THE UNIVERSITY
of EDINBURGH

The Non-Linear Universe: The Role of Simulations, Theory & Machine Learning in Weak Lensing Cosmology

Benjamin Martin Giblin

Doctor of Philosophy
The University of Edinburgh
July 2019

Declaration

I declare that no part of this thesis has been submitted for any other degree or professional qualification in this University or elsewhere.

This thesis was composed by myself, and the work contained herein is my own except where explicitly stated otherwise either here or in the text itself. In particular, the text in Section 5.2 was composed by Matteo Cataneo, second author on the paper in which Chapter 5 was published. I reproduce that Section in this thesis so as to maintain the clarity of discussion from that work. Furthermore, the ensemble of random curves designated 'SMURVES' in Section 5.3.2, where generated by Ben Moews, and the optimised Latin hypercube designs discussed in Section 5.3.3 were produced by Ben Moews and Matteo Cataneo.

The entirety of Chapter 3, parts of Chapter 4 and all of Chapter 5 of this thesis have been published in, respectively, Giblin et al. (2018), Harnois-Deraps, Giblin & Joachimi (2019) and Giblin et al. (2019).

(*Benjamin Martin Giblin, July 2019*)

Lay Summary

This thesis contains contributions to research in the arena of *cosmology*, that is, the study of what the Universe is made of, and how it has evolved, on very large scales. Breathtaking progress has been made in this field in the past century, accumulating in a standard model of cosmology, nicknamed "ΛCDM"[1], the predictions of which, are encouragingly consistent with a plethora of different observations taken of our Universe.

The key features of the ΛCDM paradigm are the following. Our Universe was once in a very hot, very dense initial state, some 13.8 billion years ago. Since then, it has been expanding (almost everything is moving away from almost everything else) and is doing so at an ever-increasing rate. Fuelling this acceleration is an unknown energy source, designated *dark energy*, which, on average, makes up around 70% of the energy density (energy per unit of volume) in the cosmos. Of the remaining portion, quite remarkably, only about 4% is contained in the 'ordinary' matter, out of which ourselves, everything in the periodic table, and all human experience is composed. The rest is some equally exotic and mysterious substance, *dark matter*, similar to ordinary matter in that it generates and responds to gravitational fields, but unlike it in all other ways, in that it is apparently completely ambivalent to electromagnetic or nuclear forces.

There are a number of ways, sometimes referred to as *cosmological probes*, via which the details of this model have been worked out. One of these probes is known as *weak gravitational lensing*, and it refers to the small distortions induced in the observed shapes of distant galaxies, caused by the fact their light was deflected slightly on it way to us, by the gravity of intervening matter (both the ordinary and the dark stuff). Those new to this idea are encouraged to imagine viewing a candle flame through a wineglass - the observed image will be distorted due to the bending of the candlelight as it passes through the glass. The amount of distortion, or 'shear', observed in the galaxy shapes, is positively correlated not only with the *amount* of matter occupying the intermediate space, but also how *concentrated* it is. Measurements of galaxy shapes, thus highlight gravitational 'hotspots', providing means to map out and weigh the large-scale structure of the Universe.

One slightly irksome finding of recent times, is that the amount of matter

[1]Pronounced 'Lambda'-CDM.

clustering measured by teams of scientists performing weak lensing observations, is considerably lower than the amount measured by scientists using a different cosmological probe, the *cosmic microwave background*. This is the name given to the low-energy sea of microwave radiation, left over from the early Universe, pervading throughout the cosmos. In this method, the tiny temperature fluctuations in the radiation are used to map the gravitational hotspots, thus informing of the matter clustering. The disagreement between these probes implies that either mistakes were made in the analyses, or, more radically, that the standard model is missing a crucial piece of physics to explain how the large-scale matter distributions of the Universe have evolved over time.

With this cosmic conundrum in mind, in this thesis, I have designed new methods to analyse and measure weak gravitational lensing. This work can be summarised in the following three projects:

- I develop a new technique called "clipping", which consists of selectively removing certain regions from weak lensing images, in order to enhance the information that can be extracted about the matter clustering from the remaining regions of the image. I test this method with data taken by a large weak lensing collaboration, the Kilo-Degree Survey, and find that it indeed leads to a 17% improvement in the precision with which the matter clustering can be measured.

- I apply *machine learning* algorithms to state-of-the-art simulations of the Universe. These computer codes study how the weak lensing measurements obtained from the simulations, depend on the simulation settings such as the amount of matter and how concentrated it is. This means that the algorithms learn to predict the weak lensing signal for *arbitrary* values of these settings, such that one does not have to compute more simulations, which would be very time-consuming.

- Finally, using new advancements in theoretical modelling, I develop a machine learning code capable of predicting the amount of matter clustering in hypothetical Universes, featuring laws of physics that differ from those of the standard, ΛCDM model. These predictions can be compared to observations of our own Universe, to see which hypothetical Universe best matches.

The methods presented in this thesis have been demonstrated to increase the amount of weak lensing information that can be extracted from data. This will facilitate more probing examination of the next-generation weak lensing observations, and help us to test whether ΛCDM really is the correct model of the Universe.

Abstract

The coherent distortions in the observed shapes of distant galaxies, a consequence of the spacetime curvature induced by the intervening large-scale structure of the Universe, is an abundant reservoir of cosmological information. Via this phenomenon of *weak gravitational lensing*, and a number of other independent cosmological probes, the parameters of the standard model, ΛCDM, have been inferred, now with uncertainties approaching the per cent level. In this era of *precision cosmology*, however, we face new challenges. Elements of tension have emerged between the measurements of the cosmological parameters from low- and high-redshift probes, seemingly implying either a failure to account for all relevant systematics, or perhaps even an incompleteness in the ΛCDM paradigm.

In this thesis, I develop novel methodologies in weak lensing, to enhance the cosmological information extracted from current and future data sets. In this pursuit, I adopt a three-pronged approach, combining new advances in theoretical modelling, cutting-edge numerical simulations and recent developments in machine learning. Applying this trinity of techniques to three distinct bodies of research, described below, I construct new routes to improving the constraining power of this cosmological probe.

A notable shortfall of the standard two-point statistics conventionally used in weak lensing, is their inability to capture all of the information contained in the non-linear cosmological fields of the real Universe. In answer to this problem, I develop the use of "clipping" transformations, which suppress the signal from the highest density regions observed. I present the first "clipped" cosmic shear measurement using data from the Kilo-Degree Survey (KiDS-450), and employ a suite of numerical simulations to calibrate and explore the cosmological dependence of this novel statistic. I show that these transformations improve constraints on $S_8 = \sigma_8(\Omega_\mathrm{m}/0.3)^{0.5}$, where Ω_m is the mass energy density and σ_8 is the amplitude of matter density fluctuations, when used in combination with conventional, "unclipped" two-point statistics, by 17% in the case of the KiDS-450 data.

Clipping is but one member of the non-Gaussian statistics family, which have great potential for improving cosmological constraints, but are reliant both on numerical simulations, and a robust means to interpolate the statistics measured in the simulations to arbitrary cosmologies for comparison to the data. In this thesis, I develop a general framework to facilitate this, by designing and

training a Gaussian process emulator, employing Bayesian supervised machine learning, on the state-of-the-art *cosmo*-SLICS suite, consisting of 26 different wCDM cosmologies. I demonstrate that this emulator achieves per cent level interpolation accuracy, in turn yielding unprecedented precision in the estimation of non-Gaussian statistics. I subsequently show how the *cosmo*-SLICS emulator might be employed within a likelihood analysis to constrain the cosmology of next-generation lensing data using these non-standard statistical probes. Taking clipped shear correlation functions as an example, I find that the low levels of noise present in the *cosmo*-SLICS emulator's predictions facilitate improved constraints on cosmological parameters when the clipped and unclipped two-point probes are combined, not only for S_8, but also for Ω_m, and the Hubble and dark energy equation of state parameters, by 18%-26%.

Finally, I combine the emulator approach with recent progress in theoretical modelling, to create a comprehensive framework for accurately predicting the non-linear matter power spectrum in arbitrary models of cosmology. This requires only a suite of vanilla ΛCDM N-body simulations with their initial conditions suitably tailored, such that the late-time non-linear power spectrum deviates from the standard model within a range permitted by observational constraints. These "pseudo" power spectra serve as the training set for the emulator, the predictions from which can be rescaled by reaction functions, analytically computed from the halo model, to obtain per cent level accurate non-linear predictions in a broad-class of beyond-ΛCDM cosmologies. In this proof-of-concept analysis, with a HALOFIT training set substituting the simulation suite, I find that the emulator recovers the power spectra corresponding to $f(R)$ gravity, massive neutrino cosmologies, combinations thereof, and even artificially generated departures from the ΛCDM prediction, with errors $\lesssim 1\%$ deep in the highly non-linear regime. This work thus demonstrates a flexible and powerful method to not only test the validity of the standard model in the non-linear regime with next-generation cosmological data, but to also limit our reliance on costly numerical simulations in the future.

Acknowledgements

As is often case in my scientific writing, I am typing this on a train. Specifically, and coincidentally, this train is currently passing through the urban pastures of Selly Oak, Birmingham, where I cut my teeth as an undergraduate in astrophysics. It is strange to think, that as I began the final year of my Master's here, I had no knowledge of Professor Catherine Heymans, the person who, more so than anyone, would impact my life in academic research. To say that I have valued Catherine's support as my PhD supervisor, is an understatement of cosmological proportions. Catherine has been known to say, *"A happy astronomer is a good astronomer"*, and whilst imposter syndrome prevents any PhD student from speaking convincingly of their competence, I can say categorically that I have been very happy to work with Catherine these past four years. Her guidance, insight, experience, kindness and morale-raising, wall-penetrating laughter, have exceeded everything I could have hoped for in a supervisor.

I consider myself unjustly lucky to have had the support of not just one, but *two* excellent advisors in my doctorate. With unwavering optimism and endless genuine interest in my work, Dr Joachim Harnois-Déraps has assisted me in two challenging Chapters of this thesis. Though Joachim's cosmological measurements may be simulated, his contributions to my work have been very real, and greatly valued.

They say it takes a village to raise a child. I believe it true also, of collaborations and PhD students. Therefore, I would like to extend my wholehearted gratitude to Matteo Cataneo, Angus Wright, Marika Asgari, Hendrik Hildebrandt, Konrad Kuijken, Fergus Simpson and everyone in the Kilo-Degree Survey for everything you have taught me. You have been fantastic collaborators and even better friends. Sincerest thanks also goes to my thesis examiners, Adam Amara and Bob Mann, for providing extremely useful comments and placing me in the running for Edinburgh's longest ever viva.

To my academic siblings, much is owed. My thanks to Vasiliy Demchenko, for the coffee, the beer, the skateboarding, the hockey, the yoga, the mindfulness classes, and the showstopping pairs of socks. To Alexandra Amon, Naomi Robertson, Clémence Fontanive, Raphaël Errani and Tom Kemp - I came to Edinburgh for the doctorate, but I stayed for your friendship.

For my bandmates in *The Darjeeling Feeling*, Joe Kennedy, Daniel Eastwood and Andy Lombard, I have boundless gratitude. You rekindled my love of playing music and gave me a much-needed creative outlet throughout my PhD. Surely the tricky second album will be our big break, and we can finally can all this cosmology nonsense, eh? Thanks as well to my teammates in the *Edinburgh Phoenix Ice Hockey Club* and their 100 mph slapshots. For 90 minutes a week, you made it impossible to worry about weak lensing.

To Mum and Dad, for your intelligence and work ethic, and your unwavering support in everything I do. Although you might not know about 'Gaussian process emulators' and 'cosmic shear', there is a piece of you behind every word in this thesis. Your contributions are incalculable. To Alex and Matt, my thanks for teaching me that the wonders of the Universe, pale in comparison to a pint with your brothers.

Most of all, I am thankful to *Dr* Laura Jean Wells, for her everyday love and encouragement. Countless times in this doctorate, both emotionally, and literally, you have picked me up when I have been miles from home with my bike tyres flat. You have attended my public outreach talks, shown interest in every incomprehensible new science result I brought home, and tolerated my borderline-unhealthy love of Bruce Springsteen. Moreover, in seeing you excel as a clinical psychologist, I have been inspired to become a better cosmologist. I look forward to many more years of being disillusioned in each other's favour.

Contents

List of Figures

List of Tables

This could be Rotterdam or anywhere,
Liverpool or Rome.

The Beautiful South

1

Cosmology

The year is 2019. Despite the formation, evolution and constituents of our Universe being the subject of speculation for millennia, miraculously, we find ourselves living in the era of *precision cosmology*, a time where these things are thought to be known to an unprecedented level of detail. One might also say our epoch falls within the *Golden Age of Astronomy*, for our observable Universe currently contains very distant galaxies, accommodating our cosmic enquiries. If indeed the cosmic expansion is accelerating, as observational evidence suggests, then these distant sources will, at some point, forever pass beyond our gaze. In that sense, what a time to be alive!

That is not to say that all questions in cosmology have been answered. Quite the contrary. Such is the reason this field remains a breathing, changing animal, the remaining (potentially substantial) mysteries of which are a privilege to investigate. Nevertheless, in the past 100 years or so, theoretical advances have culminated in a suitable mathematical framework for comprehending the behaviour of the Universe on large scales. Meanwhile, technological progress has facilitated the collection of exquisite observational evidence which has considerably constrained the free parameter space in the theory.

In this introductory Chapter, I outline the basic picture supported by our observations, that is the evolution of our Universe from a *Hot Big Bang*. I present the mathematical formalism and associated cosmological parameters describing this view of reality, which form the basis of much of the work in this thesis. I also briefly describe how model constraints are obtained from various independent observational probes. Throughout this Chapter I adopt *natural units* wherein the speed of light in a vacuum, c, is set to unity. For more in-depth discussion and bibliographies of the standard material covered here, I direct the interested reader to any of the following excellent textbooks and reviews: Liddle (1998); Peacock (1999); Weinberg et al. (2013) and Jones (2017).

1.1 A brief history of the Universe

There are many big questions in cosmology. I will begin with the more simple question of 'what?' before attempting the more complex questions of 'how?' and 'why?' That is to say, I shall begin with a *very* brief discussion of what is the best working model for the evolutionary history of our Universe, before answering how and why theory and observations compel us to champion said model.

Observations by Hubble (1929), showing galaxies receding into the cosmos with velocities proportional to their distance from us, provided the first evidence that our Universe is expanding. This result was of enormous significance, with extrapolation back in time implying a genesis to the Universe with a singularity, termed the 'Big Bang', more than 10 billion years ago.

The Hot Big Bang model sees the early Universe in a radiation-dominated, high-density, high-temperature state. Within the hot bath of radiation there were quantum fluctuations which later seeded gravitational collapse and eventual formation of large-scale structure. Prior to this however, came expansion and cooling, such that the production of subatomic particles became energetically favourable within a few minutes after the Big Bang. With the energy density of radiation falling faster than that of matter, an epoch of *matter-radiation equality* was reached before matter became the dominant energy content in the Universe.

After approximately 250,000 years, the Universe had cooled to the extent that the energetic barrier, which prevented the joining of electrons and protons to form

neutral hydrogen, was lifted. Prior to this epoch of *recombination*[1], the Universe was *opaque*, since photons could travel only short distances without absorbtion or scattering by an electron. A short time afterwards however, at the epoch of *last scattering*, electromagentic radiation propagated uninhibited through a *transparent* Universe. The relic radiation originating from this time – the *cosmic microwave background* (CMB) – was first detected by Penzias & Wilson (1965), and found to be isotropic (the same in all directions). The invocation of the *cosmological principle*, which states we reside in a typical, unremarkable region of the Universe[2], implies that if isotropy is observed from our position, so it must be from all positions. This leads us to the profound conclusion that our Universe is also homogeneous (constant density)[3] and hence, on the largest scales, will appear smooth and mostly featureless.

As is often the case with breakthroughs in cosmology, observations of the CMB simultaneously raised new questions to replace the ones it answered. Particularly puzzling was the similarity of the CMB on different regions of the sky, which had seemingly never been causally connected (i.e. at no epoch had they ever been in close enough proximity that light could travel between them within the available time offered by the Universe). A possible solution to this *horizon problem*, and several others identified in the developing cosmological model, is that the Universe underwent a period of rapid, exponential expansion, known as *inflation*, around 10^{-34}s after the Big Bang.

After the formation of neutral atoms, the large scale density perturbations, created from the quantum fluctuations through the period of inflation, invoked gravitational collapse. This began the formation of galaxies, large-scale cosmic structure and also the first stars, whose nuclear furnaces would generate the first heavy elements, of which we are composed.

[1] A misnomer, given that these subatomic particles had never been close friends at any time before this.

[2] *The Beautiful South*'s alternative articulation of this concept is also given in this Chapter's opening.

[3] This result is obtained by considering observers at two separate points, A and B, both of whom observe an isotropic density field (i.e. one in which the density at a given radial distance is constant). By evaluating the density in increasingly large overlapping shells centred on A and B, one finds that constant density everywhere is required to preserve the isotropy about either point. Global isotropy thus necessitates homogeneity.

1.2 Basics

On the largest scales, gravitation becomes the dominant factor in determining the behaviour of the Universe, owing to the strong and weak nuclear forces operating only on short length scales and the overall charge neutrality eliminating electromagnetic contribution. The General Theory of Relativity (Einstein, 1915) provides the mathematical framework for understanding the evolution of homogeneous matter distributions under these conditions. The relation between the curvature of *spacetime* and the contained energy density therein is expressed via the Einstein Gravitational Field equations,

$$R_{ab} - \frac{1}{2} R g_{ab} = -8\pi G T_{ab} \,, \tag{1.1}$$

where R_{ab} and R are the Ricci tensor and scalar respectively containing combinations of first- and second-order derivatives of the metric tensor, g_{ab}, with respect to coordinates. G is the Newtonian constant of gravitation and T_{ab} is the *energy-momentum* tensor which contains information on the distribution of energy density. The presence of tensors here means that the Field Equations are invariant under general coordinate transformations. These highly non-linear differential equations can be solved analytically in a small number of symmetric cases to give the metric, an important quantity which determines the spacetime 'distance' between events. A spatially homogeneous and isotropic universe[4] for example is described by the Friedmann-Lemaître-Robertson-Walker (FLRW) metric,

$$ds^2 = g_{ab}\, dx^a\, dx^b = dt^2 - R(t)^2 \left[dr^2 + S_k(r) \left(d\theta^2 + \sin^2(\theta) d\phi^2 \right) \right] \,, \tag{1.2}$$

where ds is a spacetime interval between two events which is universally agreed upon by all observers. The effects of gravity mean that either expansion or contraction is the *modus operandi* in such a universe, with the time-dependent scale factor[5], $R(t)$, keeping track of how the distance between two events changes in a time t, measured by observers at rest with respect to the matter in their

[4] I reserve a capital 'U' for our actual Universe, and refer to general universes with the lower case.

[5] The regrettable repetition in R notation here (Ricci scalar and scale factor) is, alas, a consequence of a lack of notation convention in cosmology as a whole. I will endeavour to keep such degeneracies to a minimum in this thesis.

vicinity. r, θ and ϕ are simply the standard spherical polar coordinates with the caveat that r is a dimensionless comoving quantity; observers moving with the expansion/contraction (termed *Hubble flow* in our Universe) do not change r, and physical distances are only given by its product with the scale factor. The term $S(k)$ is dependent on the geometry or curvature of this universe, k, according to

$$S_k(r) = \begin{cases} \sin(r), & (k=1) \\ r, & (k=0) \\ \sinh(r), & (k=-1) \end{cases}, \tag{1.3}$$

where the universe is said to have *closed*, *flat* or *open* geometry in each of these three cases respectively. The relation between k and R is given by the Friedmann Equation,

$$\left(\frac{\dot{R}}{R}\right)^2 \equiv H^2 = \frac{8\pi G}{3}\rho - \frac{k}{R^2}\,, \tag{1.4}$$

where H is the Hubble parameter and ρ is the total energy density. This is the solution to the Einstein Field Equations for the FLRW metric and a perfect fluid (diagonal energy-momentum tensor). Equation 1.4 can be regarded as a statement of energy conservation, with the left-hand side corresponding to kinetic energy, the first term on the right-hand side being potential energy, and the final term on the right-hand side being the total energy, associated with the curvature of the universe. In this sense, k is not a free parameter but determined by the total energy content.

The total energy density is given by the sum of the matter, radiation and vacuum contributions,

$$\rho = \rho_{\rm m} + \rho_{\rm r} + \rho_{\Lambda}. \tag{1.5}$$

As the universe expands or contracts, the change in the energy densities is computed via the covariant conservation of the energy-momentum, $\nabla_a T^{ab} = 0$, giving the conservation law,

$$\dot{\rho}_i + 3H(\rho_i + p_i) = 0\,, \tag{1.6}$$

where p_i is the pressure affecting a given species and $i \in \{\mathrm{m}, \mathrm{r}, \Lambda\}$. The energy density for each species evolves differently in accordance with their *equation of state*,

$$p_i = w_i \rho_i \,, \tag{1.7}$$

where $w_\mathrm{m} = 0$, $w_\mathrm{r} = 1/3$ and $w_\Lambda = -1$. The vacuum component in particular warrants further consideration: we find that equation 1.6 for this species reduces to $\dot{\rho} = 0$, hence the energy density of the vacuum counter-intuitively remains constant throughout expansion and/or contraction. This is a consequence of the peculiar *negative* equation of state: the work done by the vacuum is just sufficient to maintain the energy density at a constant level. The negative sign also means that the vacuum energy engenders a *repulsive gravitational* force. The inferred presence of such a phenomenon in our Universe has earned the nickname *dark energy*, or the *cosmological constant*, Λ, so called as it was first introduced by Einstein as an extra constant of integration term on the left-hand side of his Field Equations (Eq. (1.1)).

Introducing the dimensionless scale factor, a, defined to be unity at the present-time, t_0, via

$$a(t) = \frac{R(t)}{R_0} \tag{1.8}$$

where $R_0 \equiv R(t = t_0)$, solving equation 1.6 for the different species yields

$$\begin{aligned} &\rho_\mathrm{m} \propto a^{-3} \,, \\ &\rho_\mathrm{r} \propto a^{-4} \,, \\ &\rho_\Lambda = \text{constant}. \end{aligned} \tag{1.9}$$

The radiation component dilutes fastest in an expanding universe owing to the proportionality between radiation energy and momentum, which is being diluted in addition to the spatial density of this species. The different dependencies on a reveal that an expanding universe undergoes different phases in its evolution, with radiation dominating at earlier times (when a was very small), followed by matter and finally vacuum energy, with epochs of equality marking the transitions.

Consequently, the Friedmann equation can be rewritten in the form,

$$\left(\frac{\dot{a}}{a}\right)^2 \equiv H^2 = \frac{8\pi G}{3}\left(\frac{\rho_{\mathrm{m}_0}}{a^3} + \frac{\rho_{\mathrm{r}_0}}{a^4} + \rho_{\Lambda_0}\right) - \frac{k}{a^2 R_0^2}. \tag{1.10}$$

Defining a dimensionless energy density parameter per species,

$$\Omega_i(a) = \frac{8\pi G}{3H^2}\rho_i(a) = \frac{\rho_i(a)}{\rho_{\mathrm{cr}}(a)}\,, \tag{1.11}$$

where $\rho_{\mathrm{cr}}(a)$ is known as the *critical energy density*, allows the Friedmann equation to be cast in the convenient form,

$$\left(\frac{\dot{a}}{a}\right)^2 \equiv H^2 = H_0^2\left(\frac{\Omega_{\mathrm{m}_0}}{a^3} + \frac{\Omega_{\mathrm{r}_0}}{a^4} + \Omega_{\Lambda_0}\right) + \frac{1-\Omega_0}{a^2}\,, \tag{1.12}$$

where the total energy density parameter, $\Omega = \Omega_{\mathrm{m}} + \Omega_{\mathrm{r}} + \Omega_\Lambda$, sums to unity in the case of a flat ($k = 0$) universe. Henceforth in this thesis, I set these parameters to their present-day values and neglect the subscript '0' for brevity. In cases where I am considering their values at other cosmic epochs, time dependence will be explicitly indicated with the inclusion of the scale factor dependence, i.e. $\Omega_i(a)$.

Differentiating the Friedmann Equation produces the acceleration equation,

$$\frac{\ddot{a}}{a} = -\frac{H_0^2}{2}\left(\frac{\Omega_{\mathrm{m}}}{a^3} + 2\frac{\Omega_{\mathrm{r}}}{a^3} - 2\Omega_\Lambda\right). \tag{1.13}$$

One sees that it is in principle possible to construct a static universe by fine-tuning Ω_{m}, Ω_{r} and Ω_Λ to achieve $\ddot{a} = \dot{a} = 0$. Indeed this was Einstein's intention, out of prejudice towards ideas of expansion and contraction, when introducing the cosmological constant to the Field Equations (though he later came to rue his decision). It can be shown however that such a universe would be very unstable to perturbations, necessitating extreme fine-tuning of the smoothness in the infinite past. Such a model is therefore not particularly realistic.

For any given universe, it is possible to solve the Friedmann Equation for the relationship between the scale factor and time, although one may have to resort to numerical methods where the universe is composed of multiple species. Solutions may be reached analytically in simple cases. For example, in a universe containing

only matter we have $a \propto t^{2/3}$, and in the case of only radiation, $a \propto t^{1/2}$. In both of these cases the universe has a beginning: $a = 0$ where $t = 0$. These are the so-called *Big Bang cosmologies*. In contrast, a universe containing vacuum energy alone has $a \propto \exp(H_0 t)$ and no point of cosmic creation can be defined. Certain combinations of the *cosmological parameters*, Ω_i, permit $a = 0$ at some point in the future, meaning that the universe meets its end in a *Big Crunch*.

It is important to note that although these equations entertain the concept of $a = 0$, General Relativity is not a quantum mechanical theory and hence breaks down on small scales ($a \to 0$) and at high energies. In this regime, an as-of-yet-undiscovered "quantum theory of gravity" is expected to run the game. The caveats surrounding a singular value of the scale factor notwithstanding, $a = 0$ remains a useful concept, allowing us to define an *age*, t_0, for any given universe as the time since this epoch (if indeed $a = 0$ *does* occur in said universe). This is achieved via manipulating the Friedmann equation to give,

$$t_0 = \int_0^1 \frac{da}{aH(a)}. \tag{1.14}$$

1.2.1 Cosmological redshift

In our own expanding Universe, as the scale factor grows, so too do the wavelengths of light propagating through the expansion. Thus the Hubble flow is detectable via the shifting of light, emitted by distant receding galaxies, towards the red end of the spectrum. The *cosmological redshift* can be derived from the FLRW metric (equation 1.2). Consider the first crest of a light ray (for which $ds^2 = 0$) travelling radially ($d\theta = d\phi = 0$) from the time of emission at the source, t_e, to the time of observation, t_o. Further consider the second crest of the light ray, emitted and observed at times, $t_e + \Delta t_e$ and $t_o + \Delta t_o$, respectively. As the comoving coordinate of the source, r_s, remains unchanged with the expansion, this leads to

$$R_0 \int_0^{r_s} dr = \int_{t_e}^{t_o} \frac{dt}{a(t)} = \int_{t_e+\Delta t_e}^{t_o+\Delta t_o} \frac{dt}{a(t)}. \tag{1.15}$$

Decomposing the dt integrals into separate pieces, (from t_e to $t_e + \Delta t_e$ etc.) and

cancelling common terms yields,

$$\int_{t_e}^{t_e+\Delta t_e} \frac{dt}{a(t)} = \int_{t_o}^{t_o+\Delta t_o} \frac{dt}{a(t)}. \tag{1.16}$$

Finally if we assume that the expansion between emission (and observation) of consecutive crests is small, (i.e. $a(t) = a(t + \Delta t)$) this reduces to,

$$\frac{\lambda_e}{a_e} = \frac{\lambda_o}{a_o}\,, \tag{1.17}$$

where $\lambda = \Delta t$ is the wavelength at either emission (subscript 'e') or observation ('o'). I now define the redshift as the relative difference between the wavelengths at these two times,

$$z = \frac{\lambda_o - \lambda_e}{\lambda_e}\,, \tag{1.18}$$

and setting the time of observation to be the present-day, such that $a(t_o) = a(t_0) = 1$, sets the relationship between the dimensionless scale factor and the redshift,

$$a = \frac{1}{1+z}. \tag{1.19}$$

It is often convenient in cosmology to use redshift as a proxy for distance rather than use the comoving coordinate, r, itself. This is because the redshift is directly observable and independent of the values of the cosmological parameters (the same cannot be said of the comoving distance, as is evident in the next Section). At the time of writing the most distant galaxy ever detected resides at $z = 11.09^{+0.08}_{-0.12}$ (Oesch et al., 2016).

1.2.2 Cosmological distances

It is possible to define a number of distance measures in cosmology. In this Section I present some of the most useful as far as observations are concerned; for a more comprehensive list I direct the interested reader to Hogg (1999).

We have already encountered the comoving distance in equation 1.15, which can alternatively be expressed as an integration over scale factor or redshift,

$$r = \frac{1}{R_0} \int \frac{da}{a^2 H(a)} = \frac{1}{R_0} \int \frac{dz}{H(z)}\,, \tag{1.20}$$

where the limits depend on the context of the problem. In calculating the comoving distance to a galaxy with observed redshift z_s for instance, the limits would be $z = 0$ and $z = z_s$. The presence of the Hubble parameter here (see equation 1.12) mean that the conversion between redshift and physical distance requires knowledge of the cosmological parameters, H_0 and Ω_i. Hence, the redshift combined with an *independent* measure of distance facilitates the inference of the values of these parameters in our Universe. This is discussed further in Section 1.3.2.

Using equation 1.20, it is also possible to compute the *particle horizon distance*, r_p, defined as the maximum distance light could have travelled in the finite amount of time since $a = 0$ (naturally this is only defined in Big Bang cosmologies). This distance marks the boundary between the observable and unobservable regions of a given universe. Its present-day value given by,

$$R_0 r_p \equiv R_p = \int_0^1 \frac{da}{a^2 H(a)} = -\int_0^\infty \frac{dz}{H(z)}. \tag{1.21}$$

The particle horizon defines a scale of *causal connectivity*; events separated by greater than this distance cannot ever have influenced each other. The apparent presence of causally connected points in the CMB on super-particle-horizon scales is considered evidence for the theory of inflation.

Due to cosmic expansion (or contraction), the particle horizon is not simply given by the inverse Hubble factor (modulo a factor of c). Such calculation yields a related but different distance, termed the *Hubble horizon*, $d_H = 1/H$. This defines the boundary between particles moving slower (within the horizon) or faster (due to expansion; outside the horizon) than the speed of light relative to an observer at a given time. The size of the Hubble horizon grows with time. In a universe with zero or negative acceleration in its expansion, objects outside of the Hubble horizon at earlier times will eventually enter it at later times and become observable. In a positively accelerating universe however, objects which are observable at the current time will, at some point, irreversibly pass beyond

observation.

Another commonly used scale in cosmology is the *angular diameter distance*, $D_{\rm A}$, defined as the ratio of an object's physical size to its angular size on the sky. In flat space this is equivalent to the distance to the object. $D_{\rm A}$ is related to the comoving distance, r, via

$$D_{\rm A} = \frac{R_0 S_k(r)}{1+z}. \tag{1.22}$$

Similarly, the *luminosity distance*, $D_{\rm L}$, defined as the distance which satisfies the relationship,

$$F = \frac{L}{4\pi D_{\rm L}^2}\,, \tag{1.23}$$

between bolometric flux, F, and luminosity, L, converges with the physical distance to an object only in the case of flat-space. It is defined from the metric as,

$$D_{\rm L} = (1+z) R_0 S_k(r) = (1+z)^2 D_{\rm A}. \tag{1.24}$$

Hence in an expanding universe (where redshift is positive) the luminosity distance is always larger than the angular diameter distance. This means that objects appear dimmer and larger on the sky than they would compared to a static universe. The faintness is a consequence of the reddening of emitted wavelengths due to expansion and also the dilution of the surface brightness associated with the object's apparent increase in size (relative to a non-expanding universe). The size magnification is attributable also to *gravitational lensing* due to the curvature of spacetime. This is discussed in more detail in Sections 1.3.3 and 2.2.

1.3 Concordance cosmology

Having established the basic mathematics governing the evolution of a general universe, and how this depends on the values of the cosmological parameters, I now focus on what is known with regards to our own Universe. Via a number of independent *cosmological probes*, precise measurements of the parameters for our Universe have been made over the past two decades, which in most cases show remarkable consistency (we will, however, in Section 1.5, draw attention to recent cosmological *tensions* which have arisen between different probes at significance

levels of ~ 2–4σ).

