe effect, MNRAS, 459, 1586

Ziour, R., & Hui, L. 2008, Magnification bias corrections to galaxy-lensing cross-correlations, Phys. Rev. D, 78, 123517

Zuntz, J., Paterno, M., Jennings, E., et al. 2015, CosmoSIS: Modular cosmological parameter estimation, Astronomy and Computing, 12, 45

Zuntz, J., Sheldon, E., Samuroff, S., et al. 2018, Dark Energy Survey Year 1 results: weak lensing shape catalogues, MNRAS, 481, 1149

Zwicky, F. 1933, Die Rotverschiebung von extragalaktischen Nebeln, Helvetica Physica Acta, 6, 110