The fruit of this labour is the standard model, or 'concordance cosmology', ΛCDM, broadly illustrated by the constraints shown in Figure 1.1. The portrait of reality painted by this paradigm features flat geometry ($k \simeq 0$; $\Omega = 1$) with roughly 30% of the Universe's energy content corresponding to matter ($\Omega_{\rm m} \simeq 0.3$) and the rest corresponding to dark energy ($\Omega_\Lambda \simeq 0.7$). Of the matter component, observations suggest that only $\simeq 0.04$ is attributable to the familiar baryonic matter, of which all human experience is constructed, with the remainder corresponding to some exotic substance, termed *dark matter*. The exciting (or worrying, depending on one's outlook) implication of these results is that $\simeq 96\%$ of the energy content of our Universe resides in the mysterious dark sector (dark matter and energy), and currently lacks theoretical understanding. The dark components consequently dominate *100%* of the name given to the standard model, with 'Λ' representing dark energy and 'CDM' being an acronym for cold (i.e. non-relativistic) dark matter. In this Section, I will outline the metrics used to diagnose the cosmology of our Universe, before discussing the key pieces of observational evidence which lead us to the conclusions of the standard model.

1.3.1 Growth of structure

We can parameterise the growth of cosmic structures in terms of the dimensionless density perturbation of 3-dimensional comoving scale size, $\boldsymbol{r}$, as

$$1 + \delta(\boldsymbol{r}) = \frac{\rho(\boldsymbol{r})}{\langle\rho\rangle}\,, \tag{1.25}$$

where $\langle\rho\rangle$ is the average density within a volume V (which can be taken to be infinite if required).

How the density fluctuations for matter and radiation respond to the expansion of our Universe depends on whether these species experience *adiabatic* or *isocurvature fluctuations*. These orthogonal modes act as either fluctuations in the energy density (adiabatic) or entropy density (isocurvature). In the former case, density fluctuations are present initially (such is the prediction of inflationary cosmology, where initial quantum fluctuations are greatly enhanced by the short-lived period of exponential expansion). Adiabatic perturbations alter the matter and radiation *number* densities by the same factor, and thus the *energy* densities

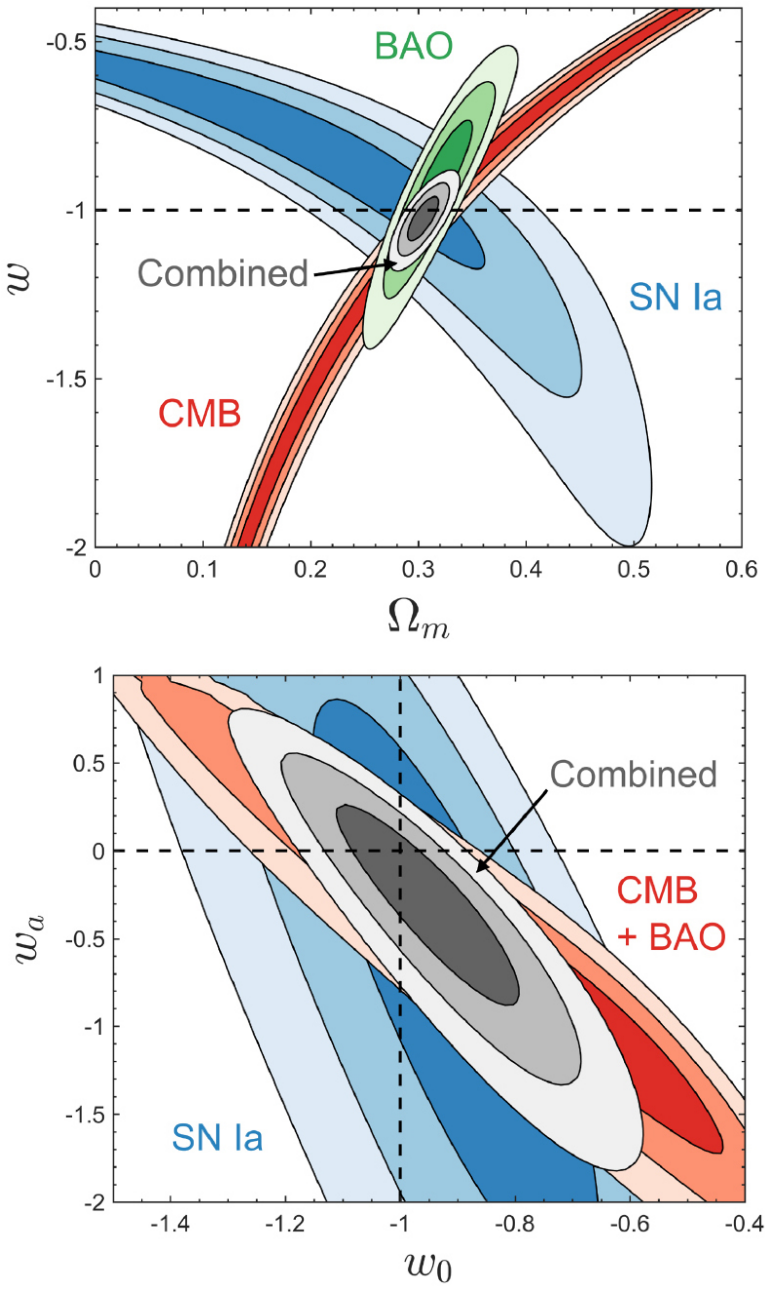

Figure 1.1 *From Huterer & Shafer (2018). Constraints on cosmological parameters from some of the observational probes discussed in this Section: Type Ia supernovae (Betoule et al., 2014, "SN Ia"), baryon acoustic oscillations (Alam et al., 2016, "BAO"), the cosmic microwave background (Planck Collaboration et al., 2016a, "CMB") and their combination (grey). Upper panel: constraints on the mass energy density, $\Omega_{\rm m}$, and the constant vacuum equation of state parameter, w. Lower panel: constraints on evolving dark energy energy, parameterised by w_0 and w_a, the present and time-dependent values of w, respectively. The contours contain 68.3%, 95.4%, and 99.7% of the likelihood, and a flat geometry is assumed, such that $\Omega_\Lambda = 1-\Omega_{\rm m}$. Our Universe is found to be consistent with one that is vacuum-dominated, flat, accelerating and containing a cosmological constant with $w = -1$.*

in different ways, $\delta_{\rm r} = 4\delta_{\rm m}/3$ (given that matter and radiation energy density are proportional to temperature to the third and fourth powers respectively). Isocurvature perturbations on the other hand, caused by stresses in the matter, alter the number densities differently but maintain the overall energy density as homogeneous, $\delta_{\rm r}\rho_{\rm r} = -\delta_{\rm m}\rho_{\rm m}$. Isocurvature perturbations move matter around in a causal way and therefore must, by definition, be on sub-Hubble-horizon scales, $d_H = 1/H$. In contrast, inflationary models change the nature of the horizon scale at early times, meaning that adiabatic fluctuations can exist on scales much larger than d_H.

As we will see in Section 1.3.2, observations of the CMB are fully consistent with purely adiabatic density perturbations. In this case, for $\delta \ll 1$, perturbations grow with time according to,

$$\delta \propto \begin{cases} a^2, & (a < a_{\rm eq},\, \lambda > d_H(a)) \\ a, & (a > a_{\rm eq}) \end{cases}, \tag{1.26}$$

where $a_{\rm eq}$ is the scale factor corresponding to $t_{\rm eq}$, the time at which equality between the energy densities of matter and radiation occurs (Peacock, 1999). The quadratic dependence of δ with scale factor prior to equality is conditional also on the wavelength of the perturbation, λ, exceeding the Hubble horizon size at the given epoch, $d_H(a)$. This is because perturbations smaller than this threshold find their growth suspended by pressure forces which oppose gravity. As the Universe expands so too does the horizon, swallowing increasingly larger scale perturbations, subsequently halting their growth until equality.

At later times, where $a \gg a_{\rm eq}$, the linear growth of structure changes once again, following a growth function, $g(a)$, according to

$$\delta(a) = \delta_0 \frac{g(a)}{g(a_0)}\,, \tag{1.27}$$

where an analytical form for $g(a)$, neglecting the effects of baryons, is derived in Carroll et al. (1992), as

$$g(a, \Omega_{\rm m}, \Omega_\Lambda) = \frac{5}{2}\Omega_{\rm m}(a)\left[\Omega_{\rm m}(a)^{4/7} - \Omega_\Lambda(a) + \left(1 + \frac{\Omega_{\rm m}(a)}{2}\right)\left(1 + \frac{\Omega_\Lambda(a)}{70}\right)\right]^{-1}. \tag{1.28}$$

After equality, the cold, collisionless dark matter undergoes gravitational collapse whilst baryonic matter remains oscillating, subject to the the radiative pressure

of photons. Finally, recombination releases baryons from the grip of radiation and they fall into the potential wells created by the dark matter. This mechanism of structure formation, referred to as 'bottom-up' with larger structures growing from smaller ones, is supported by observational evidence including the existence of dwarf galaxies (Battinelli & Demers, 2006). If instead dark matter was relativistic, or 'hot', then all small-scale structure is washed-out by *free streaming* and structure formation takes the 'top-down' mechanism, which is more difficult to reconcile with observations (White et al., 1984).

The linear matter power spectrum

So far I have purposefully referred to density fluctuations and structure growth over a range of *scales* rather than at specific *positions*. This is because the evolution of large-scale structure is inherently *random*, understood only in terms of the statistical probabilities of properties such as structure sizes and clustering. With this in mind, via Fourier analysis, we can begin to build up a general density field as a superposition of many modes. A random, isotropic density field for example, such as would be produced by primordial adiabatic perturbations, is fully described by the ensemble averaged power spectrum,

$$(2\pi)^3 \, P_\delta(k) \, \delta_D(\boldsymbol{k} - \boldsymbol{k'}) = \langle |\delta(\boldsymbol{k}) \delta^*(\boldsymbol{k})| \rangle \, , \tag{1.29}$$

where $*$ denotes the complex conjugate, δ_D is the Dirac delta function, included due to the cross-terms ($\boldsymbol{k'} \neq \boldsymbol{k}$), averaging to zero, and $\delta(\boldsymbol{k})$ is the Fourier transform of the real-space density contrast,

$$\delta(\boldsymbol{k}) = \int d^3\boldsymbol{r} \, \delta(\boldsymbol{r}) \, e^{-i\boldsymbol{r}.\boldsymbol{k}}. \tag{1.30}$$

Here the wave number[6], $k = |\boldsymbol{k}|$, is related to the wavelength of density perturbations, λ, via $k = 2\pi/\lambda$. As previously mentioned, the growth of perturbations are halted upon entering the Hubble horizon, i.e. when

$$\lambda = d_H = \frac{c}{a_{\rm enter} H(a_{\rm enter})} \, , \tag{1.31}$$

where $a_{\rm enter}$ is the scale factor at the time, $t_{\rm enter}$, when a given perturbation

[6]Not to be confused with the spatial curvature parameter, k.

enters the horizon, and the extra factor of $a_{\rm enter}$ converts the physical scale to one that is comoving. Substitution of the Hubble parameter (see equation 1.12) and considering only times dominated by either radiation ($t \ll t_{\rm eq}$) or matter ($t_{\rm eq} \ll t \ll t_0$) yields the dependence of the perturbation wavelength on the scale factor at the time of entering the horizon,

$$\lambda \propto \begin{cases} a_{\rm enter}, & (t \ll t_{\rm eq}) \\ \sqrt{a_{\rm enter}}, & (t_{\rm eq} \ll t \ll t_0) \end{cases} . \tag{1.32}$$

From this we can formulate the relationship between the power spectrum at time $t_{\rm enter}$ and the primordial power spectrum, $P_i(k)$. From the definition of the power spectrum in terms of the density contrast (equation 1.29), and the equation for the growth of δ before and after the equality epoch (equation 1.26), we obtain

$$P_{\rm enter} \propto \begin{cases} a_{\rm enter}^4 P_i(k) \propto k^{-4} P_i(k), & (t \ll t_{\rm eq}) \\ a_{\rm enter}^2 P_i(k) \propto k^{-4} P_i(k), & (t_{\rm eq} \ll t \ll t_0) \end{cases} , \tag{1.33}$$

where I have used equation 1.32 to convert the dependence on $a_{\rm enter}$ to dependence on λ and thus k. By assuming a scale-invariance to the power of density fluctuations, such that $k^3 P_{\rm enter} =$ constant, we find that the form of the initial power spectrum can be written as $P_i(k) \propto k^{n_s}$, where n_s is the *spectral index*. As we shall see in Section 1.3.2, observational evidence from the CMB is highly consistent with the linear model, $n_s = 1$, corresponding to the Harrison-Zel'dovich spectrum (Harrison, 1970; Zel'dovich, 1970).

For perturbations with $\lambda < d_H$, the suppression of the power spectrum can be parameterised as $P_\delta(k) = f_{\rm sup}^2 P_i(k)$, where the suppression factor,

$$f_{\rm sup} = \left(\frac{a_{\rm enter}}{a_{\rm eq}} \right)^2 , \tag{1.34}$$

takes the appropriate form to effectively cancel the growth expressed in equation 1.33. Since the suppression only affects small physical scales (large wave numbers), a turnover will be manifested in the late-time power spectrum caused by the unimpeded growth on large physical scales (small wave numbers). Defining k_0 as the wave number corresponding to the horizon size at $t = t_{\rm eq}$, and again employing equation 1.32 to convert the scale factor dependence of the suppression

factor to $f_{\rm sup} \propto k^{-2}$, gives

$$P_\delta(k) \propto \begin{cases} k, & (k \ll k_0) \\ k^{-3}, & (k \gg k_0) \end{cases} . \tag{1.35}$$

This relatively simple modelling framework powerfully captures the key features of the *linear* power spectrum ($\delta \ll 1$). The solid lines in Figure 1.2 show the present-day linear $P_\delta(k)$ predictions for three cosmologies with zero curvature ($\Omega_{\rm m} + \Omega_\Lambda = 1$). As $\Omega_{\rm m}$ is increased, the epoch at which matter-radiation equality occurs is pushed to earlier times, resulting in a smaller Hubble horizon, $d_H(e_{\rm eq})$. Correspondingly, we see the turnover point in the power spectrum, $k_0 = 2\pi/d_H(a_{\rm eq})$, translates to higher wave numbers for larger matter density components, with approximately linear rise and cubic decline of $P(k)$ for $k \ll k_0$ and $k \gg k_0$ respectively. The dependence of the turnover point on the present-day Hubble value is typically removed by plotting the wave number in units of h/Mpc.

The evolution of the power spectrum from its primordial form is hence sensitive to cosmological information. This dependence is modelled using the transfer function, $T(k, z)$, as

$$P_\delta(k, z) = T(k, z) P_i(k) . \tag{1.36}$$

Knowledge of the form of $T(k, z)$ is invaluable for relating the power spectrum inferred from observations to cosmological parameters. However, the mixture of different species which play a role in structure formation (collisionless dark matter and neutrinos, collisional baryons and photons) combine to form a fluid with complex behaviour. The solution lies in solving the *Boltzmann Equation*, a complicated collection of coupled differential equations involving the distribution of positions and momenta of all relevant species and incorporating the relevant scattering physics (see, e.g., Peebles, 1980). For a multiple-species fluid this has to be undertaken numerically. For a universe consisting of only cold dark matter with adiabatic perturbations, there are various formulae fit to the results of such calculations (Bond & Efstathiou, 1984; Bardeen et al., 1986; Eisenstein & Hu, 1998). For example, the Bardeen et al. (1986) transfer function takes the form,

$$T(k) = \frac{\ln(1 + 2.34q)}{2.34q} \left[1 + 3.89q + (16.81q)^2 + (5.46q)^3 + (6.71q^4)\right]^{-1/4} , \tag{1.37}$$

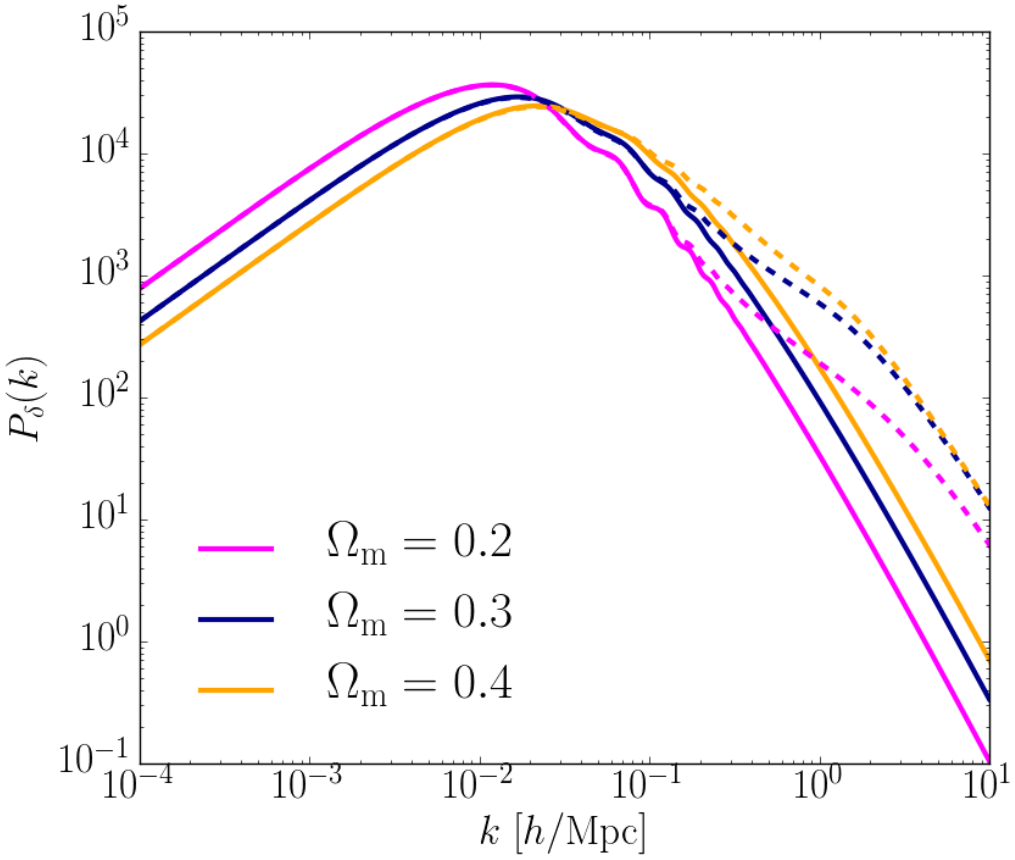

Figure 1.2 *The linear (solid lines) and non-linear (dashed lines) $P_\delta(k)$ for three cosmologies with flat geometry ($\Omega_\mathrm{m} + \Omega_\Lambda = 1$). As Ω_m is increased, the turnover point in the power spectrum translates to larger wave numbers (smaller physical scales). The boost in non-linear power at high k, computed using the Takahashi et al. (2012)* HALOFIT *approach, occurs due to the viriliasation of collapsing structure.*

where

$$q \equiv \frac{k/h\mathrm{Mpc}^{-1}}{\Gamma}\,, \tag{1.38}$$

and Γ is the shape parameter, given by Ωh in a universe with zero baryon content. For universes with a finite baryonic component or non-flat geometry however, Γ takes a more sophisticated form (Peacock & Dodds, 1994; Sugiyama, 1995; Eisenstein & Hu, 1998). The redshift dependence of the transfer function is then asserted by modulating $T(k)$ by the growth function (equation 1.28),

$$T(k,z) = a(z)\frac{g(z)}{g(z=0)}T(k). \tag{1.39}$$

The non-linear matter power spectrum

So far I have only considered the evolution of the density field in the limits of linear perturbations, ($\delta \ll 1$). The density contrast on the length scales of clusters and individual galaxies however is orders of magnitude higher than unity. Perturbations on these small scales therefore evolve *non-linearly*. A consequence of this is collapsing structures stabilise, amplifying the tail of the power spectrum at $k \gg k_0$. Accurately predicting the amplitude of this increase as a function of cosmological parameters is one of the key pursuits of our field in the modern age.

The most accurate results for the evolution of the density field are currently obtained from *N-body simulations*, in which the density field is represented by an ensemble of $(\sim 10^3)^3$ discrete particles. The equations of motion for each particle, given the gravitational field generated by all of the others, are solved at discrete time steps, leading to new particle positions and velocities at each iteration. Even with the baryonic physics neglected, these gravity-only simulations (see, e.g., Heitmann et al., 2010; Dietrich & Hartlap, 2010; Heitmann et al., 2016; Liu et al., 2018; Harnois-Deraps et al., 2019) are enormously expensive, consuming thousands of hours of computation time. Consequently, it is completely infeasible to simulate every combination of cosmological parameter values, even within the range of today's observational constraints.

One solution to this conundrum, from Peacock & Dodds (1994), is to run a feasible number of simulations and fit the free parameters of a formula, f_{NL}, mapping empirically between the linear and non-linear power spectra according to

$$\Delta^2_{\mathrm{NL}}(k_{\mathrm{NL}}) = f_{\mathrm{NL}}\left[\Delta^2_{\mathrm{L}}(k_{\mathrm{L}})\right]\,, \tag{1.40}$$

where subscripts 'L' and 'NL' denote linear and non-linear quantities respectively, and Δ^2 is the dimensionless power spectrum defined as,

$$\Delta^2(k) = \frac{k^3 P_\delta(k)}{2\pi^2}. \tag{1.41}$$

Here, the linear and non-linear wave numbers are related via,

$$k_{\rm L} = k_{\rm NL}\left[1 + \Delta^2_{\rm NL}(k_{\rm NL})\right]^{-1/3}. \tag{1.42}$$

Hence an estimation of the non-linear power spectrum for an arbitrary cosmology (within a reasonable range of those in the simulations used for calibration) can be easily computed. Improvements on this formalism have since been made by Smith et al. (2003) and later Takahashi et al. (2012) using the *halo model* approach. This framework models the total non-linear power spectrum as having contributions from two components,

$$\Delta^2_{\rm NL}(k) = \Delta^2_{\rm Q}(k) + \Delta^2_{\rm H}(k)\,, \tag{1.43}$$

where $\Delta^2_{\rm Q}(k)$, called the two-halo term, represents the power due to the clustering of haloes on large, quasi-linear scales, whereas the one-halo term, $\Delta^2_{\rm H}(k)$, is the power arising on smaller scales from the clustering of dark matter particles within the same halo. Fitting functions are obtained for these two components separately, achieving non-linear power spectra predictions which are within 5% of the results of numerical simulations for $k \leq 1\,h^{-1}$Mpc and $0 \leq z \leq 10$ (Takahashi et al., 2012). The non-linear predictions from this approach, termed HALOFIT, for the three flat-geometry cosmologies, $\Omega_{\rm m} = \{0.2, 0.3, 0.4\}$, are shown by the dashed lines in Figure 1.2. These clearly illustrate the boost in small-scale power due to the stabilisation of collapsing structures. More recently, Mead et al. (2015) extended this formalism to models beyond the standard cosmology, ΛCDM, achieving 5% precision out to higher wave numbers.

Improving the accuracy of non-linear calibration to the per cent level is achievable using an *emulator* to effectively interpolate between the predictions from numerical simulations (Habib et al., 2007; Schneider et al., 2008; Heitmann et al., 2009, 2016; Lawrence et al., 2017; Euclid Collaboration et al., 2018; Winther et al., 2019; Harnois-Deraps et al., 2019). Of this I shall say little now, returning to cosmic emulators in earnest in Section 1.4, and Chapters 4 and 5.

1.3.2 Observations

Distance-redshift

The first detection of the recession velocities of galaxies by Slipher (1917), along with the subsequent characterisation of the linear relationship between the velocities, v, and distances, x,

$$v = H_0 x \tag{1.44}$$

provided the first evidence that we live in an expanding Universe. We now know equation 1.44, expressing the famous *Hubble Law*, is in fact just the low-redshift ($z \simeq 0; a \simeq 1$) limit of the Friedmann Equation (see Eq. (1.12)). Over these small distances it is reasonable to treat the redshift as the straight-forward Doppler shift ($z = v/c$). Continuing to use the $c = 1$ convention, this gives,

$$z = H_0 x. \tag{1.45}$$

The "constant" of proportionality in these linear relations, the Hubble parameter or Hubble constant, is poorly-named given that it is in fact a function of time. However, it varies very slowly compared to human timescales, and thus can be treated as a constant over short time and distance scales. Similarly, I make no specification on which of the distance measures presented in Section 1.2.2 is represented by x, since all converge for $z \simeq 0$.

The units of H_0 are inverse-time, equivalent to inverse-distance with natural unit convention. The uncertainty in the value of H_0 is typically consolidated in the of-order-unity parameter h, such that $H_0 = 100h\,\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$.

Equation 1.45 implies that the redshift of a single galaxy combined with an independent measure of its distance is sufficient to measure the Hubble parameter. However, inhomogeneities in density (due to structures such as galaxy clusters or voids) causes the galaxies in proximity to have gravitationally-induced *peculiar velocities*. This means that the expansion rate measured locally may differ from the global expansion rate of the isotropic and homogeneous Universe (Ben-Dayan et al., 2014). In order to suppress the bias caused by local measurements of the Hubble flow, large samples of objects from a statistically representative volume of

the Universe are required. Given such a sample, reliable inference of the Hubble parameter then rests on obtaining accurate estimates of their distances.

In cases where an object is close enough to Earth to be reliably resolved, the angular size combined with a means to estimate the *absolute* size of the object provides an estimate of its distance. The 1987 supernova (SN 1987A) in the Large Magellanic Cloud (LMC) is one notable example in which this was possible. The absolute size of the SN explosion is given by the time between the onset and peak of the UV line emission (Panagia, 1999). This event proved serendipitous indeed, also providing an estimate of the distance to the Cepheid variable stars residing in the LMC ($\sim 163,000\,\mathrm{lyr}$ or $z \sim 10^{-3}$). This was useful in calibrating the tight correlation observed between the period of pulsation and luminosity for these objects, thus facilitating more precise implementation of Cepheids as cosmic distance probes.

For higher-redshift sources ($z_s \gtrsim 0.1$) the cosmological parameters begin to play a non-negligible role in the distance–redshift relation, given by

$$D_{\mathrm{L}} = \frac{1+z_s}{H_0} \int_0^{z_s} \frac{dz}{\sqrt{\Omega_{\mathrm{m}}(1+z)^3 + \Omega_\Lambda}}\,, \tag{1.46}$$

for our flat matter- and dark-energy-dominated Universe. At these larger distances, measuring the flux from *standard candles* (astronomical sources with known intrinsic luminosities) presents the most promising measure of distance. In addition to Cepheids, the extremely luminous thermonuclear explosions of *Type Ia supernovae* (SNe) are used. Unlike Type II SNe (e.g. SN 1987A), characterised by hydrogen emission lines in their spectra and believed to be caused by the core collapse of massive stars, Type Ia spectra exhibit an absence of hydrogen and the presence of ionised silicon (SiII). The progenitors of Type Ias are thought to be carbon-oxygen white dwarfs exceeding a critical (Chandrasekhar) mass of 1.44 solar masses via either accretion or merger in a binary system (Weinberg et al., 2013). Strictly speaking, Typa Ia SNe are not perfect standard candles; Hamuy et al. (1996) identified a scatter of approximatey 0.4 magnitudes in their V-band peak luminosities. A more precise description would be *standardisable candles* owing to a tight empirical correlation between peak luminosity and light curve shape, first identified by Phillips (1993). With additional corrections for the connection between SNe colour and extinction as well as redshifting effects (the K-correction), the scatter in peak luminosity can be reduced to $\simeq 0.12$ mag, leading to a distance uncertainties of $3-4\%$ per SN (Folatelli et al., 2010).

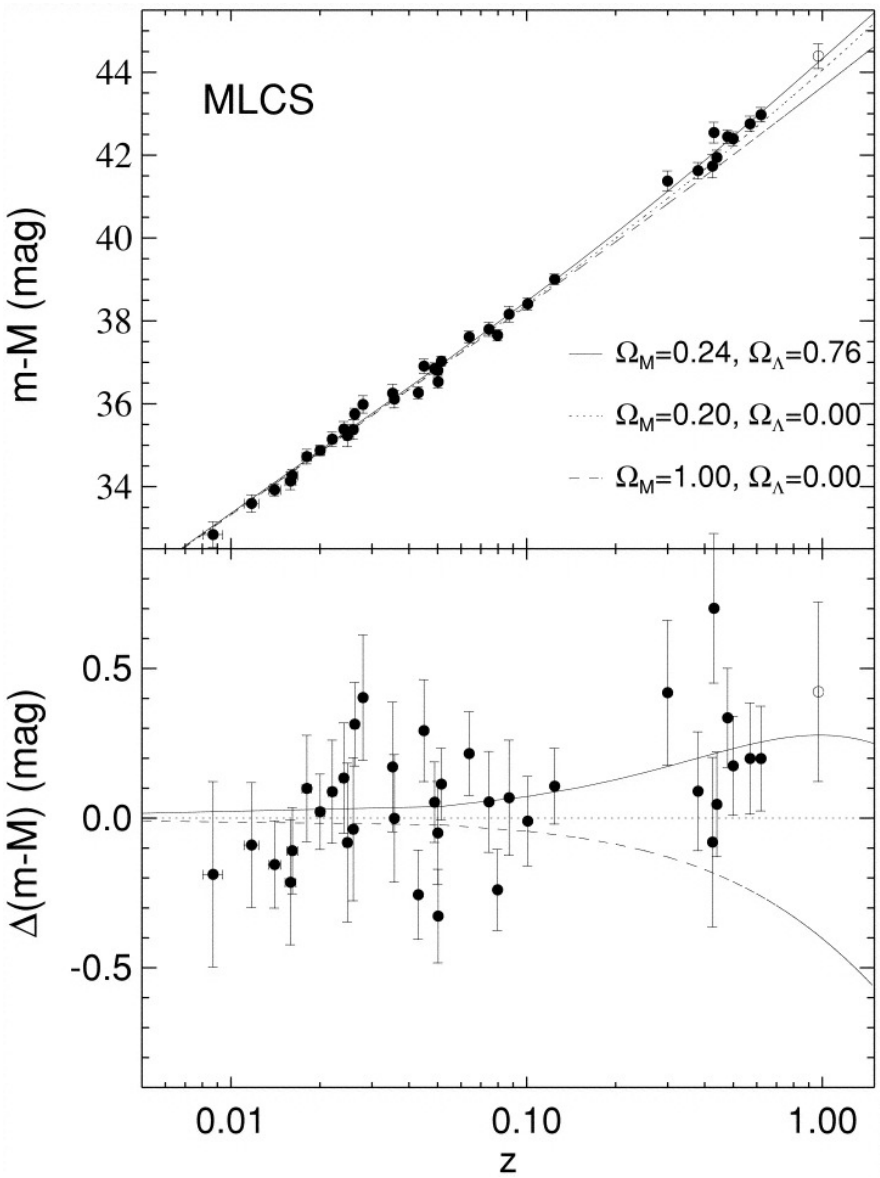

Figure 1.3 *Hubble diagram from Riess et al. (1998) showing a proxy for luminosity distance to the Type Ia SNe (the difference in apparent and absolute magnitudes) versus redshift. Upper panel: the theoretical predictions for three different cosmologies are shown for reference: "low" and "high"* $\Omega_{\rm m}$ *with* $\Omega_\Lambda = 0$*, and the best-fit result for a flat Universe,* $\Omega_{\rm m} = 0.24, \Omega_\Lambda = 0.76$*. Lower panel: the difference between the results in the upper panel and the* $\Omega_{\rm m} = 0.20, \Omega_\Lambda = 0$ *prediction.*

The discovery that the expansion of our Universe is in fact *accelerating* was made using Type Ia SNe to measure cosmological distances (Riess et al., 1998; Perlmutter et al., 1999). Figure 1.3 shows the Hubble diagram from the former study, showing a proxy for luminosity distance versus redshift for the sample of observed Type Ia SNe. Shown for reference are the theoretical predictions for two universes with zero cosmological constant ($\Omega_\Lambda = 0$) and high and low Ω_m values, as well as the best-fit cosmology ($\Omega_\mathrm{m} = 0.24, \Omega_\Lambda = 0.76$). Riess et al. (1998) found $\Omega_\Lambda = 0$ to be ruled-out by the data with 3–4σ significance (depending on the method of lightcurve fitting), increasing to 7–9σ if the curvature parameter is fixed to zero. The present-day value of the *deceleration parameter*, defined as,

$$q_0 = -\frac{\ddot{a}(t_0)}{\dot{a}^2(t_0)}\,, \tag{1.47}$$

was found to be *negative* (indicating an *accelerating Universe*) with 2.8–3.9σ significance.

The cosmic microwave background

Complimentary to SNe as a probe of cosmology is the CMB, from which a goldmine of information on the constituents and evolution of our Universe has been extracted since its discovery by Penzias & Wilson (1965).

The CMB features anisotropies as a consequence of, among other things, the Sachs-Wolfe effect (the gravitational redshift of photons caused by their ascent out of potential wells), Doppler frequency shifts (owing to the finite plasma velocities at last scattering), as well as adiabatic plasma perturbations. Cosmological information contained in these anisotropies, are encoded into the CMB power spectrum. Statistical treatment is very similar to that of the spatial density discussed in Section 1.3.1, with fractional temperature fluctuations, $\delta T/T$, replacing δ, but with the added caveats that the fluctuations are integrated radially, generating an *angular* temperature power spectrum, and the Fourier treatment is superseded by spherical harmonics when the sky curvature becomes non-negligible.

Figure 1.4 shows the full-sky temperature fluctuation map observed by the Planck Collaboration et al. (2018) in the upper panel, with the corresponding angular power spectrum, compared to the best fit ΛCDM cosmology, occupying the lower panel. The increasing size of the error bars at large angular scales (small values

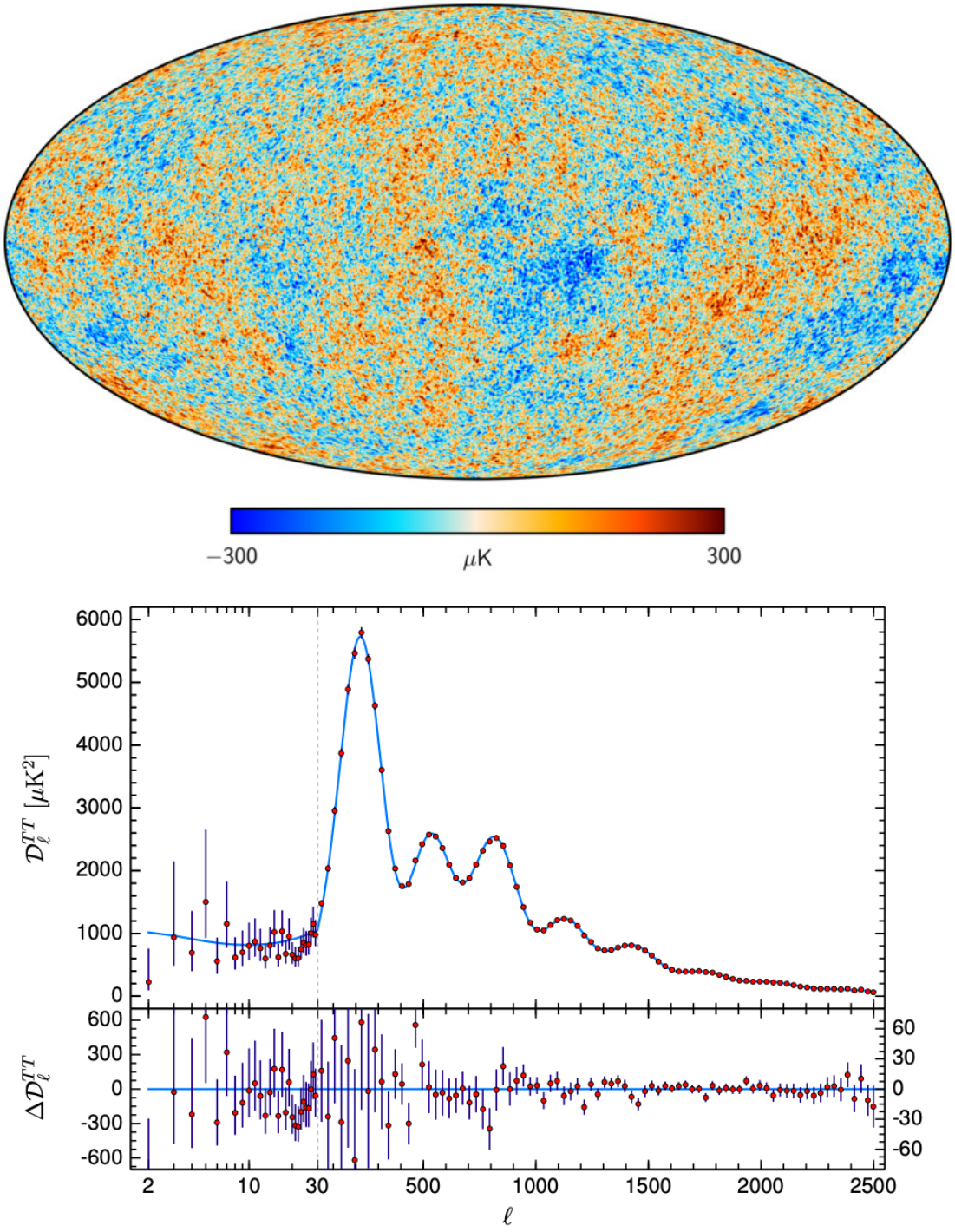

Figure 1.4 *Upper: the full-sky map of the CMB temperature fluctuations about the mean of 2.73K, observed by Planck Collaboration et al. (2018). Lower: The corresponding angular temperature fluctuation, TT, power spectrum (data points) as a function of angular multipole ℓ, relative to the best fit ΛCDM cosmology (solid line).*

of the angular multipole, ℓ) indicate the uncertainty due to cosmic variance, arising from the fact that we have but one Universe to study. Clearly visible in the angular power spectrum are the *acoustic peaks*, arising from the conditions in the photon-baryon plasma prior to the epoch of recombination. The contest between opposing forces, gravity and radiation pressure, caused the fluid to ring with relativistic sound waves, called *baryon acoustic oscillations* (BAOs). At recombination the species decouple, and the oscillatory information is frozen into the CMB power spectrum. The heights of the acoustic peaks are sensitive to the overall-matter and baryon energy densities, providing constraints on $\Omega_{\rm m}h^2$ and $\Omega_{\rm b}h^2$, where the h^2 dependence arises because it is the physical densities, rather than the dimensionless quantities (see equation 1.11), which affect these features (Weinberg et al., 2013). The CMB can hence provide tight constraints on the Hubble constant under the assumption of a flat Universe with a cosmological constant. The constraining power is weakened however if these parameters are allowed to vary, owing to their influence on the locations of the peaks in the spectrum (Doroshkevich et al., 1978; Kamionkowski et al., 1994; Hu & Dodelson, 2002).

The high level of consistency between the theoretical predictions and the measurements illustrates how successfully the CMB is described by ΛCDM. Planck Collaboration et al. (2018) obtain highly precise constraints on the energy content parameters, agreeing with other probes on a flat geometry with $\Omega_{\rm m} = 0.315 \pm 0.007$, and more than 80% of this component being composed of dark matter. The Hubble constant, tilt of the primordial power spectrum, effective number and mass sum of neutrinos, the spatial curvature and the dark energy equation of state parameter are also constrained, all consistent with a flat ΛCDM Universe featuring a cosmological constant with adiabatic, Gaussian initial conditions, as predicted by simple inflationary models.

In addition to the temperature fluctuation, 'TT', power spectrum, Planck Collaboration et al. (2018) also measure the CMB *polarisation*, 'EE', spectrum and its cross-correlation with the temperature, the 'TE' spectrum. The polarisation fluctuations are a consequence of the anisotropic radiation field experienced by the electrons at last scattering. As well as adding statistical power, polarisation helps break the degeneracy between the value of $\Omega_{\rm b}$ and whether there are *tensor*, in addition to *scalar*, contributions to the fluctuations (Peacock, 1999). The latter relate directly to mass fluctuations, such as those mentioned at the start of this Section on the CMB. The former however could be

the result of *primordial gravitational waves*, a prediction of inflation (Starobinskii, 1985). However, there has been no detection of tensor modes within the limits of present CMB experiments.

Baryon acoustic oscillations

BAOs also leave their fingerprint on the late-time clustering of galaxies and matter. The suppression of matter clustering prior to recombination at $z = z_{\rm rec}$, means that an *acoustic scale*, r_s, given by the comoving distance traversed by the sound waves between the Big Bang and this epoch,

$$r_s = -\frac{1}{R_0}\int_{z_{rec}}^{\infty}\frac{c_s(z)}{H(z)}dz\,, \tag{1.48}$$

is imprinted on large-scale structure. The acoustic scale is large, $\sim 150\,$Mpc comoving, owing to the relativistic sound speed, maxing out at $c_s = 1/\sqrt{3}$ at early times when the matter component is negligible. Consequently, there is an excess of galaxies separated by this scale, manifesting in a bump at $r = r_s$ in the galaxy correlation function, $\xi(r)$, defined as the probabilistic excess of finding a galaxy pair with separation r, and in wiggles at $k \sim 0.1\,h/$Mpc in its Fourier transform, the matter power spectrum (Peebles & Yu, 1970; Sunyaev & Zeldovich, 1970, see also Figure 1.2). Simulations suggest the size of this feature is stable to better than 1% accuracy (Weinberg et al., 2013), making it an excellent *standard ruler*, the apparent size of which can be used to probe distances and thus cosmological parameters.

One of the primary challenges of this method is statistical - the BAO feature is weak and large scale, such that large volumes of the Universe must be mapped to extract the signal. The Baryon Oscillation Spectroscopic Survey (BOSS) has amassed observations of the 3D clustering of more than 1.2 million galaxies between redshift 0.2 and 0.75, from which the BAO scale has been measured. For this data set, Ross et al. (2017) measured the dependence of the galaxy correlation function, ξ, on the redshift-space separation, s, and $\mu = \cos(\theta)$, where θ is the angle from the line of sight[7]. Figure 1.5 shows the monopole and quadrupole moments of the galaxy correlation function, obtained by integrating $\xi(s, |\mu|)$, weighted by the first two even Legendre polynomials, over μ. The excess clustering at the acoustic scale is clearly visible in both moments.

[7]Distinguishing between transverse and radial (i.e. redshift-space) separation measurements is important due to the effects of *redshift-space distortions* discussed in the next Section.

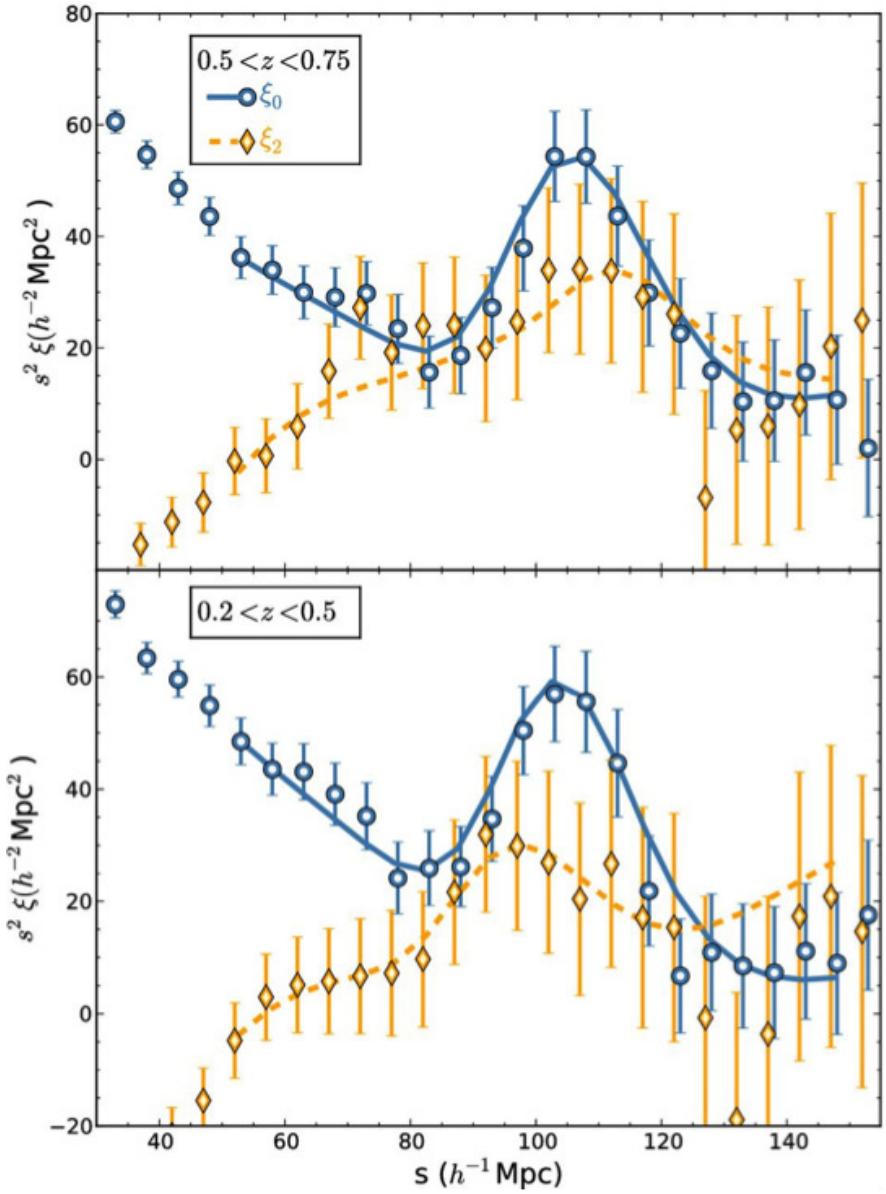

Figure 1.5 *The monopole (blue) and quadrupole (yellow) moments of the correlation function of BOSS galaxies, as a function of the redshift-space separation, s, in two reshift bins, specified in the panels (Ross et al., 2017). A bump, indicating excess clustering at the BAO scale, is clearly visible.*

Despite the relative robustness in the acoustic scale as a standard ruler, it is subject to systematics including the shift in this length as a result of second-order gravitational perturbations. For example, Padmanabhan & White (2009) compute a shift in the acoustic scale of 0.25% at $z = 0$ via these effects. Peculiar velocities also create anisotropies which inhibit measurement of the acoustic scale along the line of sight. *Galaxy bias*, the fact that the observed galaxies do not perfectly trace the underlying unobserved dark matter distribution, is an additional source of systematic error. If the bias is linear between the density contrasts of galaxies and dark matter, then only the amplitude and not the shape of the correlation function would change. However, realistic bias models must be at least somewhat non-linear; Mehta et al. (2011) numerically computed shifts in the acoustic scale of 0.1–0.8%, depending on the strength of the galaxy bias.

Redshift-space distortions

The distribution of galaxies observed by, for example, BOSS or the 2dF Galaxy Redshift Survey will appear anisotropic, as illustrated by Figure 1.6, in contrast to the recurring themes of isotropy discussed up until now. This arises because galaxy positions are measured in terms of their angular coordinates and redshifts. Converting between z and physical distance is approximate because the peculiar velocities of galaxies in virialised structures cause the redshifts to depart from those induced by the Hubble flow. This leads to *redshift-space distortions* (RSDs) - scattering of galaxies in redshift, equating to displacements along the line of sight and constructing radial 'fingers of God' that appear to point at the observer.

These effects break the isotropy of the cosmos; density perturbations along the line of sight will differ from those transverse. To separate these effects, the power spectrum can be decomposed into a function of the wave number, k, and the viewing angle cosine, μ. Given perfect tracers of the matter distribution, Kaiser (1987) related the linear redshift-space power spectrum, Δ_s^2, to its real-space counterpart,

$$\Delta_s^2(k, \mu) = (1 + f_g \mu^2)^2 \Delta^2(k)\,, \tag{1.49}$$

where f_g is the linear theory growth rate, given by

$$f_g = \frac{d \ln g(a)}{d \ln a}\,, \tag{1.50}$$

and $g(a)$ is the growth function expressed in equation 1.28. We see that for

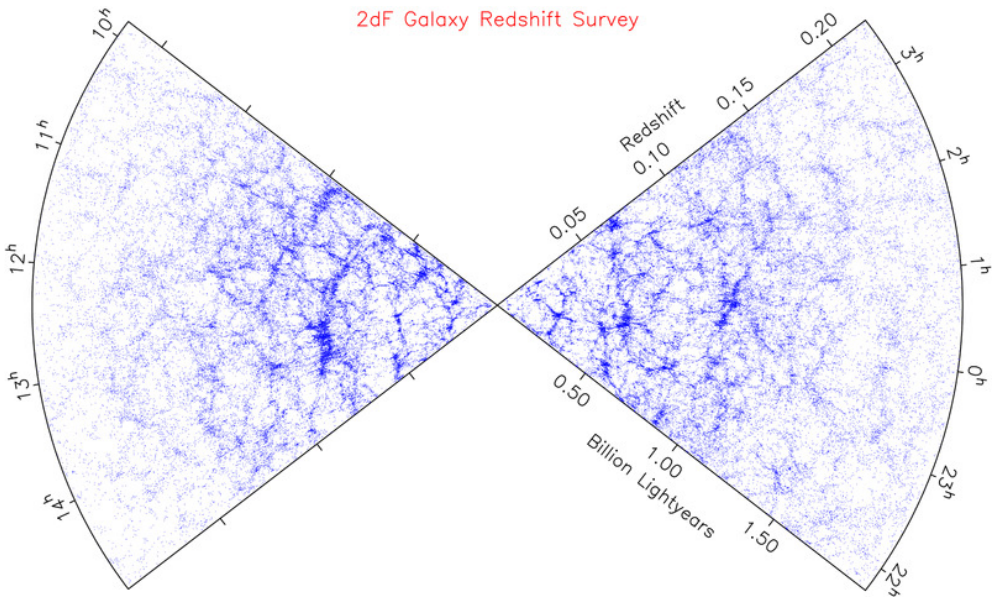

Figure 1.6 *The anisotropic distribution of galaxies observed by the 2dF Galaxy Redshift Survey. Visible is the 'great wall' on the left, empty void regions and the radial 'fingers of God' caused by the viriliased motions of galaxies within clusters, acting to scatter galaxies parallel to the line of sight.*

perturbations transverse to the line of sight, $\mu = 0$, and the redshift- and real-space power are the same. Modes with some component along the line of sight however, enjoy a boost in power, thereby enhancing radial clustering.

In actuality, we observe galaxies as biased tracers of the matter. Assuming a linear bias, b, between the galaxy and dark matter mass, the redshift-space power spectrum relation becomes,

$$\Delta^2_{s,g}(k,\mu) = b^2(1+\beta\mu^2)^2\Delta^2(k)\,, \tag{1.51}$$

where $\beta = f_g/b$. The more strongly biased galaxies exhibit the weakest RSDs, since linear bias amplifies clustering isotropically. In this expression, the strongest cosmological dependence resides in the growth rate term. Hence, given the $\Delta^2(k)$ for an approximately correct cosmology, the redshift-space power spectrum measured from the galaxy survey provides a constraint on β. Further assuming the galaxy bias can be estimated independently, these constraints can be converted onto cosmological parameters through the growth rate.

Inconveniently, non-linearities play an even larger role in redshift-space. This has led to numerous efforts to derive corrections to, or extend the range of validity of, the Kaiser (1987) approximation on non-linear scales (Peacock & Dodds, 1994; Scoccimarro, 2004; Taruya et al., 2010; Simpson et al., 2016a).

Galaxy clusters

Galaxy clusters boast a long history as probes of cosmology. From the curiously high velocity dispersions of galaxies in the Coma cluster, Zwicky (1933) inferred the presence of non-luminous material greatly outweighing the luminous, thereby providing the first evidence for dark matter. Hoessel et al. (1980) used observations of 116 nearby Abell clusters to anchor the bright end of the Hubble diagram, and in doing so provided some of the earliest evidence, albeit at low significance, for the cosmic acceleration confirmed by Riess et al. (1998). Offsets between the mass and light distributions of merging clusters is also used to test the nature of dark matter (Clowe et al., 2006).

In particular, the number density of galaxy clusters as a function of mass and redshift is a powerful probe of dark energy and the growth of structure. This is parameterised in terms of the mean space density, $n(M,z)$, in units of comoving

volume, also called the *mass function*,

$$\frac{dn}{d\ln M} = \frac{\rho_{\rm m}}{M}\left|\frac{d\ln\sigma}{d\ln M}\right| f(\sigma)\,. \tag{1.52}$$

Here $f(\sigma)$ is a model-dependent function of the variance of linearly evolved density fluctuations, filtered on a mass scale $M = 4\pi\rho_{\rm cr}R^3/3$,

$$\sigma^2(M,a) = \int \frac{d^3k}{(2\pi)^3}\, W^2(kR)\, P_\delta(k,a)\,, \tag{1.53}$$

where a typical form for the filter function is the sharp, top-hat, $W(y) = 3\left[\sin(y)/y^3 - \cos(y)/y^2\right]$ (Allen et al., 2011). Evaluating this at $8\,h^{-1}$Mpc and $a = 1$ gives the matter power spectrum normalisation parameter, σ_8, a cosmological parameter commonly constrained in data. The halo mass function can be calibrated through the $f(\sigma)$ term. For example, Tinker et al. (2008) use a large suite of N-body simulations to fit the free parameters, $\{A, a, b\}$, of the functional form from Sheth & Tormen (1999),

$$f(\sigma) = A\left[\left(\frac{\sigma}{b}\right)^{-a} + 1\right] e^{-c/\sigma^2}\,, \tag{1.54}$$

finding a mild evolution of the parameters with redshift is required to maintain statistical precision of 5% in halo number at higher redshift.

Thus, cluster abundance observations connect us with cosmological parameters, with particular sensitivity to the combination $\sigma_8\Omega_{\rm m}^q$, and $q \simeq 0.4$ (White et al., 1993). Figure 1.7 shows simulated cluster mass functions of various cosmologies relative to those inferred from optical and X-ray observations from Bahcall & Cen (1992). This provided some of the earliest evidence for a sub-critical mass energy density, ($\Omega_{\rm m} < 1$).

The principal difficulty of this method, however, is calibrating the relations between cluster masses and the actual quantities of observation (Rozo et al., 2010; Sehgal et al., 2011; von der Linden et al., 2014; Planck Collaboration et al., 2016b; Schrabback et al., 2018). In optical surveys, for example, clusters are quantified by the optical richness (the number of galaxies per luminosity and colour bin within the virial radius of the host dark matter halo). In X-ray surveys, the mass, temperature and luminosity of the hot cluster gas all provide indicators of the halo mass. CMB surveys instead detect clusters by their Sunyaev & Zeldovich (1970) signature, Y_{SZ}, i.e., the Compton up-scattering of CMB photons by electrons

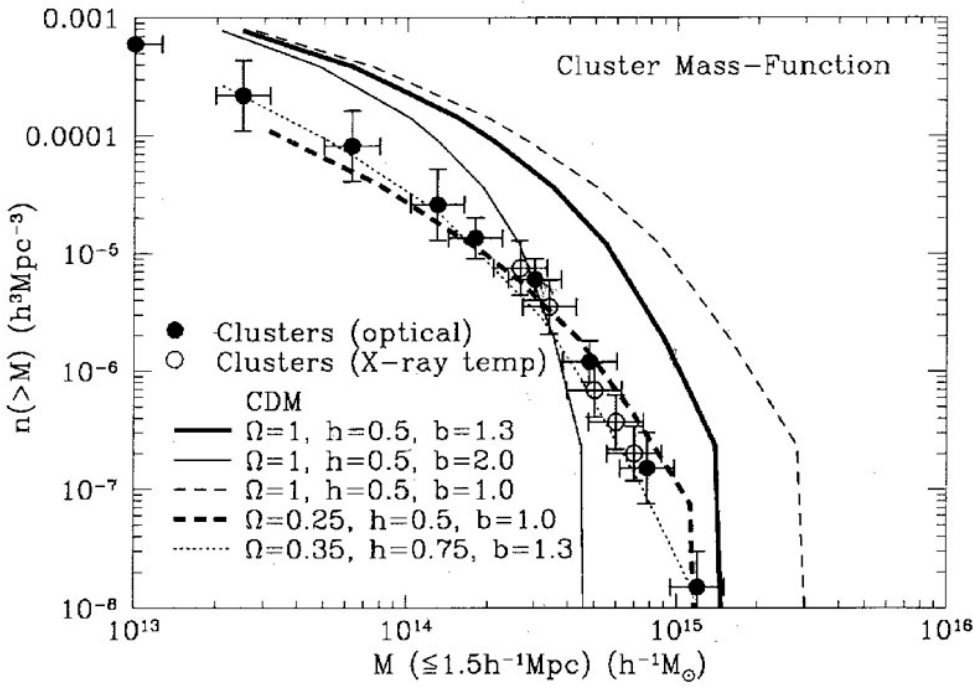

Figure 1.7 *Cluster mass functions measured in simulations for various cosmologies (solid/dashed lines) relative to those from optical and X-ray observations (Bahcall & Cen, 1992, data points). We see that models with critical values of the mass energy density (here represented by Ω) fit the data poorly.*

in the hot intra-cluster medium. Y_{SZ} has been found to scale with the X-ray observables linking to halo mass.

1.3.3 Gravitational lensing

We arrive now at the family of observations to which much of this thesis is dedicated: gravitational lensing. Here, I describe only briefly its main pillars. Detailed description of its core principles is presented in Chapter 2.

The detection of gravitational lensing by Sir Arthur Eddington in 1919 stands as one of the first experimental validations of the Theory of General Relativity. By observing a number of stars during a solar eclipse, whilst on an expedition to the Gulf of Guinea, Eddington showed their positions were altered due to the gravitational deflection of their starlight by the Sun. Most importantly, the deflection angle was larger than the expectation of Newtonian physics, which predicts half as much lensing as Einstein's theory. The *New York Times* proclaimed *"Lights All Askew in the Heavens ... Men of Science More or Less Agog ... but Nobody Need Worry"*. Einstein consequently became a household name.

Despite Newtonian theory also predicting the deflection of light's path in a gravitational field, its argument is wholly different from that of GR. Einstein's theory predicts that particles follow *geodesic* paths (those which maximise the action) through spacetime, and that the distortion of spacetime (described by, for example, the FLRW metric in equation 1.2) due to matter is what causes the observed curvature in the particle's trajectory. The greater the concentration of matter (energy density) the more severely the trajectory is perturbed. Gravitational lensing therefore has cosmological implications, providing a means to probe foreground mass structures via the deflection of light from background sources. Particularly lensing proves effective where other methods, which rely on the electromagnetic emission of foregrounds, will fail, e.g., the detection and characterisation of dark matter.

This method, however, is a broad church, with a large range of strengths in the lensing signal manifesting in a variety of different observational signatures. These can roughly be divided into three camps: *strong*, *micro-* and *weak* lensing.

Strong lensing

Strong lensing pertains to the study of massive structures as the foreground 'lenses', for example, galaxy clusters. The hallmark of strong lensing is multiple images of the same background source, possibly blended into a single arc of light, called the *Einstein Ring*, as shown in Figure 1.8. The angular positions of the multiple images depend on the ratios of the angular diameter distances between the source, lens and observer, and as such are encoded with cosmological information (see, e.g., Jullo et al., 2010). However the results are sensitive to substructure either in the lens or mass distributions along the line of sight (D'Aloisio & Natarajan, 2011). Strong lensing of intrinsically variable sources, such as quasars, also permit cosmological inference via the relative time delays in the fluctuations of the multiple images, which scale as H_0^{-1} (Weinberg et al., 2013). The distance scale is, however, degenerate with the surface mass density of the lens, necessitating an independent estimate of the latter to isolate the cosmological information (Kochanek, 2003).

Microlensing

Transient microlensing events occur when a compact massive body, such as a star, brown dwarf or black hole, is coincident with the line of sight to a background source, such as a star or quasar. The gravitational lensing caused by the foreground object briefly magnifies the source, manifesting in a spike in that body's lightcurve. The frequency of these events has been used to constrain how much dark matter might be composed of, so-called, MAssive Compact Halo Objects (*MACHOs*), i.e., ordinary, low-luminosity baryonic matter, including primordial black holes. Over a broad range of masses, microlensing observations have convincingly ruled out MACHOs comprising all of the dark matter content demanded by cosmological observations (Brandt, 2016; Green, 2016; Niikura et al., 2019). Similarly, the hopes of the *axion*, a hypothetical extension to the standard model of particle physics, as a serious contender to the dark matter component of the Universe, look doubtful given initial constraints from microlensing (Fairbairn et al., 2017).

Figure 1.8 *The mischievous cosmic smile of strong lensing, captured by the NASA/ESA Hubble Space Telescope. Whereas the eyes are two galaxies, SDSSCGB 8842.3 and SDSSCGB 8842.4, the smile lines are background sources which have been multiply-imaged and blended into an arc of light, the Einstein Ring, via strong lensing.*

1.3.4 Weak lensing

The small but coherent distortions in the images of galaxies, due to the spacetime curvature induced by the intermediary large-scale structure, falls into the weak lensing category - the primary probe considered in this thesis. Unlike its strong sister, weak lensing does not generate multiple images of sources. In fact the induced change in a given galaxy's shape and orientation, or *shear*, is only at the per cent level, in most cases dwarfed by systematics including the *intrinsic* galaxy ellipticity and the *point-spread function* of the instrument used for observation (see Section 2.6). Therefore, the weak lensing signal can only be coherently extracted by statistically correlating the observed shapes of large samples of galaxies - a method referred to as *cosmic shear*.

Cosmic shear

Commonly, cosmic shear is quantified with the two-point *shear correlation functions*, $\xi_{\pm}(\theta)$. These are estimated from large photometric galaxy surveys, as the excess correlation in the galactic ellipticities and orientations (compared to if they were random), as a function of angular separation, θ,

$$\widehat{\xi_{\pm}}(\theta) = \frac{\sum_{\rm ab} w_{\rm a} w_{\rm b} \left[\epsilon_{\rm t}(\boldsymbol{\theta}_{\rm g,a})\epsilon_{\rm t}(\boldsymbol{\theta}_{\rm g,b}) \pm \epsilon_{\times}(\boldsymbol{\theta}_{\rm g,a})\epsilon_{\times}(\boldsymbol{\theta}_{\rm g,b})\right]}{\sum_{\rm ab} w_{\rm a} w_{\rm b}}. \tag{1.55}$$

The summation here is over pairs of galaxies a and b positioned at angular coordinates $\boldsymbol{\theta}_{\rm g,a/b}$, within an interval $\Delta\theta$ about the angular separation θ (Bartelmann & Schneider, 2001). The $\epsilon_{\rm t}$ and $\epsilon_{\times}$ terms designate the tangential- and cross- components of the observed ellipticities measured relative to the vector $\boldsymbol{\theta}_{\rm g,a} - \boldsymbol{\theta}_{\rm g,b}$ connecting the galaxy pairs. w is the weight ascribed to the measurement of the ellipticity components, which are treated as unbiased estimates of the shear.

The $\widehat{\ }$ notation is included here to distinguish between the shear correlation functions estimated from data, and the theoretical predictions, related to the *convergence* power spectrum, $P_{\kappa}(\ell)$, by

$$\xi_{\pm}(\theta) = \frac{1}{2\pi} \int {\rm d}\ell\, \ell\, P_{\kappa}(\ell)\, J_{0,4}(\ell\theta)\,, \tag{1.56}$$

where the zeroth $J_0(\ell\theta)$ and fourth $J_4(\ell\theta)$ order Bessel functions of the first kind are used for ξ_+ and ξ_- respectively. The convergence power spectrum $P_\kappa(\ell)$ is in turn related to the matter power spectrum, $P_\delta(\ell)$ (see equation 1.29), via

$$P_\kappa(\ell) = \int_0^{r_\mathrm{H}} \mathrm{d}r \, \frac{q(r)^2}{D_\mathrm{A}(r)^2} \, P_\delta \left(k = \frac{[\ell + 1/2]}{D_\mathrm{A}(r)}, r \right), \tag{1.57}$$

where r_H is the comoving radial distance to the horizon and D_A is the angular diameter distance (see equation 1.22). The flat-sky first-order extended Limber approximation, discussed further in Section 2.4.2, is assumed here. The lensing efficiency, $q(r)$, is defined as

$$q(r) = \frac{3H_0^2\Omega_\mathrm{m}}{2c^2} \frac{D_\mathrm{A}(r)}{a(r)} \int_r^{r_\mathrm{H}} \mathrm{d}r' \; n(r') \frac{D_\mathrm{A}(r' - r)}{D_\mathrm{A}(r')} \, . \tag{1.58}$$

Thus we see the connection between weakly lensed galaxy images and cosmological parameters. Cosmic shear has a soft spot for Ω_m and σ_8 in particular, primarily sensitive to the combination, $S_8 = \sigma_8(\Omega_\mathrm{m}/0.3)^q$, with $q \simeq 0.5$.

Precision cosmic shear measurement in the modern age demands galaxy sample sizes in the millions, realistically only achievable by relatively large scientific collaborations. At the present time, there are three main weak lensing surveys taking and analysing independent observational data: the Kilo-Degree Survey (KiDS; de Jong et al., 2013), the Dark Dark Energy Survey (DES; Frieman & Dark Energy Survey Collaboration, 2013) and the Hyper Suprime-Cam (HSC; Mandelbaum & Hyper Suprime-Cam (HSC) Collaboration, 2017) survey. Dedicated cosmic shear analyses have recently been performed by each survey (Hildebrandt et al., 2017; Troxel et al., 2017; Hikage et al., 2018; Hildebrandt et al., 2018; Hamana et al., 2019), yielding consistent results for the large-scale clustering in our Universe.

A comparison of some of the key quantities in these analyses are presented in Table 1.1. We see that DES currently has the largest survey footprint and largest ensemble of galaxy shapes (1514 deg^2, 26 million), followed by KiDS (450 deg^2, 15 million) and then HSC (137 deg^2, 9 million). However, this trend reverses when we consider the precision and depth with which source galaxies are imaged, as parameterised by the seeing (the amount of blurring due to atmospheric and instrumental effects, discussed more in Section 2.6.1) and mean z_source statistics. HSC ranks most impressive on both of these quantities (0.58

Table 1.1 *A comparison of recent cosmic shear analyses by the three main weak lensing surveys currently operating: the Kilo-Degree Survey (Hildebrandt et al., 2018, KiDS), the Dark Energy Survey (Troxel et al., 2017, DES) and the Hyper-Suprime Cam (Hikage et al., 2018, HSC) survey. The galaxy number density is calculated using the prescription given in Chang et al. (2013), which depends on the ellipticity dispersion, here given for the case of no redshift cuts on the respective shear catalogues. The median seeing quoted corresponds to the wavelength band in which the galaxy shapes were measured (highlighted in bold on the sixth row). Note that KiDS augment their survey with five near-infrared bands, $ZYJHK_{\rm s}$, from the VISTA Kilo-Degree Infrared Galaxy Survey (VIKING).*

	KiDS(-VIKING)	DES	HSC
Telescope/camera	VST/OmegaCAM	Blanco/DECam	Subaru/HSC
Telescope diameter [m]	2.6	4.0	8.2
Current footprint [deg^2]	450	1514	137
Final footprint [deg^2]	1350	5000	1400
Galaxy density [arcmin^{-2}]	6.9	5.1	16.5
Ellipticity dispersion	0.29	0.27	0.40
Galaxy number [million]	15	26	9
Bands	$ug\boldsymbol{r}i(ZYJHK_{\rm s})$	$g\boldsymbol{r}iZY$	$gr\boldsymbol{i}ZY$
Median seeing [arcsec]	0.66	1.03	0.58
$\langle z_{\rm source}\rangle$	0.71	0.59	0.81
Inferred $S_8 = \sigma_8/\sqrt{\Omega_{\rm m}/0.3}$	$0.737^{+0.040}_{-0.036}$	$0.777^{+0.036}_{-0.038}$	$0.780^{+0.030}_{-0.033}$

arcsec, 0.81, respectively), again followed by KiDS (0.66 arcsec, 0.71), and then DES (1.03 arcsec, 0.59). Overall, the statistical power of three surveys is fairly comparable at the present time, as indicated by the relative sizes of the error bars on the inferred S_8 parameter.

Figure 1.9 shows the shear correlation functions estimated from the first 450 square-degrees of KiDS deep multi-band imaging (Hildebrandt et al., 2018). The statistical power is improved by incorporating *tomography*, i.e., auto-correlating and cross-correlating the shear in and across redshift bins, numbered 1–5 here for uniform redshift binning in the range $z \in [0.1, 1.2]$. Estimating the redshifts of source galaxies in a cosmic shear analysis is one of the chief sources of systematic error, a problem I return to in Section 2.6.

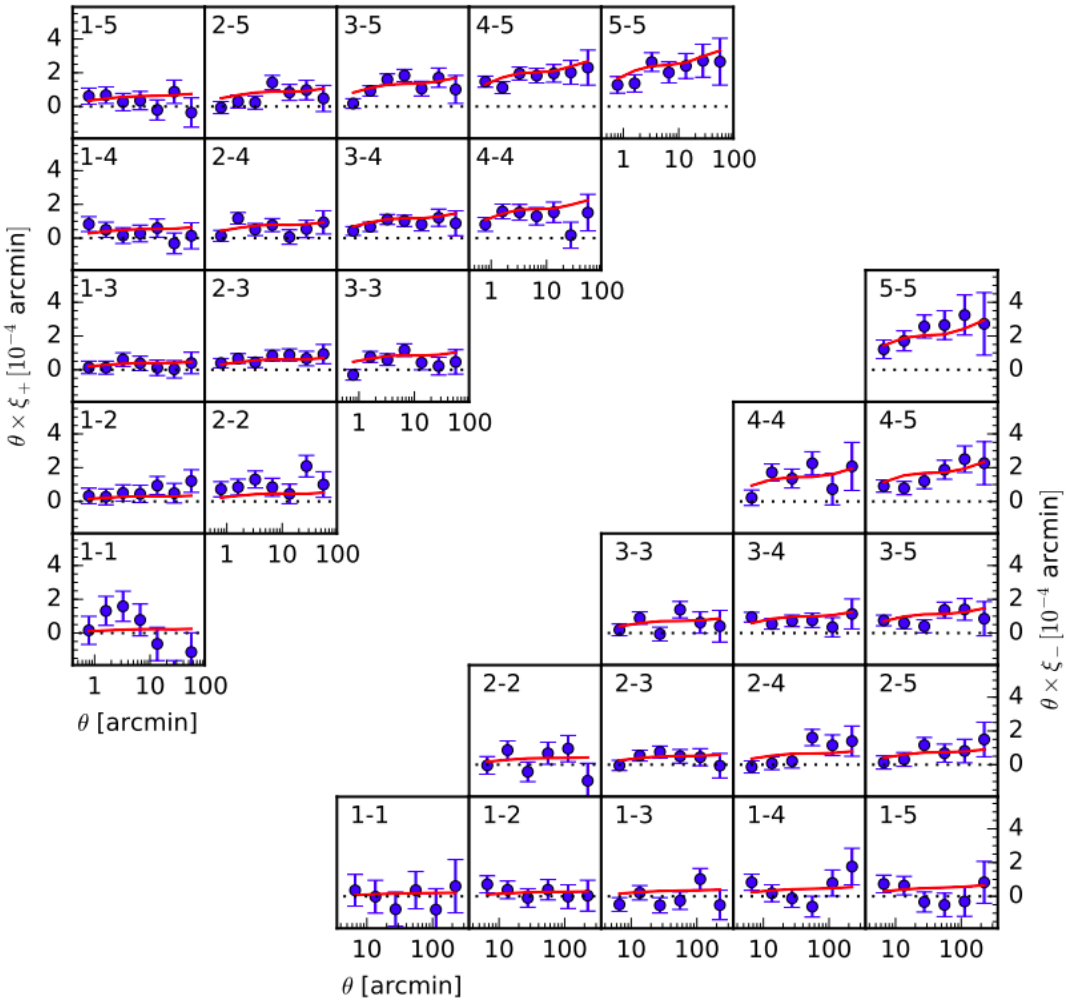

Figure 1.9 *The tomographic shear correlation functions (weighted by angular separation, θ) estimated from the first 450 square-degrees of KiDS imaging (data points) relative to the theoretical prediction for the best-fit cosmology (Hildebrandt et al., 2018).*

1.4 Machine learning cosmology

If there is but one take-home message from Section 1.3, it is that traditional, theory-based methods of inferring cosmological parameters have proven highly successful. In recent times, however, a new family of statistical methods has emerged, potentially signalling a paradigm shift in how cosmological analyses are conducted. Scientists, and not just those within cosmology, but indeed all disciplines, are increasingly exercising *machine learning* algorithms, to empirically infer unknown relationships between input and output, based on examples, known as the *training set* (see, for example, Carleo et al., 2019, for general discussion). A broad range of machine learning techniques are available, capable of performing tasks including data interpretation, pattern recognition, classification, prediction, compression and inference, all of which have their place in a cosmology analysis.

For example, artificial neural networks (NNs), based loosely on the neurons of the human brain, consist of nodes which receive, process and pass information between one another along weighted connections. An NN may have several 'hidden' layers of nodes, in which case the term *deep learning* is used to describe its functionality. NNs are effective at image classification and mapping non-linear relations. As such, they have been applied to a multitude of cosmological problems, such as: strong gravitational lens identification (Petrillo et al., 2019); photometric redshift estimation (Collister & Lahav, 2003; Vanzella et al., 2004; Bonnett, 2015; Sadeh et al., 2016, see Section 2.6.1 for further discussion); galaxy shape measurement (Ribli et al., 2019); learning the influence of baryonic gas pressure on large-scale mass distributions (Tröster et al., 2019, see Section 2.6.2); and inference of cosmological parameters directly from weak lensing and CMB maps, bypassing traditional diagnostics such as the power spectrum (Auld et al., 2007; Gupta et al., 2018; Fluri et al., 2019), to name but a few.

An alternative, arguably simpler, machine learning method lies in *Gaussian process* (GP) models, with which many results in this thesis are obtained. A GP is a collection of random variables, any subset of which is jointly Gaussian distributed (Rasmussen & Williams, 2006). Whilst NNs require a number of decisions to be made (the architecture of the layers, the functions activating connections between nodes, the learning rate etc.), GPs typically need only the specification of a mean and a kernel, which controls the covariance of the output with input variables (see Chapter 4 for in-depth discussion). In practice this is not so difficult, with only a small number of functional forms being sensible

for most problems. Furthermore, it has also been shown that very large NNs converge with GPs in the limit of infinite size (Neal, 1994).

As briefly mentioned in Section 1.3.1, GPs have been used by many authors in the construction of *emulators* to predict statistics as a function of cosmological parameters (Habib et al., 2007; Schneider et al., 2008; Heitmann et al., 2009, 2016; Lawrence et al., 2017; Euclid Collaboration et al., 2018; Winther et al., 2019; Harnois-Deraps et al., 2019; Giblin et al., 2019). My work in this area, first appearing in these latter two references, is presented in Chapters 4 and 5, respectively. GPs have also been implemented in a diverse range of other problems, such as standardising Type Ia SN lightcurves (Kim et al., 2013), inference of photometric redshift distributions (Almosallam et al., 2016; Johnson et al., 2017; Rau et al., 2019) and simulation design optimisation (Rogers et al., 2019).

The scope of applications for machine learning in scientific inquiry is difficult to overstate. There are concerns over the functionality of these algorithms however, such as a lack of interpretability, i.e., that scientific inference is made from a "black box". This may mean that in using these methods, we are open to the "Clever Hans" effect, named for the early 20^{th} century German horse who appeared capable of solving mathematics problems, but was in fact inferring the correct answer from the body language of his trainer. An example of this phenomenon in machine learning is given by Lapuschkin et al. (2019), where neural networks correctly identify images of horses, but rather than use the visual characteristics of the animal, instead base their inference on a spurious image artefact which so happened to be present in many of the training set horse images. This causes the algorithm to incorrectly label images of cars featuring the same artefact, as horses. This example poses the question, what is machine learning, really learning?

One further concern to bare in mind with some of these techniques, is their dependence on vast computational resources and subsequent energy consumption. A recent study by Strubell et al. (2019) quantified the carbon footprint of training several common large NN architectures, finding the CO_2 emissions to be nearly five times those of the average American car in its entire lifetime. With the machine learning trend currently showing no signs of slowing down, it will be interesting to see how statistics such as this influence its usage in cosmology in the future.

1.5 Discordance cosmology?

The standard cosmological model, ΛCDM, has been tested a great many times and been shown to be consistent with a plethora of observational constraints. This is especially impressive given the diversity of means through which said data was collected, using completely different instruments to observe different regions and epochs of the Universe, with many of the assumptions made being independent. Nevertheless, in the current *cosmo-political* climate, where "Tensions in Cosmology" conferences are increasingly commonplace, it would be remiss not to mention certain *discordances* which, at least for some, cast doubt on the standard model.

Most significant of these is the discrepancy between the values of the Hubble constant inferred from early- and late-time probes. Riess et al. (2018b) report a $\sim 4\sigma$ deviation between the H_0 measured using Type Ia SNe and Cepheid variables, among other objects, to calibrate the distance-redshift relation, and the value inferred from the CMB by Planck Collaboration et al. (2016a). A less significant, $\sim 2\sigma$ tension has also been found between the values of $S_8 = \sigma_8\sqrt{\Omega_\mathrm{m}/0.3}$ measured by Planck and from cosmic shear (Hildebrandt et al., 2018). Although the other major cosmic shear analyses of late report consistency with the CMB results (Troxel et al., 2017; Hikage et al., 2018), discussions in the literature with regards to the lensing-CMB tension (or lack of) are on-going (Troxel et al., 2018; Handley & Lemos, 2019; Joudaki et al., 2019).

A very straight-forward explanation for the differences in the inferred cosmological parameters between these probes is unaccounted-for systematic error(s). Though appealing in its simplicity this idea may be, the key players contributing to the measurement of the inferred tension, argue[8] that all known relevant systematics are sufficiently controlled (see, for example, Hatt et al., 2017; Kilbinger et al., 2017b; Riess et al., 2018a; Hildebrandt et al., 2018; Riess et al., 2019), meaning that no such culprit has been conclusively identified to date.

An alternative possibility, therefore, is an incompleteness in the standard model which underlies all of the aforementioned lines of enquiry. A few notable exceptions aside (Abbott et al., 2016; Abbott et al., 2017), GR has been tested extensively only within the Solar System, and so its extrapolation to cosmic

[8]Convincingly, in this author's humble opinion. I refer the reader to Section 2.6 for discussion of systematics pertaining to weak lensing, and their mitigation strategies.

distances, is analogous to assuming the physical laws governing human experience, having only knowledge of those on nuclear scales. If the domain of suitability for GR is narrower than the cosmic scales to which is has been applied, gravity and the evolution of the Universe between early- and late-times would not behave as expected, perhaps explaining the deviations between observations made of the early- (CMB) and late-time (SNe, lensing) Universe.

Various extensions to the standard model have been reviewed in the literature with aims to solve the discordance (Verde et al., 2013; Joudaki et al., 2017; Mörtsell & Dhawan, 2018; DES Collaboration et al., 2018). For example, Figure 1.10 presents results from Riess et al. (2016), which show how the size of the discrepancy between early- and late-time H_0 measurements can be reduced with: **(1)** '$\Delta w_0 = -0.1$': a 0.1 change in the vacuum equation of state parameter, w_Λ; **(2)** '$\Delta w_a = -1$': a time-dependent term in the equation of state, given by $w_\Lambda(a) = w_0 + \Delta w_a a$; **(3)** '$\Delta N_{\rm eff} = +1$': incrementing the effective number of relativistic species at decoupling, beyond the standard model prediction of $N_{\rm eff} = 3.046$ neutrinos; **(4)** '$\Delta\Omega_k = -0.01$', slightly altering the spatial curvature to favour an open Universe. Also shown are four SN-independent H_0 measures (Bonamente et al., 2006; Sorce et al., 2012; Suyu et al., 2013; Gao et al., 2016, magenta error bars) and the combined CMB-BAO constraint from Hinshaw et al. (2013, green error bar).

It is the opinion of this author, that such discrepancies strongly justify undertaking more precise cosmological investigations, to unravel whether there is any truth in the notion of physics beyond the standard model.

1.6 Setting the scene

In this introductory Chapter, I have briefly provided answers to the big 'what?', 'how?' and 'why?' questions concerning our Universe. This covered the key events in cosmic evolution (Section 1.1; the 'what?'), the mathematics facilitating an understanding of the physics which govern the Universal expansion (Sections 1.2-1.3; the 'how?'), and the observational evidence providing justification for this picture of reality (Section 1.3.2; the 'why?'). Novel machine learning approaches to cosmological analyses, made possible by 21$^{\rm st}$-century advances in computing, have been discussed (Section 1.4), and a light shone on the discordant comparison of observational constraints, perhaps casting doubt upon the standard model

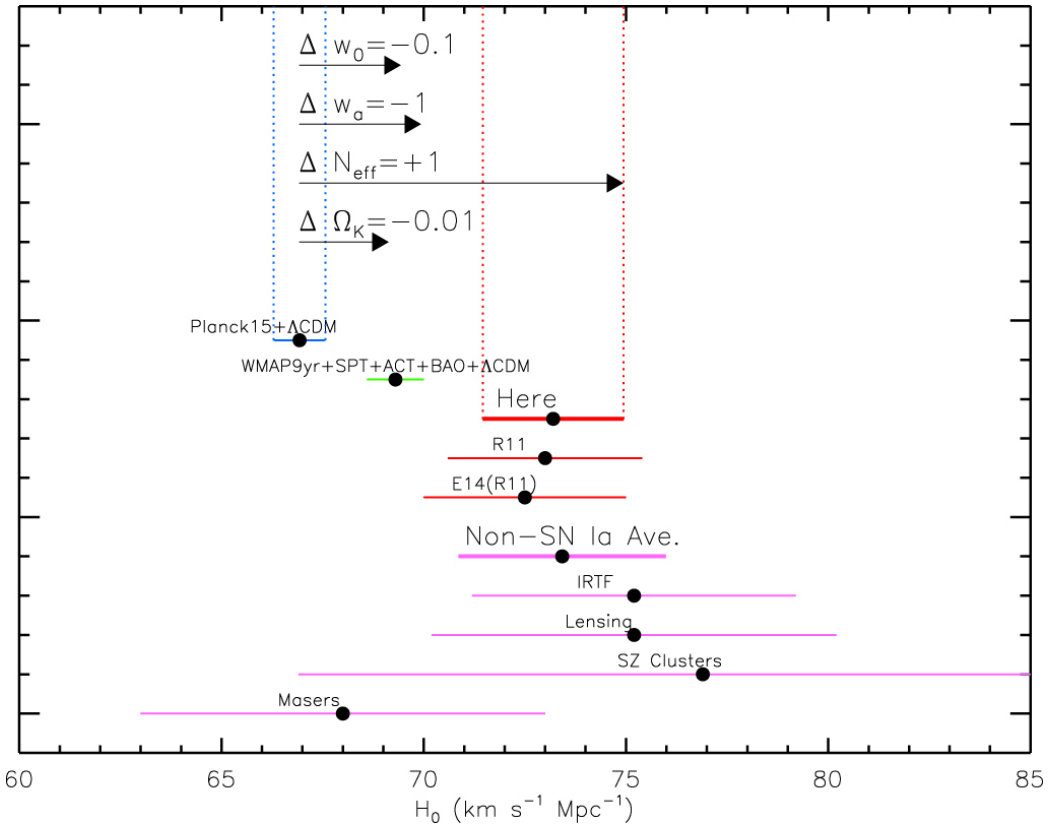

Figure 1.10 *From Riess et al. (2016). How the H_0 values inferred from early-times (Planck Collaboration et al., 2016a, blue error bar) compare with those inferred from late-times (Riess et al., 2016, 2011; Efstathiou, 2014, red error bars, 'Here', 'R11' and 'E14', respectively). Also shown are 4 SN-independent H_0 measures (Bonamente et al., 2006; Sorce et al., 2012; Suyu et al., 2013; Gao et al., 2016, magenta error bars) and a combined CMB-BAO constraint (Hinshaw et al., 2013, green error bar). The horizontal arrows indicate how the discrepancies in H_0 measures are reduced by considering various extensions (discussed in the text) to the standard model.*

(Section 1.5).

Where then, in the intricate theatre of cosmology research, does the work presented in this thesis enter? The answer is to better constrain the cosmology of the *non-linear Universe*, that which has evolved beyond the reach of our simple linear prescriptions for the properties of large-scale structure. To this end, I adopt the approach of weak gravitational lensing, a powerful probe of the virilised overdense structures at low redshift. Equipped with its theoretical framework, state-of-the-art numerical lensing simulations, and novel machine learning techniques, I develop new methodologies to test and constrain the standard cosmological model. This work ultimately aims to answer the question: how can we better exploit the cosmological information contained in the non-linear Universe?

Before I begin to contend with this question, in Chapter 2 I present a necessary in-depth discussion of weak lensing, in theory and in practice. Therein, I outline key systematic challenges facing this approach, and some of my own work in quantifying their impact on current observations from the Kilo-Degree Survey. In Chapter 3, I develop a novel "clipping" formalism, which enhances cosmic shear constraints by mitigating the barriers typically imposed by non-linearities. In Chapter 4, I present the cutting-edge *cosmo*-SLICS N-body simulations, and construct a tailored Gaussian process emulator, allowing the user to exploit these mocks in better understanding observational data. Chapter 5 sees further headway in this machine learning approach. By combining with new advances in theoretical modelling, I cultivate a general method to predict the non-linear matter power spectrum in *arbitrary* models of cosmology. Finally, in Chapter 6 I present my conclusions, and outline how this doctoral work may ultimately aid in resolving the cosmological tension with the next generation of surveys.

2

Weak Lensing in Theory and Practice

In this Chapter, I continue the discussion of weak gravitational lensing from Section 1.3.3, elaborating on both its theoretical and observational facets. The general relativistic description of this phenomenon is that photons travel on null geodesics within a distorted spacetime (e.g. the FLRW metric given in equation 1.2). I begin by presenting a simpler and commonly-used approximation for this picture, termed *gravitational lens theory*, in which an analogy is drawn between gravitational lensing and refraction via an optical lens. I will primarily adopt the notation of Bartelmann & Schneider (2001), although other excellent reviews are also available (see, e.g., Schneider, 2005; Kilbinger, 2015; Bartelmann & Maturi, 2017; Mandelbaum, 2018).

2.1 Gravitational lens theory

Figure 2.1 illustrates the core principle underpinning gravitational lens theory. Light from a source at distance D_s from the observer is deflected upon reaching a mass contained in the lens plane, at a distance D_d. The localised deflection of the light within the lens plane is valid assuming the mass in concentrated within

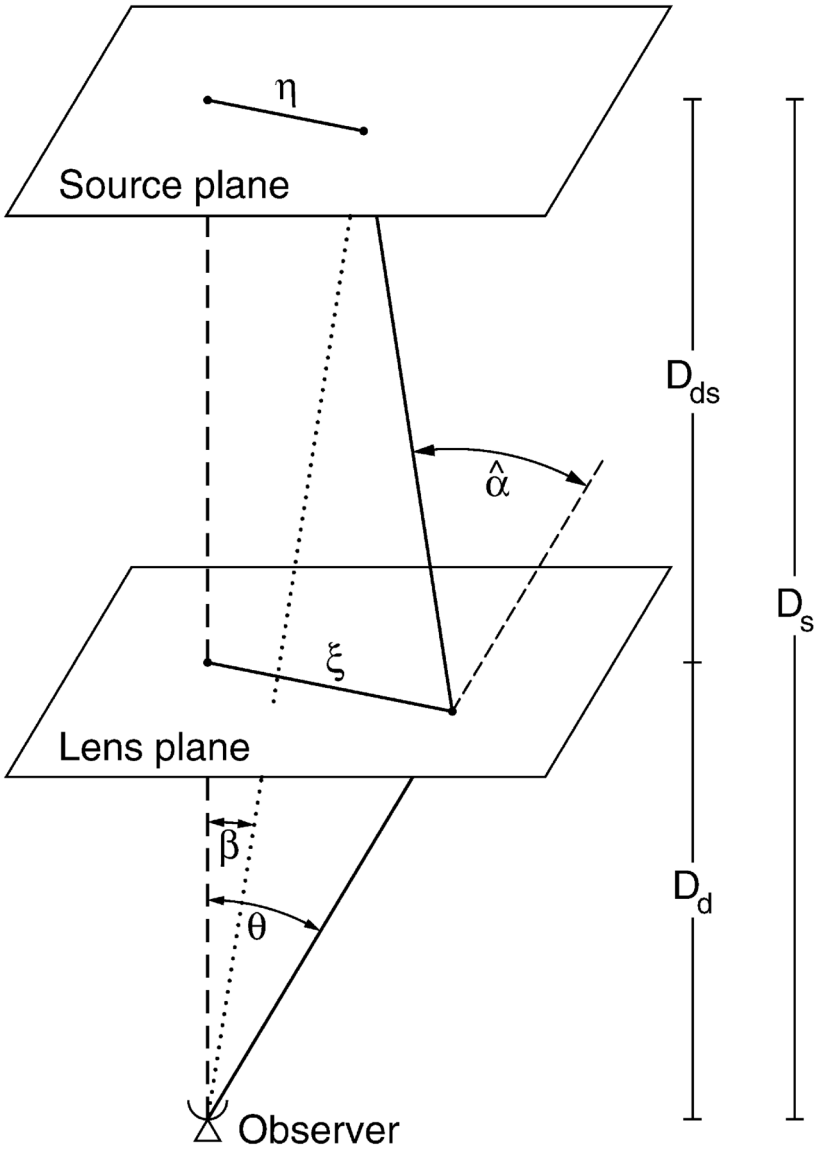

Figure 2.1 *From Bartelmann & Schneider (2001). Sketch of a gravitational lens system. Light from the source plane is deflected at the lens plane, causing the angle at which a source is observed to deviate from β (in the absence of lensing) to θ.*

a region that is small compared with the distance between the source and the deflector, $D_{\rm ds}$, and the distance between the deflector and the observer, as is the case with lensing by galaxy clusters. All distances here are angular diameter distances (see equation 1.22). The deflection causes the observed angular position of a source to change from $\boldsymbol{\beta}$, where it would be seen in the absence of lensing, to $\boldsymbol{\theta}$, in a manner given by the *lensing equation*.

From Figure 2.1 we see the relations between $\boldsymbol{\beta}, \boldsymbol{\theta}$ and the deflection angle, $\widehat{\boldsymbol{\alpha}}$, is given by,

$$\boldsymbol{\beta} D_{\rm s} + \widehat{\boldsymbol{\alpha}} D_{\rm ds} = \boldsymbol{\theta} D_{\rm s}. \tag{2.1}$$

By defining $\boldsymbol{\alpha} = \widehat{\boldsymbol{\alpha}} D_{\rm ds}/D_{\rm s}$, we obtain the lensing equation,

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}. \tag{2.2}$$

In GR, if the light impacts the lens plane at a transverse distance, ξ, which is much larger than the *Schwartzchild radius* of a deflector with mass M, then the deflection angle is given by,

$$\widehat{\boldsymbol{\alpha}} = \frac{4GM}{c^2 \xi}\,, \tag{2.3}$$

which is precisely twice as large as the Newtonian prediction and the very quantity Eddington and his crew sought to measure on their expedition in 1919. This result can be obtained by solving the Euler-Lagrange equations, which are second-order partial differential equations, for the radial and azimuthal trajectory of a photon within a spacetime geometry described by the *Schwartzchild metric* (Schwarzschild, 1916). One can subsequently obtain an expression for $r(\phi)$ and by taking the limit of very large distances from the point mass, one arrives at the quoted result for the deflection angle. Under this condition, $|\widehat{\boldsymbol{\alpha}}| \ll 1$.

If the strength of the gravitational field generated by the mass is weak, then the Einstein Field Equations (equation 1.1) can be linearised, such that the total deflection angle is given by the vectorial sum of the deflections caused by individual point masses. Consider dividing the 3-dimensional mass distribution with density $\rho(\boldsymbol{r})$ into volume elements, dV, with mass, $dm = \rho(\boldsymbol{r})dV$. A light ray approaching the mass distribution along the z-axis[1] with coordinate $(\boldsymbol{\xi}, z)$, where $\boldsymbol{\xi}$ is its 2-dimensional position on the plane of the sky, will encounter a mass element, situated at coordinate $(\boldsymbol{\xi}', z')$, with impact parameter $\boldsymbol{\xi} - \boldsymbol{\xi}'$. The light ray will be deflected in this case, but if the distortion to the trajectory is small

[1] z here denotes a radial distance coordinate itself, rather than redshift.

compared to the scales over which the mass distribution varies significantly, then the light's path can be approximated as a straight line. When applied to nuclear and atomic physics this is called the *Born approximation*. The total deflection angle is then given by the sum of the small deflections at each mass element,

$$\begin{aligned}\widehat{\alpha}(\boldsymbol{\xi}) &= \frac{4G}{c^2}\sum \frac{\boldsymbol{\xi}-\boldsymbol{\xi}'}{|\boldsymbol{\xi}-\boldsymbol{\xi}'|^2}\, dm(\boldsymbol{\xi}', z') \\ &= \frac{4G}{c^2}\int d^2\boldsymbol{\xi}' \frac{\boldsymbol{\xi}-\boldsymbol{\xi}'}{|\boldsymbol{\xi}-\boldsymbol{\xi}'|^2}\int dz' \rho(\boldsymbol{\xi}', z').\end{aligned} \tag{2.4}$$

Introducing the surface mass density of the lens plane,

$$\Sigma(\boldsymbol{\xi}) = \int dz' \rho(\boldsymbol{\xi}', z')\,, \tag{2.5}$$

allows us to write the two-dimensional deflection angle vector as

$$\widehat{\alpha}(\boldsymbol{\xi}) = \frac{4G}{c^2}\int d^2\boldsymbol{\xi}' \frac{\boldsymbol{\xi}-\boldsymbol{\xi}'}{|\boldsymbol{\xi}-\boldsymbol{\xi}'|^2}\,\Sigma(\boldsymbol{\xi}'). \tag{2.6}$$

If we consider the case of a circularly symmetric lens with a constant surface mass density Σ, this expression for the reduced deflection angle at radius $\boldsymbol{r}$ becomes

$$\widehat{\alpha}(\boldsymbol{r}) = \frac{4G}{c^2|\boldsymbol{r}|}\left(\Sigma\pi\boldsymbol{r}.\boldsymbol{r}\right). \tag{2.7}$$

Converting this to an expression on $\boldsymbol{\alpha} = \widehat{\boldsymbol{\alpha}}D_{\rm ds}/D_{\rm s}$ and recognising that $\boldsymbol{r} = D_{\rm d}\boldsymbol{\theta}$, yields

$$\alpha(\boldsymbol{\theta}) = \frac{4\pi G}{c^2}\,\frac{D_{\rm d}D_{\rm ds}}{D_{\rm s}}\Sigma\,\boldsymbol{\theta} = \frac{\Sigma}{\Sigma_{\rm cr}}\boldsymbol{\theta}. \tag{2.8}$$

Should the surface mass density of the lens equal the *critical surface density*, defined as

$$\Sigma_{\rm cr} = \frac{c^2}{4\pi G}\frac{D_{\rm s}}{D_{\rm d}D_{\rm ds}}\,, \tag{2.9}$$

$\boldsymbol{\alpha} = \boldsymbol{\theta}$ and the observer sees multiple images of the source blended into an Einstein ring, as illustrated by Figure 1.8. $\Sigma_{\rm cr}$, which depends on the redshifts of the source and lens, thus marks an approximate boundary between the strong and weak lensing regimes. Where $\Sigma \geq \Sigma_{\rm cr}$, the lensing equation gives multiple solutions for the observed angular position, $\boldsymbol{\theta}$, of a source at true position $\boldsymbol{\beta}$, and as such it is a strongly lensed system. On the other hand, if $\Sigma \ll \Sigma_{\rm cr}$, a well-

defined single solution for $\boldsymbol{\theta}$ exists for a given $\boldsymbol{\beta}$, and the lensing effect is weak. This boundary can alternatively be parameterised in terms of the dimensionless surface mass density or *convergence*, given by,

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(D_{\rm d}\boldsymbol{\theta})}{\Sigma_{\rm cr}}\,, \tag{2.10}$$

with $\kappa \ll 1$ and $\kappa \geq 1$ corresponding to weakly or strongly lensed systems respectively.

The definition of the convergence can be used to rewrite the expression for the deflection angle, given in equation 2.6, as

$$\alpha(\boldsymbol{\theta}) = \frac{1}{\pi}\int d^2\boldsymbol{\theta}'\,\kappa(\boldsymbol{\theta}')\frac{\boldsymbol{\theta}-\boldsymbol{\theta}'}{|\boldsymbol{\theta}-\boldsymbol{\theta}'|^2}\,, \tag{2.11}$$

where we have used $\boldsymbol{\xi} = D_{\rm d}\boldsymbol{\theta}$ to recast the integration over the impact parameter to one over the observation angle, $\boldsymbol{\theta}$. Noting that $\nabla_x \ln \boldsymbol{x} = \boldsymbol{x}/|\boldsymbol{x}|^2$ allows us to recognise the deflection angle as the gradient of the *lensing potential* (also called the *deflection potential*),

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi}\int d^2\boldsymbol{\theta}'\,\kappa(\boldsymbol{\theta}')\ln|\boldsymbol{\theta}-\boldsymbol{\theta}'|. \tag{2.12}$$

i.e, $\boldsymbol{\alpha} = \nabla_\theta\boldsymbol{\psi}$. This potential is the 2-dimensional analogue of the 3D Newtonian potential of the lens, Φ, and is related via

$$\psi(\boldsymbol{\xi}) = \frac{2}{c^2}\frac{D_{\rm ds}}{D_{\rm d}D_{\rm s}}\int \Phi(\boldsymbol{\xi}, z)\,dz. \tag{2.13}$$

By applying the Laplacian to this expression, and substituting in the Poisson equation, $\nabla^2_\xi\Phi(\boldsymbol{\xi}, z) = 4\pi G\rho(\boldsymbol{\xi}, z)$, we obtain

$$\nabla^2_\xi\psi(\boldsymbol{\xi}) = \frac{2}{c^2}\frac{D_{\rm ds}}{D_{\rm d}D_{\rm s}}\,4\pi G\,\Sigma(\boldsymbol{\xi})\,, \tag{2.14}$$

where we have additionally used the definition of the surface density given in equation 2.5. Finally, using our expressions for the convergence and critical surface density (equations 2.10 and 2.9), and noting that $\nabla^2_\xi = D^2_{\rm d}\nabla^2_\theta$, yields the 2-dimensional analogue of the Poisson equation,

$$\nabla^2_\theta\psi(\boldsymbol{\theta}) = 2\kappa(\boldsymbol{\theta}). \tag{2.15}$$

2.2 Magnification and shear

Liouville's theorem, and the fact that no photons are emitted or absorbed in lensing, means that the surface brightness of galaxies is conserved in this deflection process (Kristian & Sachs, 1966). Consequently, an apparent increase in the size of a galaxy via lensing corresponds to *magnification*, μ, given by

$$\mu = \frac{\text{image area}}{\text{source area}} = \frac{\delta\theta^2}{\delta\beta^2}. \tag{2.16}$$

If the angular size of the source is much smaller than the scales over which the convergence changes on the sky, then the mapping between the source and the lensed image is described by the lensing Jacobian,

$$\begin{aligned} A_{ij} =& \frac{\delta\beta_i}{\delta\theta_j} = \frac{\delta}{\delta\theta_j}\left[\theta_i - \alpha_i(\boldsymbol{\alpha})\right] , \\ =& \delta_{ij} - \frac{\delta\alpha_i(\boldsymbol{\theta})}{\delta\theta_j} , \\ =& \delta_{ij} - \frac{\delta^2\psi(\boldsymbol{\theta})}{\delta\theta_i\delta\theta_j} , \end{aligned} \tag{2.17}$$

where we have replaced the observation angle, $\alpha_i(\boldsymbol{\theta})$, with the gradient of the lensing potential. Defining the complex *shear* of magnitude $|\gamma|$ and orientation ϕ,

$$\gamma = \gamma_1 + i\gamma_2 = |\gamma| e^{2i\phi} , \tag{2.18}$$

where

$$\begin{aligned} \gamma_1 =& \frac{1}{2}\left(\frac{\delta^2\psi}{\delta\theta_1^2} - \frac{\delta^2\psi}{\delta\theta_2^2}\right) = |\gamma|\cos 2\phi , \\ \gamma_2 =& \frac{\delta^2\psi}{\delta\theta_1\delta\theta_2} = \frac{\delta^2\psi}{\delta\theta_2\delta\theta_1} = |\gamma|\sin 2\phi. \end{aligned} \tag{2.19}$$

Armed with these expressions for the shear and convergence (see equation 2.15), we can thus write the full lensing Jacobian matrix as,

$$A = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}. \tag{2.20}$$

The magnification is the result of the isotropic focussing of light by the local matter density, described by the convergence, κ, as well as the anisotropic focussing caused by the shear, γ. Mathematically this is given by the inverse

of the determinant of A,

$$\mu = \frac{1}{\det[A]} = \frac{1}{(1-\kappa)^2 + \gamma^2}. \tag{2.21}$$

The distortion in the image, on the other hand, is related to the trace-free part of the Jacobian, and is determined entirely by the shear. Using the definition of the *reduced shear*,

$$g(\boldsymbol{\theta}) = \frac{\gamma(\boldsymbol{\theta})}{1-\kappa(\boldsymbol{\theta})}\,, \tag{2.22}$$

the Jacobian is often written in the form,

$$A = (1-\kappa)\begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix}. \tag{2.23}$$

2.3 Shape Measurement

Rather than the shear, convergence or the lensing potential, it is the *shapes*, or *ellipticities* of galaxies which are directly accessible to the observer. For an isolated galaxy, its shape can be gauged with the quadrupole moment, Q_{ij}, of its surface brightness profile, $I(\boldsymbol{\theta})$, given by,

$$Q_{ij} = \frac{\int d^2\boldsymbol{\theta}\, w[I(\boldsymbol{\theta})]\, I(\boldsymbol{\theta})\, \theta_i \theta_j}{\int d^2\boldsymbol{\theta}\, w[I(\boldsymbol{\theta})]\, I(\boldsymbol{\theta})}\,, \tag{2.24}$$

where $w[I(\boldsymbol{\theta})]$ is a weight function, for example, a Heaviside step function centred on a limiting isophote. Expressions relating the complex ellipticity of the imaged galaxy, $e = e_1 + ie_2$, to the quadrupole moments, can vary subtly between authors. Bonnet & Mellier (1995), for example, present the definition,

$$\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \frac{1}{Q_{11}+Q_{22}+2(Q_{11}Q_{22}-Q_{12}^2)^{1/2}} \begin{pmatrix} Q_{11}-Q_{22} \\ 2Q_{12} \end{pmatrix}, \tag{2.25}$$

which, for a perfectly elliptical galaxy with a ratio of its semi-major and -minor axes, r, can be written

$$\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \frac{1-r}{1+r} \begin{pmatrix} \cos 2\phi \\ \sin 2\phi \end{pmatrix}. \tag{2.26}$$

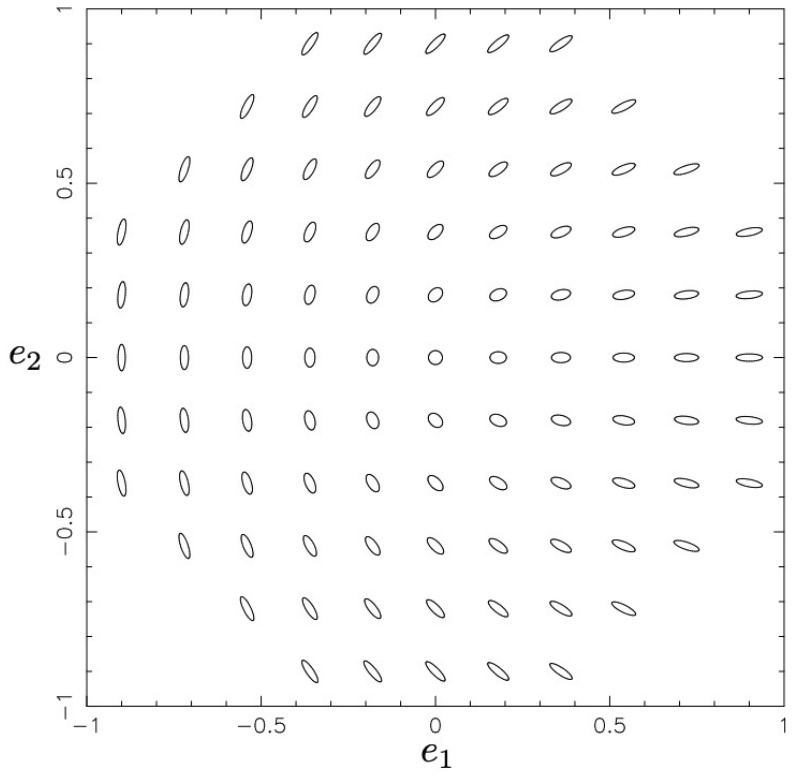

Figure 2.2 *The shapes and orientations of ellipses as a function of the ellipticity components expressed in equation 2.26 (adapted from Schneider, 2005).*

A visual illustration of how the shapes and orientations of galaxies vary as function of these ellipticity components, from Schneider (2005), is shown in Figure 2.2. Given this definition, the complex ellipticity of the source, $e^{\rm s}$, and the image are related by (Seitz & Schneider, 1997),

$$e^{\rm s} = \begin{cases} \frac{e-g}{1-g^*e}, & (|g| \leq 1) \\ \frac{1-ge^*}{e^*-g^*}, & (|g| \geq 1) \end{cases}, \tag{2.27}$$

where * denotes the complex conjugate. This expression can be reversed, such that the observed ellipticity is the subject, by swapping e and $e^{\rm s}$ and inverting the sign on g. In the weak lensing regime, $|g| \ll 1$ and $\kappa \ll 1$, such that $g \simeq \gamma$ and we obtain the simple linear relation, $e \simeq e^{\rm s} + \gamma$. In most cases, the intrinsic ellipticities of source galaxies can be assumed to be random and uncorrelated, such that $\langle e^{\rm s} \rangle = 0$ and $\langle e \rangle \simeq \langle \gamma \rangle$, although I discuss instances where this assumption is invalid in Section 2.6.2.

2.4 Weak lensing by large-scale structure

In Section 2.1, I presented gravitational lens theory, a simplified picture featuring a thin lens, where it is assumed that the angle between the incoming and deflected path is small compared to the scales over which the mass distribution of the lens varies. For lensing by large-scale structure, also called *cosmological lensing*, or cosmic shear, the mass distribution extends all the way from the observer to the source and this assumption is no longer valid. Thus, a more sophisticated approach must be adopted.

2.4.1 The deflection angle, α

The propagation equation describes how the comoving separation of two neighbouring light rays, $\boldsymbol{x} = \boldsymbol{\xi}/a$, changes with radial distance from the observer, w,

$$\frac{d^2\boldsymbol{x}}{dw^2} + K\boldsymbol{x} = 0\,, \tag{2.28}$$

where K is the spatial curvature, first seen in equation 1.4[2]. Solving this equation in unperturbed Minkowski spacetime (i.e. with no mass) with the boundary condition that the light rays converge with the propagation angle $\boldsymbol{\theta}$ at the observer ($\frac{d\boldsymbol{x}}{dw}\big|_{w=0} = \boldsymbol{\theta}$), yields (Bartelmann & Schneider, 2001),

$$\boldsymbol{x}(\boldsymbol{\theta}, w) = f_{\rm K}(w)\boldsymbol{\theta}\,, \tag{2.29}$$

where $f_{\rm K}(w)$ is the curvature-dependent comoving angular diameter distance out to radial distance w, defined in equation 1.22[3].

In the case of weakly perturbed Minkowski spacetime (i.e. where mass is present), the light is deflected according to

$$\frac{d^2\boldsymbol{x}}{dw^2} = -\frac{2}{c^2}\nabla_x\Phi\,, \tag{2.30}$$

where Φ is the Newtonian potential and $\nabla_x = \nabla_\theta/f_{\rm K}(w)$ is the transverse comoving gradient operator. The solution for the comoving distance between

[2]The curvature is rebranded in upper case here, to avoid confusion with the wave number, k.

[3]The different notation used for this quantity here follows the convention of various weak lensing reviews.

light rays propagating through such a space is given by

$$\boldsymbol{x}(\boldsymbol{\theta}, w) = f_{\rm K}(w)\boldsymbol{\theta} - \frac{2}{c^2}\int_0^w dw'\, f_{\rm K}(w-w')\,\nabla_x\Phi[\boldsymbol{x}(\boldsymbol{\theta}, w'), w']\,, \tag{2.31}$$

which trivially reduces to the unperturbed result when $\Phi = 0$. The deflection induced by the mass is given by the difference between the two solutions,

$$f_{\rm K}(w)\alpha(\boldsymbol{\theta}, w) = f_{\rm K}(w)\boldsymbol{\theta} - \boldsymbol{x}(\boldsymbol{\theta}, w). \tag{2.32}$$

Hence, the deflection angle is given by

$$\boldsymbol{\alpha}(\boldsymbol{\theta}, w) = \frac{2}{c^2}\int_0^w dw'\,\frac{f_{\rm K}(w-w')}{f_{\rm K}(w)}\,\nabla_x\Phi[f_{\rm K}(w')\boldsymbol{\theta}, w']. \tag{2.33}$$

2.4.2 The effective convergence, $\kappa_{\rm eff}$

In Section 2.1 I used gravitational lens theory to express the surface mass density contained in the lens plane, i.e. the convergence, as half the gradient of the deflection angle. Employing the same relation for the extended masses in cosmological lensing, the *effective convergence* is given by

$$\begin{aligned}\kappa_{\rm eff}(\boldsymbol{\theta}, w) &= \frac{1}{2}\nabla_\theta\alpha(\boldsymbol{\theta}, w) \\ &= \frac{1}{c^2}\int_0^w dw'\,\frac{f_{\rm K}(w-w')f_{\rm K}(w')}{f_{\rm K}(w)}\,\nabla_x^2\Phi[f_{\rm K}(w')\boldsymbol{\theta}, w']\,,\end{aligned} \tag{2.34}$$

where the extra factor of $f_{\rm K}(w')$ is introduced from the fact that $\boldsymbol{x}(w') = f_{\rm K}(w')\boldsymbol{\theta}$, and hence $\nabla_\theta = f_{\rm K}(w')\nabla_x$.

Taking the average of Poisson's equation for the Newtonian potential gives $\nabla_\xi^2\bar{\Phi} = 4\pi G\bar{\rho}$. Converting to a gradient with respect to comoving coordinates, $\nabla_x = a\nabla_\xi$, and employing the relation $\bar{\rho} = a^{-3}\bar{\rho}_0$ for the matter dominated era, we obtain

$$\nabla_x^2\Phi = \frac{3H_0^2\Omega_{\rm m}}{2a}\delta\,, \tag{2.35}$$

where we have used used the definition of the mass energy density parameter, $\Omega_{\rm m}$, expressed in equation 1.11. Using this form of Poisson's equation, we can replace the Laplacian of the potential in the equation for the effective convergence, finding

$$\kappa_{\rm eff}(\boldsymbol{\theta}, w) = \frac{3H_0^2\Omega_{\rm m}}{2c^2}\int_0^w dw'\,\frac{f_{\rm K}(w-w')f_{\rm K}(w')}{f_{\rm K}(w)}\,\frac{\delta[f_{\rm K}(w')\boldsymbol{\theta}, w']}{a(w')}. \tag{2.36}$$

We can then obtain an averaged 2-dimensional projection of the convergence, by integrating radially the 3D quantity, $\kappa_{\rm eff}(\boldsymbol{\theta}, w)$, weighted by a source redshift, $\phi(w)$, out to the horizon distance, w_H,

$$\begin{aligned}\bar{\kappa}_{\rm eff}(\boldsymbol{\theta}) &= \int_0^{w_H} dw\, \phi(w)\, \kappa_{\rm eff}(\boldsymbol{\theta}, w) \\ &= \frac{3H_0^2\Omega_{\rm m}}{2c^2} \int_0^{w_H} dw\, \phi(w) \int_0^{w_H} dw'\, \frac{f_{\rm K}(w-w')f_{\rm K}(w')}{f_{\rm K}(w)}\, \frac{\delta[f_{\rm K}(w')\boldsymbol{\theta}, w']}{a(w')} \\ &= \frac{3H_0^2\Omega_{\rm m}}{2c^2} \int_0^{w_H} dw'\, g(w')\, f_{\rm K}(w') \frac{\delta[f_{\rm K}(w')\boldsymbol{\theta}, w']}{a(w')}\,,\end{aligned} \tag{2.37}$$

where $g(w)$ is the weighting function,

$$g(w') = \int_0^{w_H} dw\, \phi(w) \frac{f_{\rm K}(w-w')}{f_{\rm K}(w)}. \tag{2.38}$$

Using the lensing efficiency, introduced in equation 1.58,

$$q(w') = \frac{3H_0^2\Omega_{\rm m}}{2c^2} g(w')\, f_K(w')/a(w')\,, \tag{2.39}$$

the effective convergence can be written in the concise form,

$$\bar{\kappa}_{\rm eff}(\boldsymbol{\theta}) = \int_0^{w_H} dw'\, q(w')\, \delta[f_{\rm K}(w')\boldsymbol{\theta}, w']. \tag{2.40}$$

2.4.3 The effective convergence power spectrum, P_κ

The two-point correlation function of two effective convergence fields is given by

$$\begin{aligned}\xi(|\boldsymbol{\theta}|) &= \langle \kappa_{\rm eff,1}(\boldsymbol{\theta}) \kappa_{\rm eff,2}(\boldsymbol{\theta}) \rangle \\ &= \int dw\, q_1(w) \int dw'\, q_2(w')\, \langle \delta[f_{\rm K}(w)\boldsymbol{\theta}, w]\, \delta[f_{\rm K}(w')\boldsymbol{\theta}, w'] \rangle.\end{aligned} \tag{2.41}$$

By replacing the real-space 3D density fluctuation fields, $\delta[f_{\rm K}(w)\boldsymbol{\theta}, w]$ with the Fourier transforms of the k-space analogues, $\delta[\boldsymbol{k}, w]$ (see equation 1.30), yields

$$\begin{aligned}\xi(|\boldsymbol{\theta}|) = \int dw\, q_1(w) \int dw'\, q_2(w') \int \frac{d^3k}{(2\pi^3)} \int \frac{d^3k'}{(2\pi^3)} \\ \langle \delta[\boldsymbol{k}, w] \delta[\boldsymbol{k'}, w'] \rangle\, e^{-if_{\rm K}(w)\boldsymbol{\theta}.\boldsymbol{k}_\perp}\, e^{-if_{\rm K}(w')\boldsymbol{\theta'}.\boldsymbol{k}_\perp{}'}\, e^{-iwk_\parallel} e^{-iw'k_\parallel{}'}\,,\end{aligned} \tag{2.42}$$

where $\boldsymbol{k} = \{\boldsymbol{k}_\perp, k_\parallel\}$. Using the definition of the matter power spectrum from equation 1.29,

$$(2\pi)^3 P_\delta(k, w)\, \delta_D(\boldsymbol{k} - \boldsymbol{k'}) = \langle |\delta(\boldsymbol{k}, w)\delta^*(\boldsymbol{k'}, w')| \rangle \,, \tag{2.43}$$

assuming that the density fluctuations are small, such that correlations are non-zero only for $w \simeq w'$, and integrating over $k_\parallel$ gives

$$\xi(|\boldsymbol{\phi}|) = \int dw\, q_1(w) q_2(w) \int \frac{d^2 k_\perp}{(2\pi)^2}\, P_\delta(k, w) e^{i f_{\rm K}(w) \boldsymbol{\phi}.\boldsymbol{k}_\perp} \,, \tag{2.44}$$

where $\boldsymbol{\phi} = \boldsymbol{\theta'} - \boldsymbol{\theta}$. The convergence power spectrum and correlation function are related via a Fourier transform,

$$\begin{aligned} P_\kappa(\ell) &= \int d^2\phi\, e^{-i\boldsymbol{\ell}.\boldsymbol{\phi}} \xi(\phi) \\ &= \int dw\, q_1(w) q_2(w) \int \frac{d^2 k_\perp}{(2\pi)^2}\, P_\delta(|k|, w) \int d^2\phi\, e^{i f_{\rm K}(w) \boldsymbol{\phi}.(\boldsymbol{k}_\perp - \boldsymbol{\ell})} \,, \end{aligned} \tag{2.45}$$

which reduces to the expression relating the convergence and matter power spectra, first presented in equation 1.57,

$$P_\kappa(\ell) = \int dw\, \frac{q_1(w) q_2(w)}{f_{\rm K}(w)^2} P_\delta\left(\frac{\ell}{f_{\rm K}(w)}, w\right). \tag{2.46}$$

Here the first-order Limber approximation, $k = \ell / f_{\rm K}(w)$, has been used to relate power spectra of the 3D and projected 2D fields (Limber, 1953). However, the more accurate second-order approximation, $k = (\ell + 1/2)/f_{\rm K}(w)$, appearing in equation 1.57, is used for the current generation of weak lensing surveys (Kilbinger et al., 2017b).

2.5 Weak lensing estimators

2.5.1 Shear-shear correlations

The most pure implementation of weak lensing as a probe of cosmology, cosmic shear, utilises the connection between the statistical correlations of galaxy shapes and the matter power spectrum (see equation 1.56), as discussed in Section 1.3.4. A number of annoying systematics notwithstanding (see Section 2.6), the two-

point shear correlation functions, expressed in equation 1.55, are measured in a straight-forward way from the data. These statistics, therefore, have been used extensively in numerous cosmic shear analyses to date (Heymans et al., 2013; Hildebrandt et al., 2017; Troxel et al., 2017; Giblin et al., 2018; Hildebrandt et al., 2018; Hamana et al., 2019). However, other diagnostics are available, including shear-shear or convergence power spectra (Köhlinger et al., 2017; van Uitert et al., 2018; Hikage et al., 2018) or the convergence correlation function (Van Waerbeke et al., 2013).

No thesis on the prowess of cosmic shear would be complete without presenting the "cosmic banana" constraints on $\Omega_{\rm m}$ and σ_8 obtained via this method. These are shown in the upper panel of Figure 2.3 for the three major weak lensing surveys currently observing, the Kilo-Degree Survey (KiDS), the Dark Energy Survey (DES) and the Hyper-Suprime Cam (HSC) survey, relative to those derived from the CMB observations of Planck Collaboration et al. (2018). The lower panel shows the corresponding "cosmic avocado" constraints in the $S_8 = \sigma_8\sqrt{\Omega_{\rm m}/0.3}$ vs $\Omega_{\rm m}$ plane. The tightness in the constraints in the lower relative to the upper panel, is a consequence of the fact that S_8 changes in the direction orthogonal to the degeneracy direction of the banana. The moderate disagreement we see between the Planck and KiDS results is a visual illustration of the lensing-CMB tension discussed in Section 1.5.

2.5.2 Shear-position correlations

An alternative to measuring the correlations in galaxy shapes induced by cosmological lensing, is to measure the shape distortions induced by individual galaxies. The method of *galaxy-galaxy lensing* (GGL) seeks to detect the tangential alignment of background galaxies about foreground lenses (Valdes et al., 1983). As with cosmic shear, this signal is weak, and only extractable via the measurement of shear-position correlations over large galaxy ensembles. The first attempt at measuring this effect in data was by Tyson et al. (1984), with many successful detections having been made since (see, for example, Brainerd et al., 1996; DES Collaboration et al., 2017; van Uitert et al., 2018).

The shear defined by equation 2.18 corresponds to a Cartesian coordinate system. For measuring the GGL signal, however, it is more convenient to have components

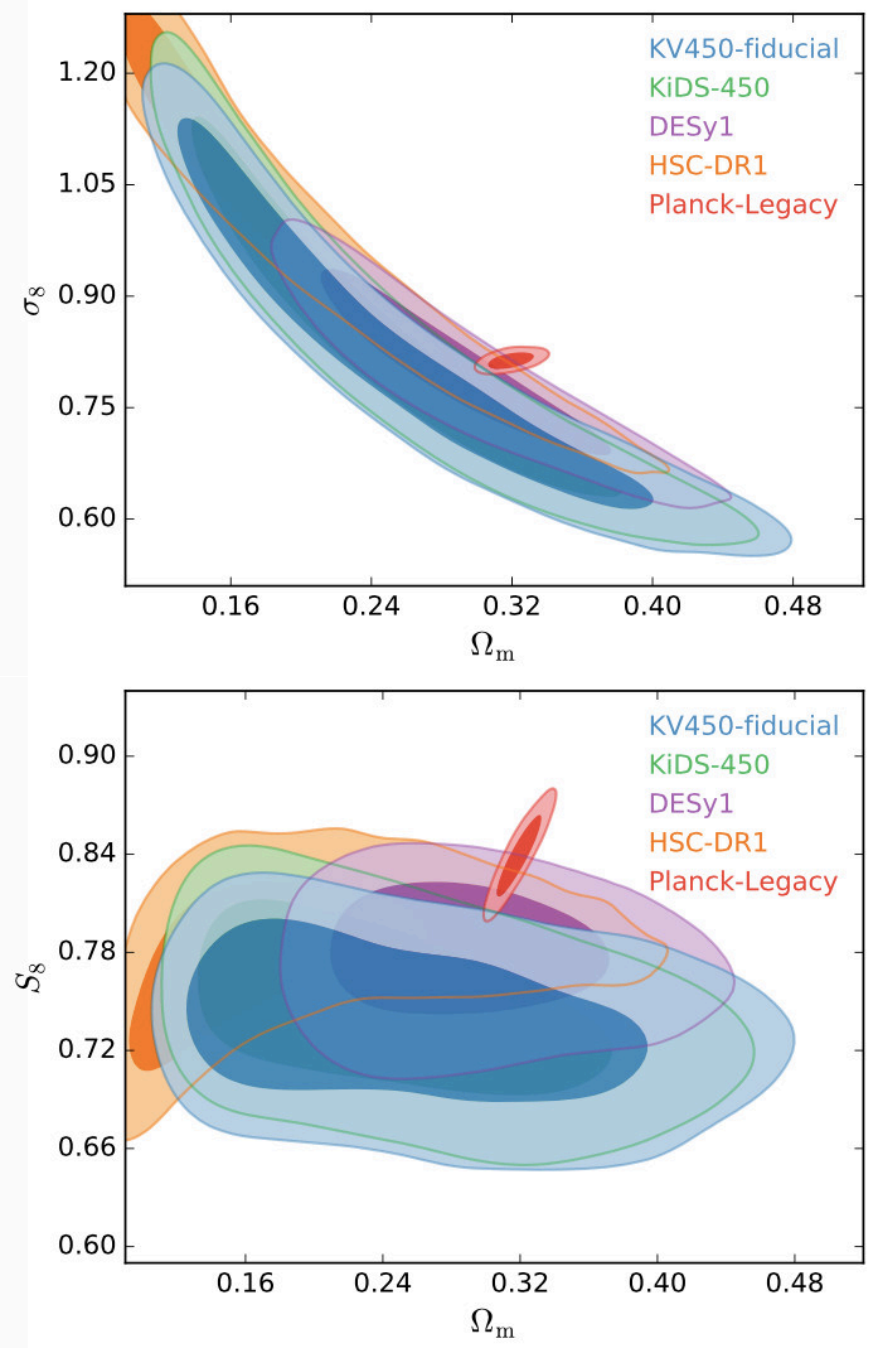

Figure 2.3 *From Hildebrandt et al. (2018). The cosmic shear constraints on* σ_8 *vs* Ω_m *(upper) and* $S_8 = \sigma_8/\sqrt{\Omega_m/0.3}$ *vs* Ω_m *(lower) from the three main weak lensing surveys currently operating, the Kilo-Degree Survey (Hildebrandt et al., 2017, 2018, "KiDS-450" and "KV450 fiducial", respectively), the Dark Energy Survey (Troxel et al., 2017, "DESy1") and the Hyper-Suprime Cam survey (Hikage et al., 2018, "HSC-DR1"), compared with the CMB constraints from Planck Collaboration et al. (2018).*

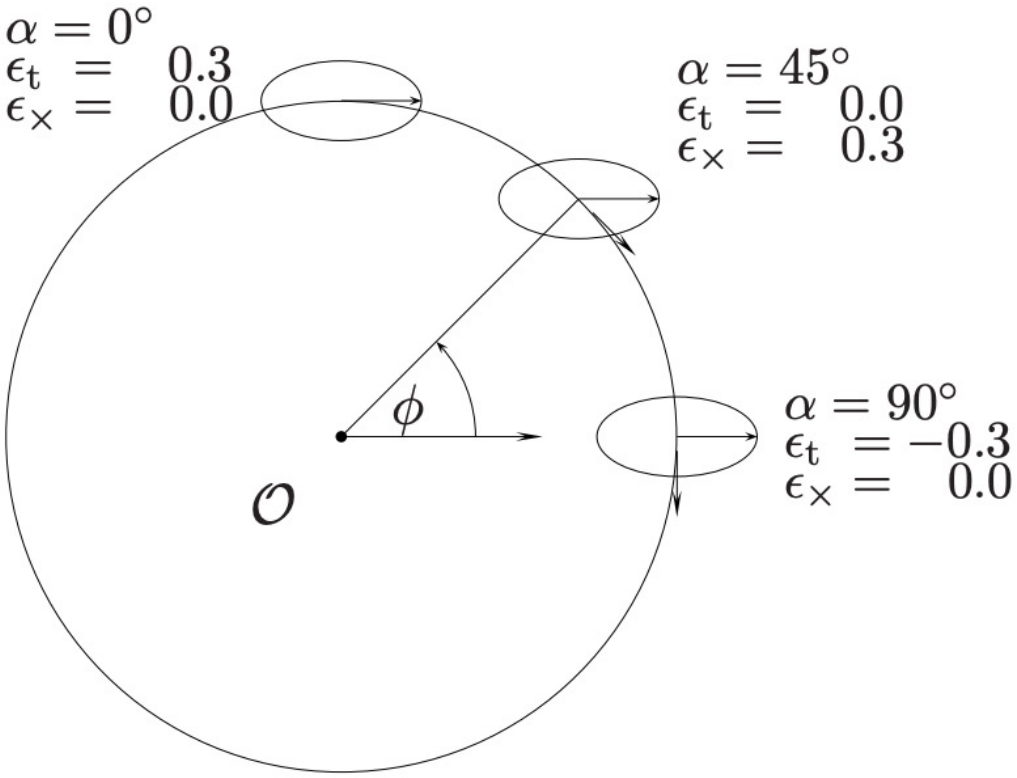

Figure 2.4 *An illustration of the γ_t and $\gamma_\times$ components (see equation 2.47) for a galaxy with $e_1 = 0.3, e_2 = 0$ from Schneider (2005). α is the angle between the semi-major axis of the source galaxy and the tangent to the lens. When $\alpha = 0°$, the galaxy has pure tangential alignment, positive γ_t and $\gamma_\times = 0$. When $\alpha = 90°$, γ_t is negative and the source is radially aligned with the lens. $\gamma_t = 0$ occurs when the source is at a 45° angle with the lens.*

defined as tangential to the lens, γ_t, and at a 45° angle to the tangent, $\gamma_\times$,

$$\begin{aligned} \gamma_t &= -\operatorname{Re}[\gamma e^{-2i\phi}] = -\gamma_1 \cos 2\phi - \gamma_2 \sin 2\phi \,, \\ \gamma_\times &= -\operatorname{Im}[\gamma e^{-2i\phi}] = \gamma_1 \sin 2\phi - \gamma_2 \cos 2\phi \,, \end{aligned} \tag{2.47}$$

where ϕ is measured relative to some axis of the lens. Figure 2.4 illustrates how the γ_t and $\gamma_\times$ components vary with the orientation of the source galaxy to the lens. The lensing of galaxies by overdensities, traced by other galaxies or galaxy clusters, will engender a tangential alignment and positive γ_t, whereas that which is induced by underdensities, i.e. voids, result in radial alignment and negative γ_t signal.

The usefulness of GGL, as far as cosmological constraints are concerned, stems from the connection between the tangential shear and the galaxy-to-matter

angular power spectrum, given by (van Uitert et al., 2018)

$$P^{\rm gm}(\ell) = b \int dw \, \frac{n(w)q(w)}{a(w)f_{\rm K}(w)} P_\delta \left(\frac{\ell + 1/2}{f_{\rm K}(w), w} \right) , \tag{2.48}$$

where b is the galaxy bias, $q(w)$ is the lensing efficiency (see equation 2.39) and $n(w)$ is the redshift distribution of the foreground lenses. Note the similarity between this quantity and the lensing power spectrum (equation 2.46). Cosmological constraints can be reached via the comparison of the theoretical galaxy-matter power spectrum and the *estimator* for this quantity measured from the data,

$$\widehat{P^{\rm gm}}(\ell) = 2\pi \int d\theta \, \theta \, \widehat{\gamma_{\rm t}}(\theta) J_2(\ell\theta) \, , \tag{2.49}$$

where $J_2(\ell\theta)$ is the second order Bessel function of the first kind, and

$$\widehat{\gamma_{\rm t}}(\theta) = \frac{\sum_i e_{{\rm t},i} w_i}{\sum_i w_i} \, , \tag{2.50}$$

is the signal estimated from the tangential ellipticities in the data, in bins of angular separation from the lens, θ, with the weights, w_i, depending on the signal-to-noise of the galaxy shape measurements.

Incidentally, the constraining power of cosmic shear is found to be augmented significantly by combining said method not only with GGL, but also with *position-position*, i.e. clustering, correlations, resulting in what is known as a "3×2-point" analysis (van Uitert et al., 2018; DES Collaboration et al., 2017). The theoretical prediction and estimator for the clustering power spectrum are given, respectively, by

$$P^{\rm gg}(\ell) = b^2 \int dw \, \frac{n^2(w)}{a(w)f_{\rm K}^2(w)} P_\delta \left(\frac{\ell + 1/2}{f_{\rm K}(w), w} \right) , \tag{2.51}$$

and

$$\widehat{P^{\rm gg}}(\ell) = 2\pi \int d\theta \, \theta \, \widehat{\xi^{\rm gg}}(\theta) \, J_2(\ell\theta). \tag{2.52}$$

Here $\widehat{\xi^{\rm gg}}(\theta)$ is the galaxy clustering correlation function estimated from the data,

$$\xi^{\rm gg}(\theta) = \frac{DD - 2DR + RR}{RR} \, , \tag{2.53}$$

where, for a given angular separation bin, DD is the number of galaxy pairs, RR is the number of pairs within an ensemble of randomly generated points, and DR the number of galaxy-random point pairs (Landy & Szalay, 1993).

Figure 2.5 shows how the constraints in the $\Omega_{\rm m} - \sigma_8$ plane from the first 450 deg^2

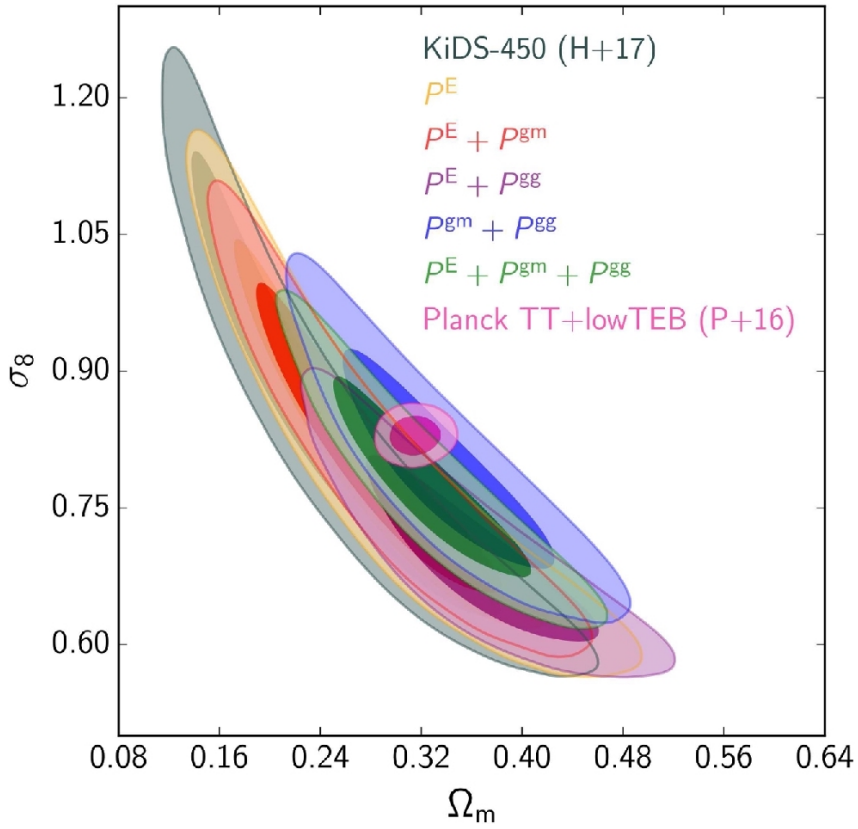

Figure 2.5 *From van Uitert et al. (2018). A comparison of the $\Omega_{\rm m} - \sigma_8$ constraints, for the KiDS-450 data, using various combinations of the lensing (shear-shear) power spectrum, $P^{\rm E}$, the galaxy-matter (shear-position) power spectrum, $P^{\rm gm}$, and the clustering (position-position), $P^{\rm gg}$ power spectrum. Also shown is the result from shear-shear correlation functions, (Hildebrandt et al., 2017, "KiDS-450 (H+17)"), and the Planck Collaboration et al. (2016a) constraints.*

of KiDS imaging (KiDS-450), vary with the application of different combinations of the shear-shear, shear-position and position-position estimators, relative to the Hildebrandt et al. (2017) and Planck Collaboration et al. (2016a) constraints. The full 3×2-point result (green contour) yields considerably improved statistical power over the cosmic shear analysis alone.

2.5.3 Aperture mass, $M_{\rm ap}$

The aperture mass statistic, introduced by Schneider (1996), detects the fluctuations in the surface mass density which are responsible for the weak lensing signal. This is defined as the convolution of the effective convergence (see equation 2.40) with an aperture, $U(\vartheta)$,

$$M_{\rm ap}(\theta) = \int_0^\theta d^2\boldsymbol{\vartheta}\, U(\vartheta)\, \bar{\kappa}_{\rm eff}(\boldsymbol{\vartheta})\,, \tag{2.54}$$

where $\vartheta \equiv |\boldsymbol{\vartheta}|$ and the aperture is endued with the property,

$$\int_0^\theta \vartheta\, d\vartheta\, U(\vartheta) = 0\,, \tag{2.55}$$

such that a finite $M_{\rm ap}$ is measured only given variations in $\bar{\kappa}_{\rm eff}$. By considering the relation between the convergence and the shear (discussed in Section 3.3.1), it can also be shown that the aperture mass is related to the 2-dimensional tangential shear field (equation 2.47), thusly

$$M_{\rm ap}(\boldsymbol{\theta}) = \int_0^\theta d^2\boldsymbol{\vartheta}\, \gamma_{\rm t}(\boldsymbol{\vartheta})\, Q(\vartheta)\,, \tag{2.56}$$

where $Q(\vartheta)$ is a second radially-symmetric weighting function, defined as

$$Q(\vartheta) = \frac{2}{\vartheta^2}\int_0^\vartheta \vartheta'\, d\vartheta'\, U(\vartheta') - U(\vartheta). \tag{2.57}$$

Schneider (1996) show that the optimal signal-to-noise measurement of the aperture mass is obtained when $Q(\vartheta)$ matches, as closely as possible, the profile of the tangential shear. As this is not known *a priori*, families of functions have been proposed in the literature (see, for example, Schneider et al., 1998; van Waerbeke, 1998; Jarvis et al., 2004).

The importance of the aperture mass, as far as cosmology is concerned, lies in its

variance, related to the lensing power spectrum via (Schneider et al., 1998),

$$\langle M_{\rm ap}^2 \rangle(\theta) = 2\pi \int_0^\infty \ell \, d\ell \, P_\kappa(\ell) \left[I_\ell(\ell\theta) \right]^2 \, , \tag{2.58}$$

where

$$I_\ell(\ell\theta) = \int_0^\theta \vartheta \, d\vartheta \, U(\vartheta) \, J_0(\ell\vartheta). \tag{2.59}$$

This can be converted into the following dependence on the shear correlation functions (for details, see Schneider et al., 2002a),

$$\langle M_{\rm ap}^2 \rangle(\theta) = \int_0^\infty \frac{d\vartheta \, \vartheta}{2\theta^2} \left[\xi_+(\vartheta) T_+ \left(\frac{\vartheta}{\theta} \right) + \xi_-(\vartheta) T_- \left(\frac{\vartheta}{\theta} \right) \right] \, , \tag{2.60}$$

where the filter functions, given by

$$\begin{aligned} T_+(x) &= 576 \int_0^\infty \frac{dt}{t^3} \, J_0(xt) \left[J_4(t) \right]^2 \, , \\ T_-(x) &= 576 \int_0^\infty \frac{dt}{t^3} \, J_4(xt) \left[J_4(t) \right]^2 \, , \end{aligned} \tag{2.61}$$

are finite for $x \leq 2$, limiting the range of integration to $0 \leq \vartheta \leq 2\theta$. Cosmological constraints, therefore, can be obtained by integrating the shear correlations measured from the data.

2.5.4 Peak statistics

As discussed in Section 1.3.2, the halo mass function is an effective means to constrain cosmological parameters. In analogy, the peaks in the projected surface mass density, observed via weak lensing, also probe the matter power spectrum. However, where analytical models exist for the former statistic (Press & Schechter, 1974; Sheth & Tormen, 2002), not so, in the case of the latter, since weak lensing peak projections may be composed of uncollapsed, yet highly non-linear structures in the cosmic web. Consequently, numerical simulations have been used by numerous authors, to explore the cosmological dependence of this probe (Dietrich & Hartlap, 2010; Kratochvil et al., 2010; Marian et al., 2012, 2013; Liu et al., 2014; Martinet et al., 2015), and indeed, to constrain observations (Liu et al., 2015, 2016; Kacprzak et al., 2016; Martinet et al., 2018; Shan et al., 2018).

The statistic measured in these analyses, generally, is the number of lensing peaks as a function of signal-to-noise, although the means of defining peaks differ

between authors. Previous studies primarily fall into two camps, depending on if peaks are identified as local maxima in the convergence field, which in the case of observations must be reconstructed from the shear (see Section 3.3.1), or identified in the shear field itself. In the latter case, peaks are associated with regions of particularly significant aperture mass (see Section 2.5.3). This has the benefit of avoiding systematics, which occur from reconstructing the projected surface mass density with observations featuring masks (a challenge relevant to my own work, as we shall see in 3.3.4). By considering the aperture mass centred on an angular coordinate $\boldsymbol{\theta}_0$, the estimator for the peak signal-to-noise is calculated from the tangential ellipticities in the data as (Martinet et al., 2018),

$$\frac{S}{N}(\boldsymbol{\theta}_0) = \frac{\sqrt{2}\sum_i Q(\boldsymbol{\theta}_i - \boldsymbol{\theta}_0)\, w(\boldsymbol{\theta}_i)\, e_{\rm t}(\boldsymbol{\theta}_i, \boldsymbol{\theta}_0)}{\sqrt{\sum_i w(\boldsymbol{\theta}_i)^2\, |e(\boldsymbol{\theta}_i)|^2\, Q^2(\boldsymbol{\theta}_i - \boldsymbol{\theta}_0)}} \frac{\sum_i w(\boldsymbol{\theta}_i)}{\sum_i w(\boldsymbol{\theta}_i)\, [1 + m(\boldsymbol{\theta}_i)]}\,, \tag{2.62}$$

where Q is the shear-space aperture mass filter, defined in equation 2.57, and m is a term to correct for multiplicative bias in the shear measurement, discussed further in Section 2.6. The comparison of this quantity, derived from observations and from a suite of numerical simulations with varying input cosmologies, in a likelihood analysis, is used to constrain the cosmological parameters of our Universe.

2.6 Systematics

This Chapter has, so far, focussed on the theoretical explanation for why constraints on cosmological parameters are obtainable, given perfect measurements of the shapes of weakly lensed galaxies. I now proceed to briefly discuss the myriad of systematic effects which, if left unaccounted for, confound our attempts to extract cosmological information from this method.

Many of these systematics give rise to either additive or multiplicative bias terms, c and m respectively, on the observed galaxy ellipticities, parameterised by $e_j^{\rm obs} = (1 + m_j)e_j^{\rm true} + c_j$ (Hildebrandt et al., 2017), where $e_j^{\rm true}$ is the j'th component of the ellipticity which would be measured in the absence of systematics. In general, the former arise where the systematic is uncorrelated with $e_j^{\rm true}$, such that the observed ellipticity has a contribution from the summation of individual 2-component distortion fields corresponding to each source of additive bias. The primary component of which is due to anisotropy in the point-spread function

(PSF, discussed in detail in this Section; Mandelbaum, 2018), but may also have contributions from artefacts caused by, e.g., cosmic rays, asteroids and non-linear response of the charge-coupled device (Fenech Conti et al., 2017). Multiplicative bias on the other hand is a consequence of contaminants to the shear signal which do correlate with galaxy ellipticity. This can arise due to imperfect PSF modelling, which serves to dilute the shear signal, or selecting galaxies for analysis in a way that correlates with their shear.

Discussion in this Section generally follows the format of Mandelbaum (2018), wherein systematics are split into two categories, depending on if they arise in the conversion of galaxy images to shear catalogues, or between the making of these catalogues and the extraction of scientific results. This division reflects where families of systematics are typically dealt with. Mitigations associated with the measurement of galaxies are primarily employed in the first stage, images-to-catalogues, although a second round of corrections may also be applied in the second stage, catalogues-to-science. Biases associated with an incomplete theoretical knowledge are exclusively attended to in the second stage however. When discussing the systematics in the images-to-catalogues stage, where possible, I will present results from my work in this area with the new release of 1000 deg^2 of KiDS imaging (K1000; Kuijken et al., 2019, Giblin et al., in prep.).

2.6.1 From images to catalogues

Image processing

The raw images taken at the telescope for a weak lensing analysis, feature a number of imperfections which require dedicated processing to achieve science-ready data products. In the course of my PhD, I performed such processing, using the THELI pipeline (Erben et al., 2005; Schirmer, 2013), for the new K1000 data release. As well as applying the general photometric (bias/dark/flat field/illumination) corrections to calibrate images for instrumental defects, and stacking exposures, this also consisted of processing specifically required of a weak lensing analysis.

The left panel of Figure 2.6 shows a single square-degree exposure contributing to the K1000 data, captured by the wide-field camera, OmegaCAM, on the VLT

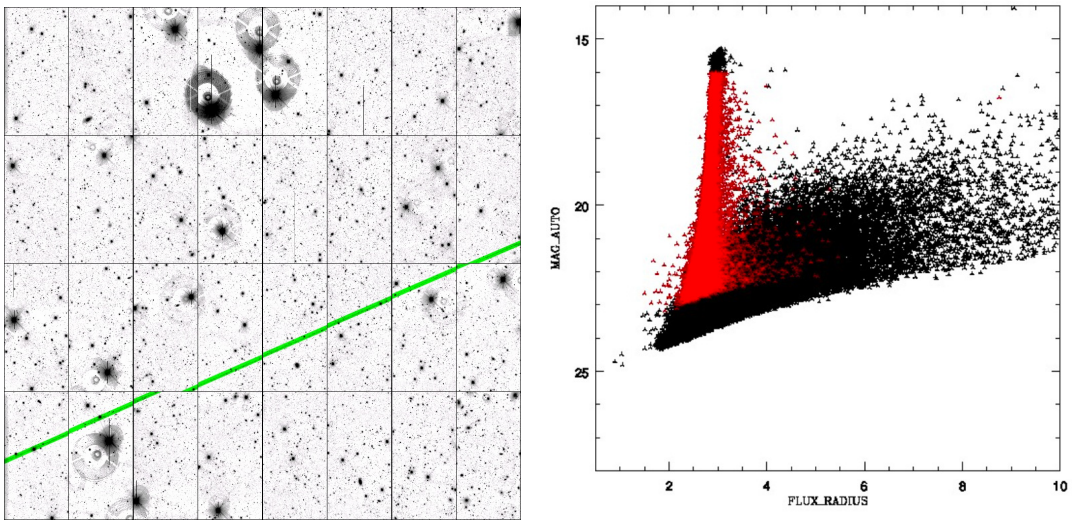

Figure 2.6 *Left panel: A 1 deg^2 r-band exposure, taken by the wide-field camera, OmegaCAM, installed on VST. This image contributed to the newest KiDS data release, for which I co-led the processing and quality control. The black disks and diagonal green line indicate, respectively, regions of bright foreground stars and where a satellite or plane crossed the image during the exposure. Both of these features require masking. Right panel: The MAG_AUTO versus FLUX_RADIUS measured with* SExtractor*, for the sources in the five stacked exposures for this tile. A distinct population of sources with magnitudes apparently independent of size are identified as stars and highlighted by the red points. The remaining sources, shown in black, are designated galaxies. The reason for the absence of a sharp divide between the two classes of object is discussed in the text.*

Survey Telescope (VST). This image features very bright foreground stars, the light from which infringes upon efforts to measure background galaxy shapes. Also unhelpful to this endeavour, is the trail of a satellite or plane which inconsiderately crossed the imaged portion of the sky during the exposure. Both of these artefacts, therefore, require masking (the host pixels given null signal), as shown by the black disks and green diagonal line, respectively.

Processing these images via the THELI pipeline also consists of characterising the sources as either stars or galaxies, the latter to be used in shape measurement. The right panel of Figure 2.6 shows the magnitude-size distribution of objects detected with SExtractor (Bertin & Arnouts, 1996), across the five exposures of the portion of sky shown in the left panel. The very bright but relatively small sources are designated stars (red points) whereas generally dimmer, but larger

sources are identified as galaxies (black points)[4].

I also co-led the quality control for this new data set. This consists of visually inspecting the performance of the photometric corrections, star/galaxy classification, masking and measurement of the point-spread function (discussed later in this Section), executed by the THELI pipeline for each KiDS tile. Should spurious patterns or artefacts be present in the processed exposures/co-added (stacked) images, or should the check-plots, such as the one shown in Figure 2.6, deviate from expectation, then the tile is flagged for further investigation and reprocessing. In the small number of cases where the tile fails to achieve the desired performance in each of these tests, as judged by the examiners, the tile is omitted from further analysis. With at least two members of the collaboration examining the products of processing for each tile, a high-quality KiDS lensing data set is assembled.

Modelling the point-spread function

The most obvious challenge one must overcome in a weak lensing analysis, is the very fact that the signal is *weak*. As previously mentioned in Section 1.3.4, the gravitationally-induced shear in a given galaxy is but $\sim 1\%$ the magnitude of its intrinsic ellipticity. Further complicating attempts to detect this are a number of additional effects, illustrated by Figure 2.7. From left to right, the upper row shows how the intrinsic shape of a galaxy translates into the imperfect realisation we observe. The shear signal for this galaxy (exaggerated in the second panel), is diluted by: a) the blurring of the image caused by Earth's atmosphere (unless the telescope is in space) and the convolution of the galaxy's light profile with the telescope's aperture (third panel); b) the pixelisation of the image by the charge coupled device (CCD) detectors in the instrument's camera (fourth panel); and c) additional sources of noise, such as caused by the finite number of photons incident on each pixel (Poisson noise), and Gaussian noise due to detector defects (fifth panel).

How the light from a point source (e.g. a distant star) spreads out in response to each of these effects is described by a PSF. This is illustrated by the lower row in Figure 2.7. In practice, the atmospheric PSF, the PSF due to the imperfect

[4]One does not observe a sharp distinction between the two types of object here. This is because the quantities shown on the horizontal and vertical axes are proxies for those used in the actual classification, which are derived from the photometry after homogenising the point-spread function across the multiple bands (for more information, see Kuijken et al., 2019).

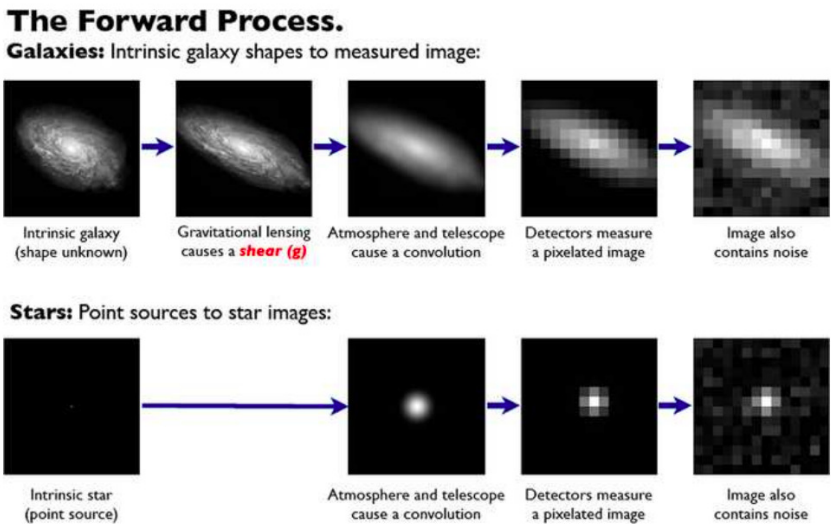

Figure 2.7 *From Bridle et al. (2009). Upper panel: how the intrinsic light profile of a galaxy is transformed, by shear, atmospheric and detector convolution, pixelisation and sources of noise, into the lower resolution realisation observed by a ground-based telescope. Lower panel: how these latter effects change the shape of a point-source (a star) used to estimate the PSF.*

response of the CCD pixels to incident photons, and that of the sub-optimal CCD charge diffusion, can all be considered as contributors to the overall effective PSF one must correct for. The PSF is modelled based on the light profile of calibration stars at various positions in the field of view. The model is then interpolated to other locations on the sky, such that galaxy shapes can be correct for this effect (see, for example, Bergé et al., 2012; Gentile et al., 2013; Kitching et al., 2013; Lu et al., 2017).

For the KiDS imaging, a PSF model, taking the form of 2-dimensional polynomial of order n, is fit across the whole field of view, with the coefficients up to order n_c given freedom to vary between each of the 32 CCD detectors in OmegaCAM. This allows for flexible spatial variation (including discontinuities) in the PSF. The total number of model coefficients per field of view is given by (Kuijken et al., 2015),

$$N_\mathrm{coeff} = \frac{1}{2}\left[(n+1)(n+2) + (N_\mathrm{D}-1)(n_\mathrm{c}+1)(n_\mathrm{c}+2)\right] , \tag{2.63}$$

where $N_\mathrm{D} = 32$ is the number of CCD detectors in OmegaCAM. Additionally, the flux and position of each calibration star are also allowed to vary in the PSF fitting, permitting further flexibility to the model. The total number of coefficients is large, but is sufficiently well constrained by the the number of data

points, equal to the number of pixels times the number of identified stars per exposure. The model is initially fit to stars identified in data from the Gaia Collaboration et al. (2018) down to the magnitude limit of this survey, and then refined using fainter stars in the KiDS imaging via the method in Kuijken et al. (2019).

Errors in estimating the size and shape of the PSF result in multiplicative and additive shear biases (see equation 2.62 for an example of how the former folds through to cosmological statistics). The PSF impacts shear estimation also via its spatial correlations. The effect these have on the ξ_+ shear correlation function (see equation 1.55) can be quantified using the ρ statistics (Rowe, 2010; Jarvis et al., 2016),

$$\begin{aligned}
\rho_1(\theta) &\equiv \langle \delta e^*_{\rm PSF}(\boldsymbol{x})\, \delta e_{\rm PSF}(\boldsymbol{x}+\boldsymbol{\theta})\rangle\,, \\
\rho_2(\theta) &\equiv \langle e^*_{\rm PSF}(\boldsymbol{x})\, \delta e_{\rm PSF}(\boldsymbol{x}+\boldsymbol{\theta})\rangle\,, \\
\rho_3(\theta) &\equiv \left\langle \left(e^*_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\boldsymbol{x})\, \left(e_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\boldsymbol{x}+\boldsymbol{\theta})\right\rangle, \\
\rho_4(\theta) &\equiv \left\langle \delta e^*_{\rm PSF}(\boldsymbol{x})\, \left(e_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\boldsymbol{x}+\boldsymbol{\theta})\right\rangle, \\
\rho_5(\theta) &\equiv \left\langle \delta e^*_{\rm PSF}(\boldsymbol{x})\, \left(e_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\boldsymbol{x}+\boldsymbol{\theta})\right\rangle,
\end{aligned} \tag{2.64}$$

where $e_{\rm PSF}$ and $\delta e_{\rm PSF} \equiv e_{\rm PSF} - e_{\rm model}$ are the ellipticity of the PSF and its residual, measured at the positions of stars, respectively. $T_{\rm PSF}$ is the size of the PSF, taken to be the trace of the second moments of the surface brightness profile, defined in equation 2.24, and $\delta T_{\rm PSF} \equiv T_{\rm PSF} - T_{\rm model}$. The correlations are measured with pairs of stars separated by the angle θ. The additive systematic error to ξ_+ caused by finite ρ contributions is given by (Jarvis et al., 2016),

$$\delta\xi_+(\theta) = \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}}\right\rangle^2 (\rho_1(\theta) + \rho_3(\theta) + \rho_4(\theta)) - \alpha \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}}\right\rangle (\rho_2(\theta) + \rho_5(\theta))\,, \tag{2.65}$$

where $T_{\rm gal}$ is the intrinsic galaxy size, unconvolved with the PSF. α is the amount of PSF "leakage", with contributions, a and b, which give rise to an additive shear bias,

$$c_i = a_i e_{{\rm model},i} + b_i \delta e_{{\rm PSF},i}\,, \tag{2.66}$$

where i denotes the component of the ellipticity. The first term is associated with the deconvolution of the PSF from the galaxy ellipticities, such that $a \neq 0$ if this

is done imperfectly. The second term is associated with how well the model fits the true effective PSF. A value of order -1 is expected for b, since PSF model errors will propagate into an error of the same magnitude, but opposite sign, in the shear (Paulin-Henriksson et al., 2008). These terms can be estimated from the gradient of the galaxy ellipticities, and the PSF residuals, both with respect to those of the PSF model (Zuntz et al., 2018).

Figure 2.8 shows the ρ statistics and the corresponding systematic bias, $\delta\xi_+$, measured from a 65 deg^2 subset of the K1000 survey footprint, assuming a very large amount of PSF leakage, $\alpha = 0.03$, arbitrarily selected as being ~ 30 times the additive bias in the previous data release, KiDS-450 (Hildebrandt et al., 2017). I calculate these correlations for four different PSF models, characterised by the polynomial orders $n : n_\mathrm{c}$ (see equation 2.63). The yellow band, included in the plot to guide the eye, shows $\pm 10\%$ of the weakest of the ξ_+ measurements from the KiDS tomographic bins mentioned in Section 1.3.4. This corresponds to the photometric redshift range $z_\mathrm{B} \in [0.1, 0.3]$. The error bars (too small to be seen in some cases) come from a jackknife resampling, whereby the field is divided into N_jk segments which are removed one-by-one, with the ρ statistics being calculated from the remaining $N_\mathrm{jk} - 1$ segments at each iteration.

Figure 2.8 shows that although some of the individual ρ-correlations for the various PSF models exceed the yellow band, showing 10% of the weakest ξ_+ measurement, the overall additive bias, $\delta\xi_+$, is much smaller, at the sub-per-cent level, for all models considered. This is an indication that the impact of the PSF is captured sufficiently well by all models to use the data for cosmic shear. These updated measurements of the ρ-statistics, however, suggest that the 3:1 model used to characterise the PSF with previous KiDS data releases (Kuijken et al., 2015), does not fit the data as well as the more complicated polynomials tested here. I therefore identify the 4:1 model as the most appropriate for applying the PSF correction to K1000, given it has the least number of coefficients, N_coeff, of the three remaining models which perform comparably.

A prudent sanity-check to evaluate the success of the PSF correction, is to examine the distribution of PSF ellipticities and residuals across the focal plane. Figure 2.9 shows the results of this test for the entire K1000 data set, assuming the 4:1 PSF model. For imperfect PSF modelling and deconvolution, coherent structure may be observable in the residual across the plane. Despite the prominent patterns in the PSF ellipticity, there are no obvious structures in the PSF residuals and the distribution is consistent with noise.

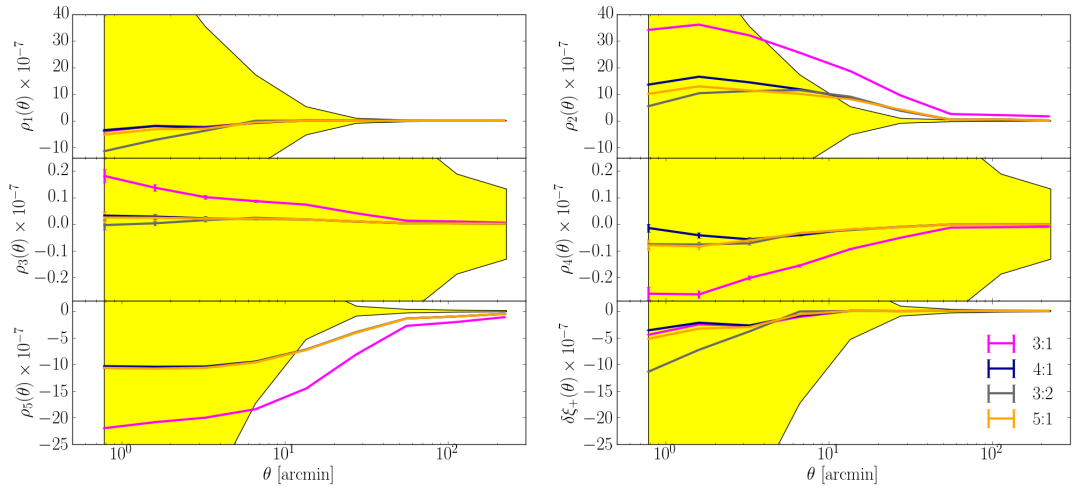

Figure 2.8 *From the upper-left to bottom right panel, the ρ statistics (see equation 2.64) and corresponding bias on ξ_+ (equation 2.65), measured from a 65 deg^2 subset of the K1000 data set, assuming a very high level of PSF leakage, $\alpha = 0.03$, arbitrarily selected. The various coloured lines are associated with different PSF models, defined by the polynomial orders $n : n_c$ (see equation 2.63). The yellow band, serving as a general guide rather than a requirement, shows $\pm 10\%$ of the weakest of the ξ_+ signals measured in each of the tomographic bins. This corresponds to the photometric redshift range $z_B \in [0.1, 0.3]$. The errors (too small to be seen in some cases) come from jackknife resampling. Note the changing scales in each panel.*

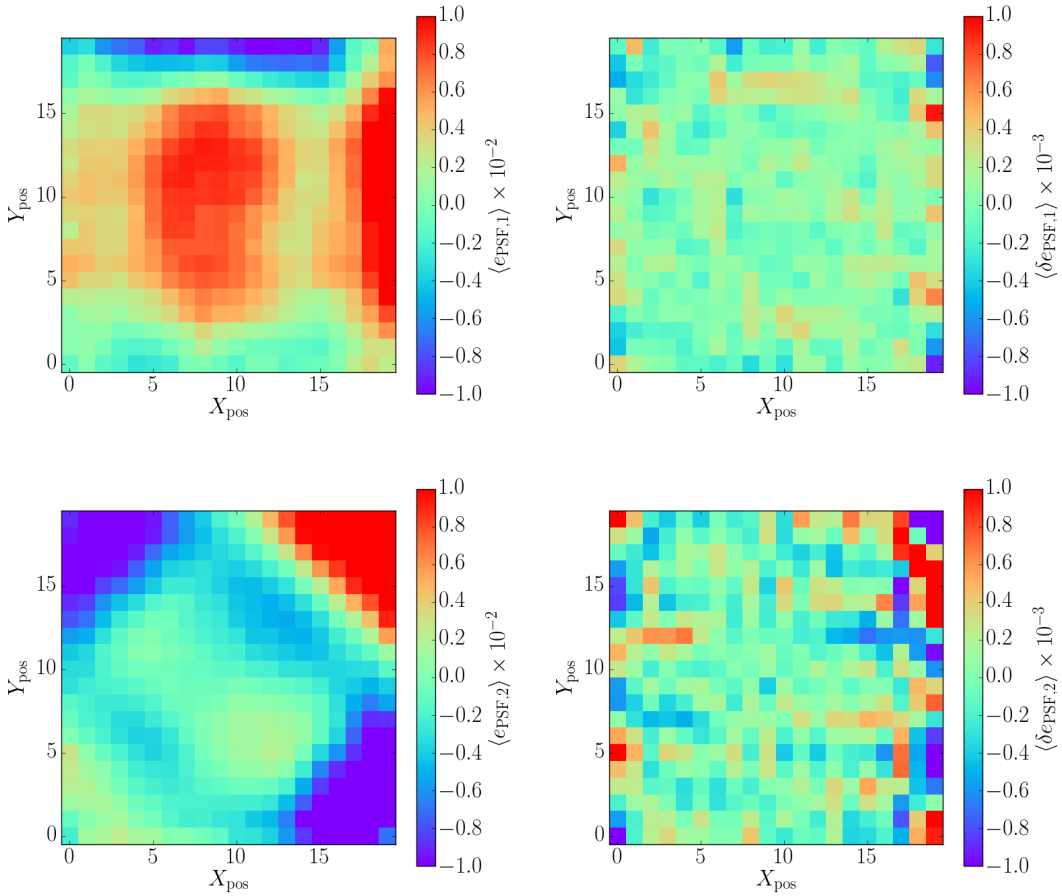

Figure 2.9 *The mean K1000 PSF ellipticity (left panels) and residual, assuming the PSF model with the 4:1 polynomial order (right panels; see equation 2.63), for the e_1 and e_2 components (upper/lower, respectively), binned by position on the focal plane. Note there is a difference in the colour scales between the left and right panels to better visualise the weaker residual PSF ellipticities.*

Shear estimation

Even given a perfect correction for the PSF, the diverse range of galaxy morphologies, combined with the noise and pixelisation associated with observations (see Figure 2.7), imply that measuring galaxy shapes, and thus estimating the shear, is a non-trivial task. Early methods consisted of simply calibrating the second moments of galactic surface brightness profiles (see Section 2.3) for those of the PSF (Kaiser et al., 1995; Luppino & Kaiser, 1997; Hoekstra et al., 1998). Later approaches included fitting parametric models to galaxy images (see Massey et al., 2007, for discussion), to going beyond a shape-per-galaxy formalism by constructing a Bayesian posterior over the shear field (Bernstein & Armstrong, 2014; Schneider et al., 2015).

*Lens*fit (Miller et al., 2013) and METACALIBRATION (Huff & Mandelbaum, 2017; Sheldon & Huff, 2017), utilised by KiDS and DES respectively, are examples of shape measurement algorithms employing different techniques. *Lens*fit models galaxy shapes as a combination of two Sersíc profiles - a disk and a bulge component - and fits for seven free parameters - ellipticity (2), position (2), size (1), flux (1) and bulge fraction (1) - in a Bayesian framework. The shape model is convolved with that of the local PSF, and a likelihood is constructed as a function of the two ellipticity components, marginalising out the five other nuisance parameters. This is modulated by the prior on the ellipticity components, which in turn is based on observations from the Hubble Space Telescope and the Sloan Digital Sky Survey, to produce the overall Bayesian posterior for the galaxy shape. This method involves dedicated image simulations to quantify and correct for multiplicative bias in the shear estimation (Fenech Conti et al., 2017; Kannawadi et al., 2019).

On the other hand, METACALIBRATION, exercised by DES on their Year 1 data release in Zuntz et al. (2018), recalibrates shape measurements using the *response* of the shear estimator, to the shear itself. Any estimator, in principle, can be used, and so DES employ the NGMIX Gaussian model-fitting software (Sheldon, 2015). METACALIBRATION determines the response by injecting the observed images with known shears, and measuring the change in the inferred galaxy ellipticities. The shear is estimated from the weighted average of the ellipticities, as usual, with the caveat that the ellipticities are additionally modulated by the inverse of the response matrix per galaxy. This method, therefore, uses the data to self-calibrate the shear estimation, negating the need for image simulations.

These approaches face a number of obstacles, such as noise bias, model bias and selection bias. The former, often dominant source of shear measurement bias, arises when the maximum-likelihood values of the per-galaxy ellipticities are used as the point-wise estimator for its shape. This is biased because the noise, which is inevitably present in observations, changes the shape of the likelihood surface (Bernstein & Jarvis, 2002; Hirata et al., 2004; Kacprzak et al., 2012; Melchior & Viola, 2012; Refregier et al., 2012). Methods to mitigate this effect have been developed in Miller et al. (2007) and Bernstein & Armstrong (2014). Model bias, on the other hand, is caused by the failure of the model to accurately describe galaxy shapes in the data (Voigt & Bridle, 2010). This is thought to be a negligible effect given the statistical power of current lensing surveys (Kacprzak et al., 2014). Finally, selection bias is caused by excluding galaxies from the sample based on properties which correlate with their intrinsic shapes or shear. If round galaxies are preferentially included in the analysis for example, as would be the case if signal-to-noise were the sole selection criterion, the lensing signal will be systematically underestimated. To minimise this, cuts made to the data can be informed by image simulations (Fenech Conti et al., 2017; Zuntz et al., 2018; Kannawadi et al., 2019). Although the systematic biases in these algorithms are sufficiently small to be used with current lensing data, they will no longer be tolerable given the statistical power of the next generation of surveys (Laureijs et al., 2011; LSST Dark Energy Science Collaboration, 2012, discussed further in Section 6.2).

Finally, a paradigm shift in shear estimation has been realised in more recent approaches. This school of thought recognises that the averaging of point-wise per-galaxy shape measurements is mathematically flawed, and instead aims to infer the ensemble Bayesian posterior over the shear field. To this end, Schneider et al. (2015) develop a hierarchical statistical framework, incorporating nuisance parameters describing intrinsic galaxy properties, and fitting parametric models to the data at the pixel level. This results in a mathematically consistent formalism, but incurs potentially considerable computational expense. The cost is mitigated somewhat in the method proposed by Bernstein & Armstrong (2014), developed further in Bernstein et al. (2016), where shear estimates are similarly obtained from a Bayesian posterior, by compressing the pixel data into a small number of moments in Fourier space. In the calculation, a prior on the noiseless distribution of target galaxy property parameters is required. This can be obtained by taking long integration exposures of a subset of the survey. Whilst these Bayesian approaches to shear estimation appear promising, they are still in

their relative infancy compared to per-galaxy model-fitting techniques. As such, further work is needed to demonstrate their performance with real data.

Photometric redshift estimation

Through equation 1.58, we see the dependence of the cosmic shear signal on the redshift distribution of source galaxies. In a perfect world, we would have highly accurate spectroscopic redshifts for the $\sim 10^7$ galaxies used for shear estimation in concurrent weak lensing surveys. In reality this is infeasible due to the task of measuring spectra being so time-intensive. Indeed the world's largest redshift survey, BOSS, (see Section 1.3.2) has collected spectra for relatively fewer galaxies ($\sim$1.2 million), and to lower depth than those used to measure cosmological lensing. To overcome this obstacle, the photometric redshifts ("photo-z's") must be estimated using the flux of sources in a small number of broad-band filters. The unbiased estimation of galaxy colours (equal to the difference of the magnitudes in different wavelength bands) is challenging in itself, given that the PSF likely varies between bands. This effect can be reduced with careful homogenisation of the PSF across the bands (Hildebrandt et al., 2012).

From the photometry, a photo-z (point-estimate or probability distribution) per galaxy is estimated. Methods to perform this calculation fall into two main classes: template-fitting and machine learning. With the former, an ensemble of template galaxy spectral energy distributions (SEDs) are used to fit the observed photometry. The photo-z is determined by the relative wavelength shift between the templates and data (see, for example, Benítez, 2000; Ilbert et al., 2006). A potential pitfall in this case, is the template ensemble may not sufficiently cover the range of observed galaxy types. Machine learning methods, on the other hand, constrain the unknown mapping between photometry and redshift based on a training sample (e.g., Collister & Lahav, 2003; Vanzella et al., 2004; Bonnett, 2015; Sadeh et al., 2016). A key challenge here is that the training set may be unrepresentative of the data. Both of these issues are a consequence of the the fact that we lack sufficiently representative spectra of galaxies to weak lensing depths (Newman et al., 2015). I return to the business of *calibrating* the photometric redshift distributions - a catalogues-to-science problem - in the next Section.

2.6.2 From catalogues to science

Photometric redshift calibration

The estimates of the photometric redshifts are used to sort source galaxies into tomographic bins. Template-fitting photo-z estimation methods, however, are not sufficiently accurate for inferred distributions to be used in producing the theoretical cosmic shear statistics (see equation 1.58) for comparison with the data[5] (Hildebrandt et al., 2010; Newman et al., 2015). Hence, spectroscopic redshift samples are required to either calibrate the photo-z distributions from these codes, or to serve as training sets in machine-learning based methods.

In the case of calibration, two options are available to us. The first is to reweight the spectroscopic calibration sample such that the colours and magnitudes match the photometric sample (Lima et al., 2008). The reweighted spectroscopic redshift distribution is then taken to be an unbiased estimate of the ensemble distribution of source galaxies. The second option is to cross-correlate the photometric redshift sample with a non-representative spectroscopic sample of equivalent or greater depth, and with large enough area and sampling rate for the cross-correlation to be determined (Mandelbaum, 2018). In this method, the excess angular clustering between the photometric and spectroscopic samples is used to infer the probability distribution function for the redshift within each tomographic bin (Newman, 2008; Benjamin et al., 2010; McQuinn & White, 2013; Ménard et al., 2013; Schmidt et al., 2013).

Both routes to calibrating photo-z distributions have been exercised by current weak lensing surveys (Hildebrandt et al., 2017; Hoyle et al., 2018; Hikage et al., 2018, KiDS, DES and HSC, respectively), although DES use high-resolution, 30-band COSMOS photo-z's, rather than spectroscopic redshifts, in calibrating their distributions estimated from the low-resolution photometry. Accurate photo-z calibration is one of the most challenging and important systematics within a weak lensing analysis. Indeed, recent papers have suggested that this factor alone, may be causing the discrepancy in the estimates of the amount of tension between CMB-inferred cosmological parameters, and those inferred in different cosmic shear analyses (Hildebrandt et al., 2018; Joudaki et al., 2019, see Section 1.5 for details.).

[5]Indeed, it has been shown that the commonly-used stacking of per-object photo-z probability distributions violates the definition of probability (Malz et al., in prep.).

Intrinsic alignments

The *intrinsic alignments* (IAs) of galaxy shapes due to gravitation, is a contaminant to the apparent alignments induced by lensing (Croft & Metzler, 2000; Heavens et al., 2000; Lee & Pen, 2000; Catelan et al., 2001). The difficulty in correcting for IAs lies in our incomplete theoretical understanding of this effect. Although (gravity-only) N-body simulations reveal correlations in the orientation of dark matter haloes, galaxy shapes are dependent on baryonic physics, a component absent from such simulations. Insights have been made from observations however, suggesting that the shapes of red galaxies exhibit large-scale alignments with those in the inner regions of the same haloes, whereas shape alignments for blue galaxies so far have yielded null detections (Mandelbaum et al., 2006; Hirata et al., 2007; Joachimi et al., 2011; Li et al., 2013; Singh et al., 2015). This understanding has facilitated the tailoring of N-body simulations to include red galaxies with realistic alignments (Schneider et al., 2012), although IAs have also been pursued with high-resolution, large-volume *hydrodynamic* simulations, which include the physics of baryons and galaxy formation (Velliscig et al., 2015; Tenneti et al., 2016; Chisari et al., 2017; Hilbert et al., 2017). The results in these cases are broadly in agreement with observations, although not all observed alignment trends are reproduced by the simulations.

Early attempts to mitigate the impact of IAs in weak lensing measurements, included down-weighting galaxy pairs with small 3-dimensional separations (King & Schneider, 2002; Heymans & Heavens, 2003). However, Hirata & Seljak (2004) pointed out that there are coherent (anti-)alignments even between the shapes of galaxies with large separations, due to their coupling to the large-scale tidal field. This, it turns out, is the dominant source of IA contamination to the weak lensing signal in observations, and it cannot be mitigated by removing or down-weighting galaxies at the same redshift (Mandelbaum, 2018).

In current weak lensing analyses, the impact of IAs are factored in at the level of the science-statistic. For example, the estimated shear correlation functions are modelled as having contributions from two terms in addition to the lensing-induced signal,

$$\widehat{\xi_\pm} = \xi_\pm + \xi_\pm^{\rm GI} + \xi_\pm^{\rm II}. \tag{2.67}$$

The 'gravitational-intrinsic' (GI), and 'intrinsic-intrinsic' (II), terms here, are correlations caused by a linear and quadratic coupling of the intrinsic galaxy shapes to the tidal field (Hirata & Seljak, 2004). These two terms, respectively,

cause a large-scale suppression and a small-scale boost in the measured $\widehat{\xi_{\pm}}$, with the former contribution being the stronger, unaffected by the purging of same-redshift galaxy pairs. These terms are calculated from their corresponding power spectra, given by (Bridle & King, 2007),

$$\begin{aligned} P_{\mathrm{GI}}(k,z) &= F(z)P_{\delta}(k,z)\,, \\ P_{\mathrm{II}}(k,z) &= F^2(z)P_{\delta}(k,z)\,, \end{aligned} \tag{2.68}$$

in the same way as the lensing-induced shear correlations (equation 1.56). Here, $F(z)$ is given by (Joachimi et al., 2011),

$$F(z) = -A_{\mathrm{IA}}\, C_1\, \rho_{\mathrm{cr}}\, \frac{\Omega_{\mathrm{m}}}{g(z)} \left(\frac{1+z}{1+z_0}\right)^{\eta} \left(\frac{L}{L_0}\right)^{\beta} \tag{2.69}$$

where A_{IA} is the dimensionless IA amplitude, C_1 is a normalisation constant, ρ_{cr} is the $z = 0$ critical density and $g(z)$ is the linear growth function (see, e.g., equation 1.28) normalised to unity at the present-time. η and β are optionally-free parameters introduced to model redshift and luminosity dependence of the IAs about arbitrary pivot points, z_0 and L_0. A_{IA}, η, and β can then be folded into the likelihood analysis as 'nuisance parameters' to be constrained by the data, although it is not uncommon to see one or more of the power terms fixed to zero, depending on whether authors deem their data sensitive enough to detect variations in IAs with redshift/luminosity.

Baryonic feedback

The distribution of matter on cosmological scales is, in reality, influenced by baryonic physics, such as energetic feedback from stars and active galactic nuclei (AGN). These processes 'heat' dark matter haloes, causing them to have lower central densities than those predicted by gravity-only CDM simulations. The impact this can have on the small scales of the matter power spectrum is considerable (Semboloni et al., 2011b). Consider, for example, Figure 2.10, which shows the ratio of matter power spectra including and excluding the impact of AGN feedback at multiple redshifts, as measured by Chisari et al. (2018) using the Horizon cosmological hydrodynamical simulation. We see that the feedback suppresses power, by as much as 50% at $k \simeq 10\,h/\mathrm{Mpc}$ and $z = 0$. The magnitude of this suppression diminishes slightly between $z = 1$ and $z = 0$, owing to the feedback being insufficient to expel gas from the more massive haloes at later

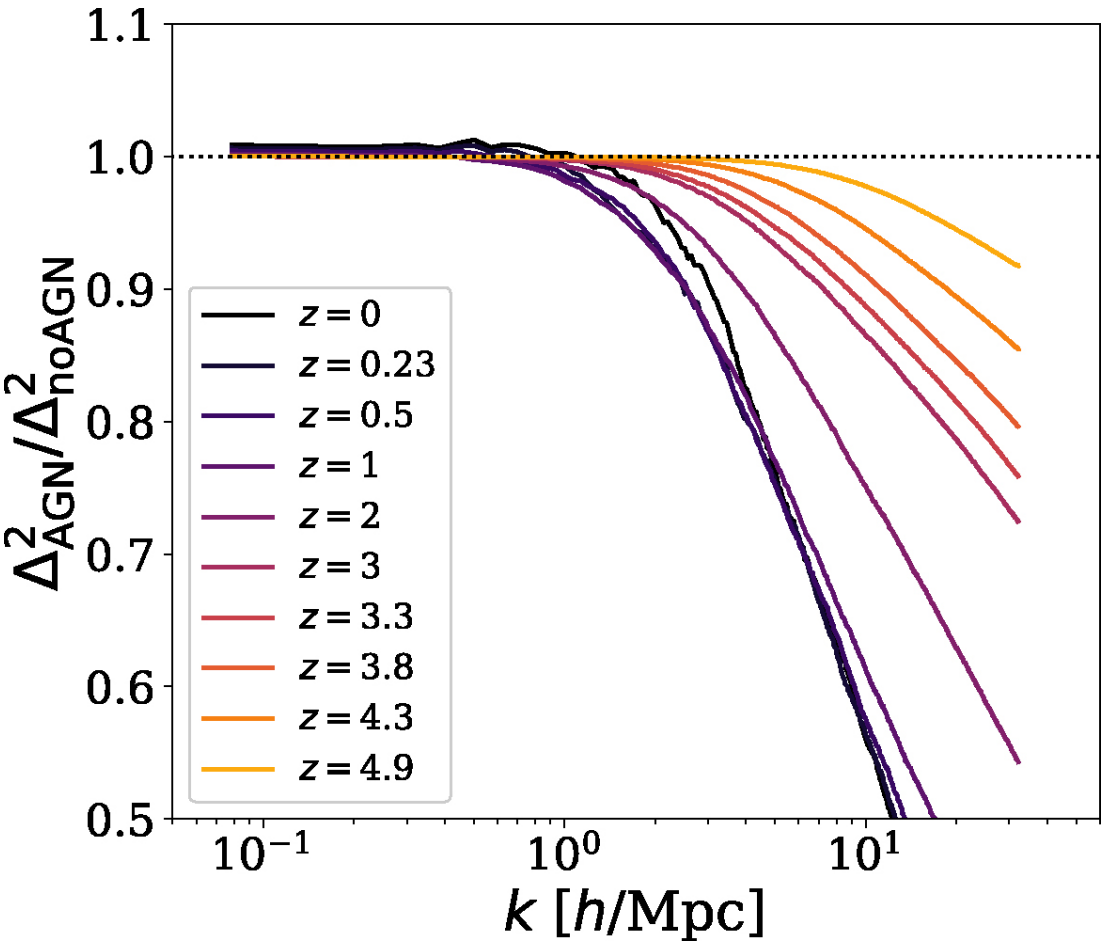

Figure 2.10 *From Chisari et al. (2018). The impact of AGN feedback, on the matter power spectrum at multiple redshifts, as measured in the Horizon hydrodynamical simulation.*

times.

Indeed, thorough investigation on this topic necessitates extremely time-intensive simulations of this ilk, such as the Horizon (Dubois et al., 2014; Dubois et al., 2016), Illustris (Vogelsberger et al., 2014), EAGLE (Schaye et al., 2015), MassiveBlack-II (Khandai et al., 2015) or BAHAMAS (McCarthy et al., 2017) simulations. These resources owe their computational expense to the requirements for realistically complicated galaxy formation models, high-resolution to resolve individual galaxies, and large volumes to infer the matter distribution on cosmological scales.

Attempts to reduce the price of prescriptions for the effects of feedback on cosmological statistics, include introducing extra parameters describing baryonic physics to the halo model (see equation 1.43). These parameters facilitate changes

to the halo concentrations (Semboloni et al., 2011b; Zentner et al., 2013) with further revisions also including a mass-dependent halo 'bloating' parameter (Mead et al., 2015). The latter work was found to fit the results of hydrodynamical simulations with accuracies of $\simeq 5\%$ for $k \leq 10\,h/\mathrm{Mpc}$ and $z \leq 2$, for a variety of ΛCDM and wCDM cosmologies. The Mead et al. (2015) methodology has thus been adopted by the KiDS survey in including the impact of baryonic feedback on the shear correlation functions (Hildebrandt et al., 2017, 2018), with the parameter for the amplitude of the halo-mass concentration being marginalised over at the level of cosmological parameter inference. An alternative strategy, adopted by DES (Troxel et al., 2017) and HSC (Hamana et al., 2019), is to simply exclude, from the likelihood analysis, the small scales in $\xi_{\pm}$ which are intolerably afflicted with the symptoms of baryonic physics.

2.7 Summary

This Chapter contained an in-depth examination of weak gravitational lensing as a cosmological probe, in theory and in practice. This began with discussion of the mathematical framework describing the lensing induced by a discrete mass confined to a thin plane (Section 2.1), how this both magnifies and shears source galaxies (Section 2.2), and how the shapes of said galaxies might be measured to estimate the magnitude of the lensing signal (Section 2.3). In Section 2.4, I presented a more general description in which the mass distribution is continuous, as is the case in reality with lensing via large-scale structure, i.e, cosmic shear. Typical estimators used to extract the cosmological information from observations of this phenomenon, were presented in Section 2.5. Finally, Section 2.6 covered the range of systematic biases which serve to confound cosmological inference via this method, in practice. Therein, I presented my own work in quantifying some of these effects in new KiDS data, and commented on various ways in which the remaining biases have been mitigated in recent cosmic shear analyses.

The previous and current Chapters respectively, discussed the status quo within the wider field of cosmology, and the more specific domain of weak lensing. Having laid this necessary groundwork, in Chapter 3-5, I proceed to present my own contributions to research in this area.

3

KiDS-450: Enhanced Cosmic Shear with Clipping

In this Chapter, I present the first "clipped" cosmic shear measurement using data from the Kilo-Degree Survey (KiDS-450). "Clipping" transformations suppress the signal from the highest density, non-linear regions of cosmological fields. I demonstrate that these transformations improve constraints on $S_8 = \sigma_8(\Omega_\mathrm{m}/0.3)^{0.5}$ when used in combination with conventional two-point statistics. For the KiDS-450 data, I find that the combined measurements improve the constraints on S_8 by 17%, compared to shear correlation functions alone. I determine the expectation value of the clipped shear correlation function using a suite of numerical simulations, and develop methodology to mitigate the impact of masking and shot noise. Future improvements in numerical simulations and mass reconstruction methodology will permit the precise calibration of clipped cosmic shear statistics such that clipping can become a standard tool in weak lensing analyses. This work has been peer-reviewed and published in Giblin et al. (2018).

3.1 Introduction

The use of two-point statistics in extracting information from cosmological fields has been eminently successful to date. Observations of the CMB temperature and polarisation power spectra (Planck Collaboration et al., 2018), weak lensing shear-shear correlation functions (Hildebrandt et al., 2017; Troxel et al., 2017) and shear-shear/convergence power spectra (Köhlinger et al., 2017; van Uitert et al., 2018), for example, have placed meaningful constraints on the cosmological model, helping forge our current understanding of the Universe. Some degree of tension has emerged, however, between state-of-the-art results from weak lensing and CMB cosmological probes, as discussed in Section 1.5. Constraints from the Kilo Degree Survey (KiDS; Hildebrandt et al., 2017) and the Canada France Hawaii Telescope Lensing Survey (CFHTLenS; Heymans et al., 2013), whilst consistent with each other are in some tension with those of the Planck Collaboration et al. (2018). The Year 1 cosmology results from the Dark Energy Survey (Troxel et al., 2017; DES Collaboration et al., 2017) "bridge the gap" between the aforementioned studies, being broadly in agreement with all, as is also the case with the Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP9; Hinshaw et al., 2013). On the other hand, the cosmic shear measurements from the Deep Lens Survey (DLS; Yoon et al., 2018) are fully consistent with Planck and are in some tension with KiDS and CFHTLenS. The range of results on this subject highlights the necessity for more precise and accurate cosmological parameter constraints, thereby affirming whether or not the existing tension is a signature of an exotic form of dark energy or new physics within our Universe (see for example Verde et al., 2013; Joudaki et al., 2017; DES Collaboration et al., 2018). It is with regards to this necessity that I review our employment of two-point statistics for cosmology.

When considering alternatives to two-point statistics, the computational- and time-intensiveness of collecting and reducing observations in the era of precision cosmology must also be considered. Two-point statistics alone fail to exploit the full wealth of information within these expensive data sets, on account of the presence of regions of non-linear gravitational collapse. Consequently, it is crucial that we employ all possible statistical tools to capitalise on the available data sets.

Indeed, the sub-optimality of two-point statistics has driven research involving

non-Gaussian statistics. Counting the abundance of high-density regions, known as "peak statistics" (Jain & Van Waerbeke, 2000, see Section 2.5.4), as well as extending the cosmological analysis to third and higher order statistics (Takada & Jain, 2002; Bernardeau, 2005; Kilbinger & Schneider, 2005; Semboloni et al., 2011a; Fu et al., 2014) have been shown to yield improved constraints on cosmology. In addition, one can perform transformations to enhance the linearity of the cosmological field in question, improving the capacity of two-point statistics to contrain cosmology. For example, Neyrinck et al. (2009) and Seo et al. (2011) found various logarithmic transformations are sufficient for this purpose.

In particular, "clipping" transformations have been shown to be beneficial to a number of analyses. Clipping truncates the peaks above a given threshold within a density field, thereby suppressing the contributions of high-density regions to the power spectrum. This methodology was successfully applied to galaxy number counts within numerical simulations, and found to increase the range of Fourier modes in which the power spectrum and bispectrum can be related with tree-level perturbation theory, leading to precise determination of the galaxy bias and the amplitude of matter perturbations σ_8 (Simpson et al., 2011, 2013). Furthermore, Simpson et al. (2016a) clip galaxy number counts from the Galaxy and Mass Assembly Survey (GAMA), to reduce the impact of non-linear processes and galaxy bias on the analysis, allowing for reliable constraints on the rate of growth of structure in the Universe. Wilson (2016) employed clipping in estimating the growth rate of structure from the VIMOS Public Extragalactic Redshift Survey as part of a redshift-space distortion analysis. Lombriser et al. (2015) also demonstrate that clipping density fields allows for modified gravity models to more easily be distinguished from concordance cosmology.

Clipping can also be combined with standard cosmological statistics, as demonstrated by Simpson et al. (2016b, henceforth 'S16') in a weak lensing analysis. They truncate the peaks in simulated fields of the projected surface density, i.e. the convergence (see equation 2.10), and measure the effect on the convergence power spectrum (equation 2.46). The objective of clipping in this context is to reduce the correlations between the Fourier modes in the convergence power spectrum in order to unlock previously inaccessible cosmological information. An alternative interpretation of the information gain in clipping, is that it is analogous to that which is found in peak statistics analyses, since both methods selectively target high-density regions. Via a Fisher matrix analysis, S16 predict the constraints on the amplitude of matter perturbations, σ_8, and the matter

density parameter, $\Omega_{\rm m}$, one would obtain from the "clipped" and the conventional "unclipped" convergence power spectra. They find that clipping engenders a small clockwise rotation of the clipped contours relative to the unclipped, breaking the degeneracy in the $\Omega_{\rm m}$-σ_8 parameter space (see Figure 2 of S16). The consequence of this is that when the contours from the two power spectra are combined (taking into account the cross-covariance of the clipped and unclipped statistics, so as to avoid double-counting) the constraints on $\Omega_{\rm m}$ and σ_8 are increased overall by more than a factor of three. Moreover, clipping is found to be more constraining than the alternative logarithmic transforms proposed by Neyrinck et al. (2009).

A crucial aspect of clipping convergence fields containing regions of non-linear gravitational collapse, is the fact that there currently exists no analytical prescription for the clipped statistics one will subsequently measure. This means that numerical simulations are necessary for establishing their cosmological dependence. This is not a disadvantage specific to clipping, given that peak statistics (Jain & Van Waerbeke, 2000; Kacprzak et al., 2016; Martinet et al., 2018) and higher order statistics (Takada & Jain, 2002; Semboloni et al., 2011a), similarly necessitate simulations for calibration. Furthermore, simulations are also required for investigating the behaviour of standard cosmological statistics on non-linear scales (Smith et al., 2003; Takahashi et al., 2012).

In this work I apply clipping to weak lensing convergence fields measured from the first 450 square degrees of r-band data from the Kilo-Degree Survey (hereafter 'KiDS-450'). In contrast to S16, rather than determine the effect of clipping on the convergence power spectrum, I investigate for the first time the properties of the clipped two-point shear correlation functions, expressed in equation 1.56. This is to facilitate a direct comparison of the clipped statistics to the conventional shear correlation functions used in constraining cosmology in the Hildebrandt et al. (2017) analysis. By exploring the cosmological dependence of clipping with the Dietrich & Hartlap (2010, hereafter 'DH10') simulations, and by measuring the covariance of these new statistics using the Scinet Light Cone Simulations (SLICS) from Harnois-Déraps et al. (2018), I constrain the cosmology of the KiDS-450 data. I also characterise how clipping is affected by masking and shape noise, and demonstrate how these can be accounted for. The format of this Chapter is as follows; in Section 3.2 I discuss the KiDS-450 data and the N-body simulations at my disposal, in Section 3.3 I explain my methodology for measuring the clipped shear correlation functions and discuss calibration corrections, in

Section 3.4 I present my results, and finally I conclude in Section 3.5.

3.2 Data and Simulations

The Kilo Degree Survey (KiDS) is an ESO public survey which will span 1350 square degrees upon completion. KiDS observes with the VLT Survey Telescope (VST) in the *ugri* bands, with science goals pertaining to cosmology and galaxy evolution. In this Chapter I focus on the KiDS-450 data release, containing the first 450 square degrees of four-band coverage (Hildebrandt et al., 2017, hereafter 'H17'). The KiDS-450 data is divided between five patches, G9, G12, G15, G23 and GS (de Jong et al., 2017) and consists of *lens*fit (Miller et al., 2013) shear estimates for ∼15 million galaxies. The effective number of galaxies per square arcminute in the data is 8.53 and the galaxy ellipticities have a dispersion of $\sigma_e = 0.29$ per component. The photometric redshifts of the background galaxies are estimated from the four-band photometry using the Bayesian photometric redshift BPZ code from Benítez (2000), as described in Hildebrandt et al. (2012). In addition, three different techniques for calibrating the effective redshift distribution $n(z)$ are investigated in H17 and found to produce consistent cosmic shear results. In constraining the KiDS-450 cosmology in this analysis, I adopt the method favoured in H17 – the weighted direct calibration ("DIR"). This follows the methodology of Lima et al. (2008), where a subsample of galaxies with spectroscopic redshifts are reweighted such that the photometric observables (e.g. colours, magnitudes) of the reweighted sample match the larger sample of galaxies with photometric redshifts only. The reweighted spectroscopic redshift distribution is then taken to be representative of the whole sample. I refer the reader to Kuijken et al. (2015) for more technical discussion of the survey.

The shapes of galaxies in KiDS-450, characterised by two ellipticity components, are measured with the *lens*fit algorithm (Miller et al., 2013) from the r-band data, as described in Fenech Conti et al. (2017). *Lens*fit models the point spread function (PSF) at the pixel level for individual exposures, and then measures the ellipticity components by fitting a PSF-convolved disc and bulge model to each galaxy via a likelihood-based method. Weights for the shape measurement are then derived from the likelihood surface. I calibrate the shape measurements with the additive and multiplicative corrections detailed in Appendix D of H17. The former correction is determined empirically by averaging the observed ellipticities

in the data, whereas the latter is quantified with image simulations resembling the KiDS-450 r-band.

The absence of an analytical prescription for clipped statistics means that in order to use clipping to constrain cosmological parameters, I require a suite of numerical simulations for various cosmologies to determine how clipping responds to changes in said parameters. In addition, this task requires that the covariance of my clipped statistic is accurately measured, which necessitates a large number of independent realisations for a given cosmology. These requirements are at odds with one another; given the computational expense, simulators typically must choose between producing simulations for a large range of cosmological configurations, or producing many realisations for a single cosmology. Therefore I am compelled to use two different simulation suites to satisfy these two criteria – DH10 and SLICS.

The DH10 suite (Dietrich & Hartlap, 2010) consists of numerical N-body simulations ran with the TREEPM code GADGET-2 (Springel, 2005) and initial conditions generated with the Eisenstein & Hu (1998) transfer function. There are 192 DH10 simulations spanning 158 different flat ΛCDM cosmologies. Each simulation has 256^3 dark matter particles in a box with sides of length 140 h^{-1}Mpc, evolved from $z = 50$ to $z = 0$. The lightcone area per simulation is 6×6 square-degrees, and the particle mass varies from $\mathrm{m_p} = 9.3 \times 10^9\,\mathrm{M_\odot}$ for $\Omega_\mathrm{m} = 0.07$, to $\mathrm{m_p} = 8.2 \times 10^{10}\,\mathrm{M_\odot}$ for $\Omega_\mathrm{m} = 0.62$. 35 of the simulations have the fiducial cosmological parameters given by $\boldsymbol{\pi_0} = \{\Omega_\mathrm{m} = 0.27, \Omega_\Lambda = 0.73, \Omega_\mathrm{b} = 0.04, \sigma_8 = 0.78, n_\mathrm{s} = 1.0, h = 0.7, w_0 = -1\}$. The remaining 157 cosmologies, each of which comprise a single N-body simulation, differ only in Ω_m and σ_8, the range of which is displayed in Figure 3.1. Hence, in this work I only demonstrate the power of clipping in constraining $S_8 = \sigma_8(\Omega_\mathrm{m}/0.3)^{0.5}$, which probes the Ω_m-σ_8 parameter space in the direction approximately perpendicular to the degeneracy between these parameters, for a flat ΛCDM Universe. These constraints are obtained with the other cosmological parameters fixed to their fiducial values.

Catalogues of the noise-free shear components for galaxies are produced by ray-tracing through each DH10 N-body simulation. This consists of propagating light rays through the matter distribution constructed by the N-body simulation, from galaxies with a given distribution in redshift. The matter distribution exists in the form of mass snapshots at various redshifts; the deflection of light rays by these mass planes determines the shear of the mock galaxies. Five pseudo-independent

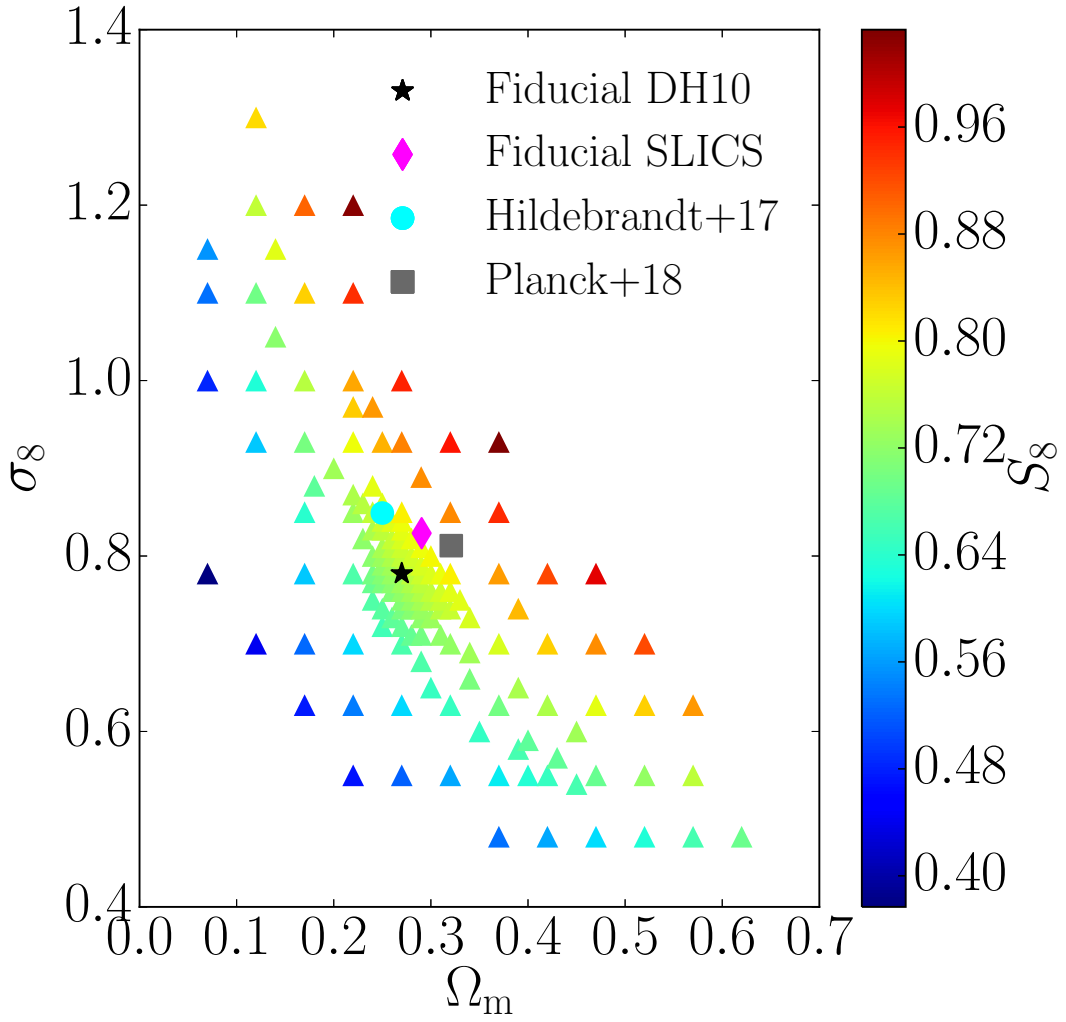

Figure 3.1 *The 158 cosmologies of the DH10 simulations in the Ω_m-σ_8 plane (triangles), colour coded by $S_8 = \sigma_8(\Omega_m/0.3)^{0.5}$. The fiducial cosmologies of DH10 and SLICS are shown by the black star and magenta diamond, respectively. The cyan circle and grey square designate the best-fit (Ω_m, σ_8) determined from the KiDS-450 data in H17, and from the TT+lowE analysis of the Planck data in Planck Collaboration et al. (2018), respectively.*

shear catalogues are obtained for a given simulation by ray-tracing through five different random angles. Thus, in this work I am using 35×5 shear catalogues for the fiducial cosmological parameters, and 1×5 shear catalogues for the remaining 157 cosmologies.

In order to measure the covariance of clipped statistics, I employ the public[1] Scinet Light Cone Simulations (SLICS) of Harnois-Déraps et al. (2018). The SLICS suite evolved 1536^3 particles of mass $m_{\rm p} = 4.17 \times 10^9\,{\rm M}_\odot$, from $z = 120$ to $z = 0$ in a box with sides of length 505 h^{-1}Mpc. They were created using the CUBEP^{3}M N-body code (Harnois-Déraps et al., 2013), with initial conditions selected from the Zel'dovich displacement of particles based on a transfer function from CAMB (Lewis et al., 2000). The SLICS consist of just three cosmologies and are therefore unable to determine the cosmological dependence of clipping. However, on account of there being 932 realisations of 100 deg^2 light cones for the fiducial cosmology, $\{\Omega_{\rm m} = 0.2905,\ \Omega_\Lambda = 0.7095,\ \Omega_b = 0.0473,\ h = 0.6898,\ \sigma_8 = 0.826, n_{\rm s} = 0.969$ and $w_0 = -1\}$, SLICS are very well suited to covariance estimation. In this work I use only the SLICS with the fiducial cosmology, and assume that the covariance measured from these realisations is robust to changes in cosmology. This is a commonly made approximation, as neglecting the cosmological dependence of the covariance has been shown to have little effect on the best-fit value of S_8 if the fiducial cosmology is sufficiently close to that of the best-fit (Eifler et al., 2009). In this case, the SLICS cosmological parameters are close to the best-fit from the H17 analysis of the KiDS-450 data, the fiducial cosmology of DH10, and the best-fit from Planck Collaboration et al. (2018), as is shown in Figure 3.1. Thus my approximation of a cosmology-independent covariance matrix is reasonable given the data I am working with. A comparison of the DH10 and SLICS specifications is presented in Table 3.1. Both suites consist of dark matter particles only.

The fact that galaxies can be intrinsically aligned through gravitational interaction, rather than have their alignments induced by weak gravitational lensing, poses a systematic bias to cosmological inference (Bridle & King, 2007), as discussed in Section 2.6.2. In order to reduce the influence of intrinsic alignments in this work, I follow Benjamin et al. (2013) and restrict my analysis to the 0.5–0.9 photometric redshift range in the KiDS-450 data. Within this tomographic interval, the density of source galaxies is 3.32 gal/arcmin^2 and the galaxy ellipticities have a dispersion of $\sigma_e = 0.28$ per component. I downsample the

[1]SLICS N-body simulations; `http://slics.roe.ac.uk`

Table 3.1 *A comparison of the specifications of the SLICS and DH10 suites used in this Chapter. These simulations are used for estimation of the covariance, and the dependence on cosmological parameters, of the clipped shear correlation functions, $\xi_{\pm}^{\rm clip}$, respectively.*

	SLICS	DH10
Science case	Covariance Matrices	Cosmological Dependence
Cosmologies	1	158
Realisations per cosmology	932	35(Fiducial)+1(Other)
Lightcone area [deg^2]	100	36
Box size [Mpc/h]^3	505^3	140^3
Particles	1536^3	256^3
Particle Mass [$M_\odot$]	4.17×10^9	9.3×10^9–8.2×10^{10}

SLICS and DH10 mock catalogues so as to have the same source density and redshift distribution of the data, which I take to be the KiDS-450 DIR-calibrated redshift distribution (H17), which has mean and standard deviation of 0.76 and 0.29, respectively, in my chosen redshift bin. I also introduce Gaussian-distributed galaxy ellipticities to the mocks, with standard deviation, σ_e, equal to that of the KiDS-450 data. I do not truncate the Gaussian distribution to ellipticities between -1 and 1, since less than 0.05% of mock galaxies are allocated ellipticities outside of this range, and their contributions to the correlation functions are negligible. I also verified that using ellipticities directly sampled from the distribution in the data, instead of from a Gaussian, does not affect my results. Matching the shape noise (which in this work I use to refer to all factors contributing to the measured galaxy shape, bar the shear itself) and source densities, means that the noise in the covariance matrices and the clipped predictions from the mocks reflect that of KiDS-450. The effect of baryonic physics on the shear correlation functions is another source of bias in weak lensing analyses (Semboloni et al., 2011b), and could in principle affect clipped statistics differently than the unclipped. For this first proof-of-concept analysis however, I do not contend with baryonic effects in this work.

3.3 Methodology

In this Section, I describe the pipeline in which I apply clipping transformations to the mocks and KiDS-450 data, and subsequently measure the "clipped" two-point shear correlation functions $\xi_{\pm}^{\rm clip}$. Measuring these statistics allows for a comparison to the conventional "unclipped" shear correlation functions, which

are directly calculated from the observed galaxy ellipticities in the data. I begin with a very brief summary of the key steps in my method for easy referral. I discuss these steps in greater detail in the Sections that follow. In my description of the methodology, I also reiterate many of the relevant cosmic shear equations from Chapter 2 without proof, and refer the reader to there for more in-depth discussion.

- My pipeline takes as input catalogues of the ellipticities and positions of galaxies. I project these onto a Cartesian grid of pixels with a resolution of 5 arcseconds, smooth these maps with a Gaussian filter and reconstruct the projected surface mass density, i.e. the convergence, κ, following Kaiser & Squires (1993).

- I subject these convergence maps to clipping; anywhere the convergence exceeds a certain threshold value, I set the convergence equal to that threshold.

- The resulting "clipped" convergence map is subtracted from the "unclipped" thereby generating a map containing the projected surface density exceeding the threshold, and zeroes elsewhere. On this "residual" convergence map, I invert the mass reconstruction process and recover the shear corresponding to these projected peaks.

- This "residual" shear is subtracted from the original shear values yielding the "clipped" shear. From the clipped shear, I calculate the clipped shear correlation functions, $\xi_{\pm}^{\rm clip}$, using TreeCorr (Jarvis, 2015). To measure the unclipped shear correlation functions, $\xi_{\pm}^{\rm unclip}$, I feed the catalogues of the observed ellipticities to TreeCorr directly.

- I repeat this process for successive SLICS realisations to measure the covariance of the $\xi_{\pm}^{\rm clip}$ and $\xi_{\pm}^{\rm unclip}$ statistics, and for successive DH10 realisations to determine the cosmological dependence of the $\xi_{\pm}^{\rm clip}$.

3.3.1 Mass reconstruction

In order to clip the densest non-linear regions from my analysis, I first produce maps of the projected surface mass density, or convergence, κ, using the methodology of Kaiser & Squires (1993, 'KS93' hereafter). In this analysis, the

process of "mass reconstruction" begins with the observed ellipticities, which can be written in the complex form $\epsilon^{\rm obs} = \epsilon_1^{\rm obs} + i\epsilon_2^{\rm obs}$ (Seitz & Schneider, 1996). The observed ellipticities have contributions from the reduced shear g, the intrinsic ellipticity $\epsilon^{\rm int}$ and the shape measurement noise η via

$$\epsilon^{\rm obs} = \frac{g + \epsilon^{\rm int}}{1 + g^* \epsilon^{\rm int}} + \eta \,, \tag{3.1}$$

where g^* is the complex conjugate of g. The reduced shear is related to the shear γ and the convergence κ by $g = \gamma/(1-\kappa)$. In a weak lensing analysis, we assume that the magnitudes of both the shear and the convergence are much smaller than unity, such that the average of the observed ellipticities $\langle \epsilon^{\rm obs} \rangle \simeq g \simeq \gamma$. In this case, it is possible to reconstruct the convergence from the observed ellipticities via the KS93 inversion method. To begin with, consider the gravitational deflection potential, $\Psi(\boldsymbol{\theta})$. This is related to the convergence, κ, for a particular source redshift and angular coordinate on the sky $\boldsymbol{\theta} = (\theta_1, \theta_2)$, via Poisson's equation (first presented in equation 2.15),

$$\nabla^2 \Psi(\boldsymbol{\theta}) = 2\kappa(\boldsymbol{\theta}) \,, \tag{3.2}$$

where $\Psi(\boldsymbol{\theta})$ is given by the line of sight integral over the 3D matter gravitational potential Φ,

$$\Psi(\boldsymbol{\theta}) = \int_0^{\chi_s} d\chi' \frac{f_{\rm K}(\chi - \chi')}{f_{\rm K}(\chi) f_{\rm K}(\chi')} \Phi \left[f_{\rm K}(\chi')\boldsymbol{\theta}, \chi' \right] . \tag{3.3}$$

Here χ is the comoving radial distance, χ_s is the comoving radial distance to the source, and $f_{\rm K}(\chi)$ is the comoving angular diameter distance. The potential $\Psi(\boldsymbol{\theta})$ is related to the shear components $\gamma_i(\boldsymbol{\theta})$ via

$$\gamma_i(\boldsymbol{\theta}) = D_i \Psi(\boldsymbol{\theta}) \,, \tag{3.4}$$

where

$$\begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \partial^2/\partial\theta_1\partial\theta_1 - \partial^2/\partial\theta_2\partial\theta_2 \\ 2\partial^2/\partial\theta_1\partial\theta_2 \end{pmatrix} , \tag{3.5}$$

and ∂ denotes partial derivatives. Combining equations 3.2 and 3.4 and taking the Fourier transform yields

$$\tilde{\gamma}_i(\boldsymbol{\ell}) = F_i(\boldsymbol{\ell})\tilde{\kappa}(\boldsymbol{\ell})\,, \tag{3.6}$$

where

$$\begin{pmatrix} F_1 \\ F_2 \end{pmatrix} \equiv \begin{pmatrix} ({\ell_1}^2 - {\ell_2}^2)/\ell^2 \\ 2\ell_1\ell_2/\ell^2 \end{pmatrix}\,, \tag{3.7}$$

and $\boldsymbol{\ell} = (\ell_1, \ell_2)$ is the 2D Fourier conjugate of $\boldsymbol{\theta}$.

From equation 3.6 we see that, in principle, either $\tilde{\gamma}_1(\boldsymbol{\ell})/F_1(\boldsymbol{\ell})$ or $\tilde{\gamma}_2(\boldsymbol{\ell})/F_2(\boldsymbol{\ell})$ would suffice to give an estimate of $\tilde{\kappa}(\boldsymbol{\ell})$, which can then be inverse-Fourier transformed to recover $\kappa(\boldsymbol{\theta})$. Both $F_1(\boldsymbol{\ell})$ and $F_2(\boldsymbol{\ell})$ vanish for particular directions however, so instead I sum over the $\tilde{\gamma}_i(\boldsymbol{\ell})$ components weighted by $F_i(\boldsymbol{\ell})$ to obtain the convergence,

$$\sum_{i=1}^{2} F_i(\boldsymbol{\ell})\tilde{\gamma}_i(\boldsymbol{\ell}) = \sum_{i=1}^{2} |F_i(\boldsymbol{\ell})|^2\tilde{\kappa}(\boldsymbol{\ell}) = \tilde{\kappa}(\boldsymbol{\ell})\,, \tag{3.8}$$

where I have employed the fact that $\sum_{i=1}^{2} |F_i(\boldsymbol{\ell})|^2$ is equal to unity (Kaiser, 1992). An inverse-Fourier transform is performed to reconstruct the $\kappa(\boldsymbol{\theta})$ map, the real part of which contains the E-modes, whereas the imaginary part contains the B-modes[2] (Schneider et al., 2002b).

The KS93 mass reconstruction can be summarised in the following:

- The shear is projected onto a Cartesian grid and smoothed with a Gaussian filter with width σ_s to reduce the impact of mask features (which remove artefacts) on the reconstruction.
- A border of zero values is added to the smoothed shear map, increasing the dimensions by 1 deg on each side, before Fourier transforming the field.

[2]Hildebrandt et al. (2017) and van Uitert et al. (2018) report significant B-modes within the KiDS-450 data but as these are at such a low-level in comparison to the E-mode signal I do not consider them in this analysis.

The border serves to reduce edge effects in the transform (Van Waerbeke et al., 2013).

- $\tilde{\kappa}(\boldsymbol{\ell})$ is computed via equation 3.8.

- An inverse-Fourier transform is performed to reconstruct the $\kappa(\boldsymbol{\theta})$ map.

The steps I take in mass reconstruction follow this recipe. However, in this analysis I am working with real data and simulations tailored to the data in terms of the redshift distribution, source density and galaxy shape noise. My observed ellipticities (see equation 3.1), smoothed with the Gaussian filter, are treated as an unbiased estimator for the shear and take the place of γ in the above equations. Furthermore, the KiDS-450 data has masked regions leading to gaps in the observed patches. The Gaussian smoothing accounts for the number of masked pixels within the smoothing window, to minimise the bias in the resultant smoothed ellipticity (see Van Waerbeke et al., 2013, for more details). The effect of masking on the clipped shear correlation functions $\xi_{\pm}^{\rm clip}$ is discussed in Section 3.3.4. I refer to the width of the Gaussian smoothing filter as the *smoothing scale*, $\sigma_{\rm s}$, hereafter.

The KS93 methodology has been shown to be accurate for relatively small fields ($\lesssim 100$ deg^2) which may be approximated as flat (Van Waerbeke et al., 2013). Other mass reconstruction methods do exist; for example Seitz & Schneider (1996) generalise the KS93 technique into the lensing regime where the $\kappa \ll 1$ approximation no longer holds, whereas Chang et al. (2017) conduct curved-sky mass reconstruction with a spherical harmonic formalism. The KS93 methodology is sufficiently accurate for my purposes however, since the KiDS-450 patches, DH10 mocks and SLICS are well described by the flat-sky approximation, and the convergence is sufficiently small (see Section 3.3.3). Future clipping analyses, especially those involving data sets with larger sky coverage, will require these improved methodologies. Convergence maps for the KiDS-450 patches created following KS93 are presented in Appendix 3.C.

3.3.2 Clipping methodology

After the convergence field is generated it is clipped if above a given threshold $\kappa^{\rm c}$ according to

$$\kappa_s^{\rm clip}(\boldsymbol{\theta}) = \begin{cases} \kappa^{\rm c}, & \text{if } \kappa_s(\boldsymbol{\theta}) \geq \kappa^{\rm c} \\ \kappa_s(\boldsymbol{\theta}), & \text{otherwise} \end{cases}, \tag{3.9}$$

where the 's' subscript is used to denote fields either directly smoothed with the Gaussian filter, or those derived from fields which have been directly smoothed. I calculate the "residual" convergence $\Delta\kappa_s$, given by

$$\Delta\kappa_s(\boldsymbol{\theta}) = \kappa_s(\boldsymbol{\theta}) - \kappa_s^{\rm clip}(\boldsymbol{\theta}). \tag{3.10}$$

The $\Delta\kappa_s$ map features the projected surface density exceeding the threshold $\kappa^{\rm c}$, and zeroes elsewhere. I subject this map to an inversion of the mass reconstruction process following equation 3.6. This generates the "residual" ellipticity maps $\Delta\epsilon_s$, which exhibit the strongest signal around the positions of the peaks, and weaker signal elsewhere. The residual ellipticities are defined on a grid; in order to obtain $\Delta\epsilon_s$ at the locations of the galaxies in the original, "unclipped" ellipticity catalogue, $\boldsymbol{\theta}_{\rm g}$, I perform 2D linear interpolation from the $\Delta\epsilon_s$ maps. The clipped ellipticity $\epsilon_s^{\rm clip}$ is the difference between the observed (unclipped) ellipticity $\epsilon^{\rm obs}$ and the residual ellipticity $\Delta\epsilon_s$,

$$\epsilon_s^{\rm clip}(\boldsymbol{\theta}_{\rm g}) = \epsilon^{\rm obs}(\boldsymbol{\theta}_{\rm g}) - \Delta\epsilon_s(\boldsymbol{\theta}_{\rm g}). \tag{3.11}$$

It is inadvisable to recover the clipped ellipticity, $\epsilon_s^{\rm clip}$, by conducting inverse mass reconstruction *directly* on the clipped convergence map, $\kappa_s^{\rm clip}$. This is because $\kappa_s^{\rm clip}$ has been affected by smoothing in all regions where the convergence is below the clipping threshold $\kappa^{\rm c}$ (those regions with convergence above $\kappa^{\rm c}$ are set to the constant threshold itself), and smoothing incurs a loss of signal. This corresponds to $\sim$90% of the area of $\kappa_s^{\rm clip}$ being affected by smoothing, for the $\kappa^{\rm c}$ and smoothing scale, $\sigma_{\rm s}$, values I identify in Section 3.3.3. In contrast, if I invert the mass reconstruction on the $\Delta\kappa_s$, only $\sim$10% of the area of which is smoothed, and subtract the $\Delta\epsilon_s$ from the unsmoothed observed ellipticities, $\epsilon^{\rm obs}$, I minimise the

impact of smoothing on my overall signal.

After computing the clipped ellipticity components via equation 3.11, using TreeCorr (Jarvis, 2015) I calculate estimators for the clipped and unclipped angular shear correlation functions in nine logarithmically spaced angular bins, θ, with bin centres from 0.78 to 219 arcmin. Following equation 1.55, these estimators, within a single tomographic bin, are defined as,

$$\widehat{\xi}_{\pm}(\theta) = \frac{\sum_{\mathrm{ab}} w_{\mathrm{a}} w_{\mathrm{b}} \left[\epsilon_{\mathrm{t}}(\boldsymbol{\theta}_{\mathrm{g,a}})\epsilon_{\mathrm{t}}(\boldsymbol{\theta}_{\mathrm{g,b}}) \pm \epsilon_{\times}(\boldsymbol{\theta}_{\mathrm{g,a}})\epsilon_{\times}(\boldsymbol{\theta}_{\mathrm{g,b}})\right]}{\sum_{\mathrm{ab}} w_{\mathrm{a}} w_{\mathrm{b}}}\,, \tag{3.12}$$

where the summation is over pairs of galaxies a and b positioned at angular coordinates $\boldsymbol{\theta}_{\mathrm{g,a/b}}$, within an interval $\Delta\theta$ about the angular separation θ (Bartelmann & Schneider, 2001). The ϵ_{t} and $\epsilon_{\times}$ terms designate the tangential- and cross- components of the clipped ellipticities (in the case of the $\widehat{\xi}_{\pm}^{\mathrm{clip}}$ estimator) or the observed ellipticities (in the case of the unclipped estimator $\widehat{\xi}_{\pm}^{\mathrm{unclip}}$, see equation 2.47) measured relative to the vector $\boldsymbol{\theta}_{\mathrm{g,a}} - \boldsymbol{\theta}_{\mathrm{g,b}}$ connecting the galaxy pairs. w is the weight ascribed to the measurement of the ellipticity components, which comes from the *lens*fit algorithm in the case of KiDS-450 (refer to Section 3.2 for more details) or takes the value of unity in the case of the mocks.

I treat the observed ellipticities, a combination of the shear and shape noise via equation 3.1, in the mocks and data as unbiased estimators for the shear. Accordingly I treat $\widehat{\xi}_{\pm}^{\mathrm{unclip}}$ as an unbiased estimator of the theoretical unclipped shear correlation functions, $\xi_{\pm}^{\mathrm{unclip}}$, defined in equation 3.13. Consequently, in this work I follow H17 and refer to the *estimators* for the unclipped shear correlation functions simply as the unclipped shear correlation functions, and omit the $\widehat{\ }$ notation. There is currently no established theoretical prediction for $\xi_{\pm}^{\mathrm{clip}}$. Thus it is not meaningful to include the$\widehat{\ }$notation nor "estimator" prefix for my measured clipped statistics, and I similarly drop this nomenclature henceforth. However, I encourage the reader not to regard the clipped statistics measured from the mocks as unbiased estimators of the clipped measurement made in the absence of shape noise (as is done with the unclipped statistic). The clipped statistics I measure not only depend on the level of shape noise, but also the clipping threshold and level of smoothing applied in the analysis (see Section 3.3.3).

To reiterate equations 1.56-1.58, the theoretical unclipped shear correlation

functions $\xi_{\pm}^{\rm unclip}$ are related to the convergence power spectrum $P_\kappa(\ell)$ via

$$\xi_{\pm}^{\rm unclip}(\theta) = \frac{1}{2\pi}\int {\rm d}\ell\, \ell\, P_\kappa(\ell)\, J_{0,4}(\ell\theta)\,, \tag{3.13}$$

where the zeroth $J_0(\ell\theta)$ and fourth $J_4(\ell\theta)$ order Bessel functions of the first kind are used for $\xi_{+}^{\rm unclip}$ and $\xi_{-}^{\rm unclip}$ respectively. The convergence power spectrum $P_\kappa(\ell)$ is in turn related to the matter power spectrum $P_\delta(\ell)$ via

$$P_\kappa(\ell) = \int_0^{\chi_{\rm H}} {\rm d}\chi\, \frac{q(\chi)^2}{f_{\rm K}(\chi)^2}\, P_\delta\left(k = \frac{[\ell+1/2]}{f_{\rm K}(\chi)}, \chi\right), \tag{3.14}$$

where $\chi_{\rm H}$ is the comoving radial distance to the horizon and k is the Fourier conjugate of χ. Here I have used the flat-sky first-order extended Limber approximation, which is sufficiently accurate for the KiDS-450 data (see Kilbinger et al., 2017a). The lensing efficiency, $q(\chi)$, is defined as

$$q(\chi) = \frac{3H_0^2\Omega_{\rm m}}{2c^2}\frac{f_{\rm K}(\chi)}{a(\chi)}\int_\chi^{\chi_{\rm H}} {\rm d}\chi'\; n(\chi')\frac{f_{\rm K}(\chi'-\chi)}{f_{\rm K}(\chi')}\,, \tag{3.15}$$

where a is the scale factor, $n(\chi)$ is the probability density of galaxies as a function of χ, H_0 is the Hubble constant and c is the speed of light.

Constraining the cosmology of the KiDS-450 data requires covariance matrices for the clipped and unclipped $\xi_\pm$. I measure the covariance of these statistics across N ∼900 independent SLICS realisations. The $i^{\rm th}$ and $j^{\rm th}$ elements of the covariance matrices are given by

$$C_\pm(\theta_i,\theta_j) = \sum_k^N \frac{(\xi_\pm^k(\theta_i) - \overline{\xi_\pm}(\theta_i))(\xi_\pm^k(\theta_j) - \overline{\xi_\pm}(\theta_j))}{N-1}\,, \tag{3.16}$$

where $\overline{\xi_\pm}(\theta_i)$ refers to either the mean clipped or mean unclipped $\xi_\pm$, across N realisations each numerated by k, within the $i^{\rm th}$ angular separation bin, given by $\sum_k^N \xi_\pm^k(\theta_i)/N$. When computing the auto-covariance of the clipped (or unclipped) statistic, all correlation functions in equation 3.16 correspond to $\xi_\pm^{\rm clip}$ (or $\xi_\pm^{\rm unclip}$). When computing the cross-covariance between the clipped and unclipped, the $\xi_\pm$ correspond to clipped in one bracket, and to unclipped in the other. In order to

constrain the cosmology of KiDS-450, I scale the covariance matrices measured from SLICS by the ratio of the areas of SLICS and KiDS-450 (Schneider et al., 2002c). I note that this is an approximation and does not account for the survey geometry, as is discussed in Troxel et al. (2018). Correlation coefficient matrices, calculated from the SLICS covariance matrices, are present in Appendix 3.A.

3.3.3 Choosing the clipping threshold and smoothing scale

In a clipping analysis, the values of the convergence threshold, $\kappa^{\rm c}$, at which peaks are truncated and the width of the Gaussian with which the ellipticity maps are smoothed, i.e. the smoothing scale $\sigma_{\rm s}$, are free parameters. Thus an important aspect of clipping is to identify values which are appropriate for the data one wishes to analyse. Suitable choices of these parameters depend on the depth and resolution of the data. These parameters are also degenerate with one another; for a given value of $\kappa^{\rm c}$, a lower level of smoothing results in more of the convergence field exceeding the clipping threshold. Similarly, for a fixed $\sigma_{\rm s}$, lesser values of $\kappa^{\rm c}$ correspond to more aggressive clipping. The interplay of these parameters means that the optimal values for constraining cosmology are costly to determine. Consequently, in this work I only determine values which are well suited to the KiDS-450 data. I also investigate the effect of different choices of the smoothing scale and clipping threshold on the clipped correlation functions.

I first establish a clipping threshold which targets the most non-linear regions of the field, without over-clipping the linear field. An intuitive way of doing this is to first fix the smoothing scale and determine where the PDF of the convergence deviates from Gaussian. However, I find that even for relatively large values of the smoothing scale, the KiDS-450 PDF(κ) is too noisy for this test. I therefore use the SLICS, the fiducial $\Omega_{\rm m}$ and σ_8 of which are similar to the best-fit values from the H17 analysis of the KiDS-data (see Figure 3.1). In the upper panel of Figure 3.2 I compare the unclipped PDF(κ) measured from 50 SLICS with a smoothing scale of 6.6 arcmin (orange) to a Gaussian fit (dashed blue). Also shown is the corresponding clipped distribution for an example clipping threshold of 0.010, which is by definition identical to the unclipped at $\kappa < 0.010$ (see equation 3.9). The middle panel shows the fractional difference between the Gaussian fit and the unclipped SLICS PDF(κ), with the deviations between the two at $\kappa = 0.005$, 0.010, 0.015 and 0.020 highlighted by the dotted lines and annotations. Finally, in the lower panel I show the unclipped PDF(κ) for the five

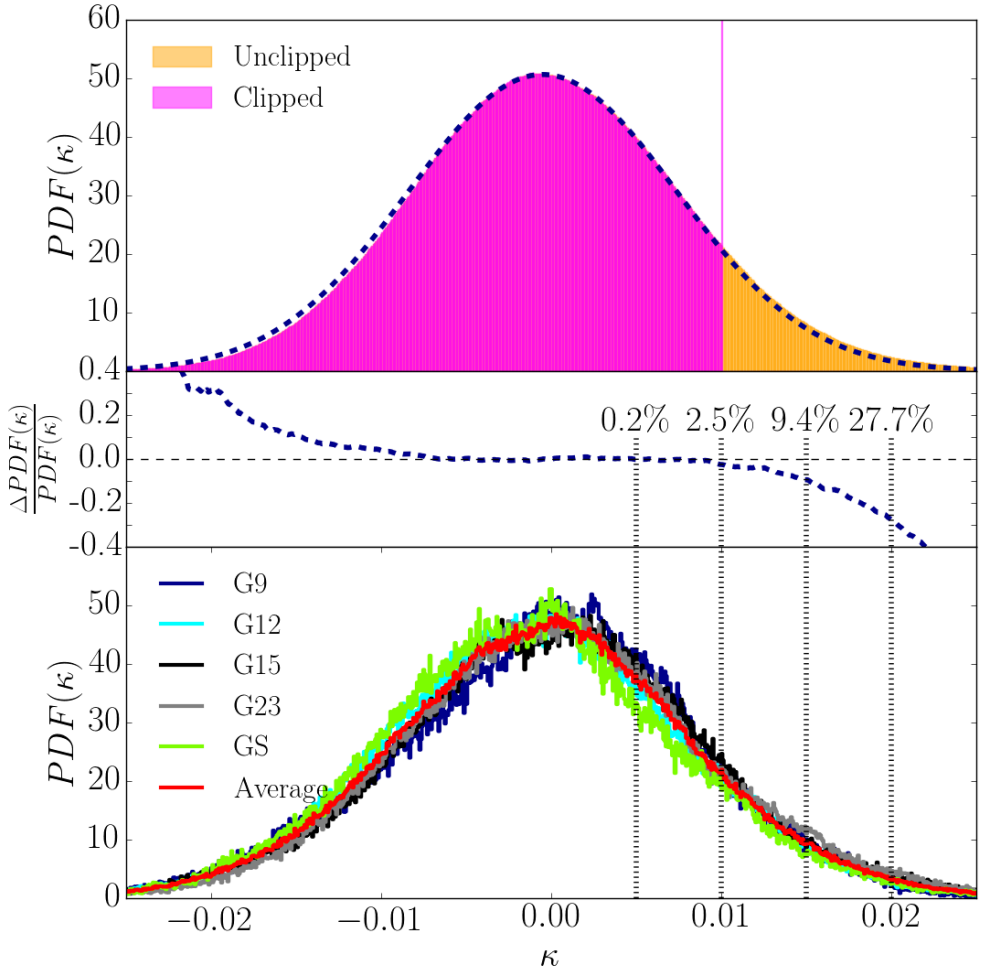

Figure 3.2 *Upper: PDF of the unclipped (orange) and clipped (magenta) convergence, κ, from 50 SLICS realisations, which are identical below the applied clipping threshold of $\kappa^{\rm c} = 0.010$. Also shown is a Gaussian fit to the unclipped distribution (dashed blue). Middle: the fractional difference between the Gaussian fit and the unclipped SLICS PDF(κ). The percentage deviations at $\kappa = \{0.005, 0.010, 0.015, 0.020\}$, designated by the dotted lines, are detailed in the annotations. Lower: The PDFs of the five KiDS-450 patches and their average. A smoothing scale of $\sigma_{\rm s} = 6.6$ arcmin is applied in all cases.*

KiDS-450 patches and their average.

I find that in the range $-0.005 \leq \kappa \leq 0.005$, the PDF of the unclipped SLICS convergence is well described by the Gaussian, but deviations of a few per cent or more arise at $\kappa \gtrsim 0.010$. At the high-end tail of the convergence, the unclipped SLICS PDF is considerably non-Gaussian, differing by $\gtrsim 30\%$. This suggests that a clipping threshold $\kappa^{\rm c} \gtrsim 0.010$ is appropriate for isolating non-linear features of the field. I note that the clipped distribution corresponding to this threshold, shown in the upper panel, being asymmetric and featuring the prominent spike at $\kappa = 0.010$, is indeed less Gaussian than the unclipped. The spike corresponds to the projected peaks exceeding the threshold in the unclipped map, which are distributed throughout the simulated fields of view. The strength of applying this transformation therefore, lies not in Gaussianising the field, but in its clean separation of the linear and non-linear lensing signal.

In setting the value of $\sigma_{\rm s}$, one should aim to reduce the prominence of peaks caused solely by noise fluctuations, but not to the extent that we lose a significant amount of the cosmological information. A comparison of the SLICS convergence maps when clipped at different smoothing scales, with and without intrinsic galaxy shape noise, serves as a useful visual indicator of whether $\sigma_{\rm s}$ is appropriate for the data. Figure 3.3 illustrates the unclipped (left column) and clipped (right column) convergence fields from a single 100 deg^2 SLICS realisation, with a smoothing scale of 2.2 arcmin (upper two panels) and 6.6 arcmin (lower two panels). I chose these values of $\sigma_{\rm s}$, simply to illustrate the substantial differences in the clipped convergence fields these scales facilitate. The first and third panels have no shape noise ($\sigma_e = 0$), whereas the second and fourth panels have shape noise at the level of KiDS-450 (Gaussian distributed with mean zero and $\sigma_e = 0.28$). The clipped fields here have a convergence threshold of $\kappa^{\rm c} = 0.010$. Comparing the first and second panels, smoothed with $\sigma_{\rm s} = 2.2$ arcmin, we see that the features within the clipped and unclipped maps change dramatically when shape noise is introduced. The third and fourth panels however show that the maps change less dramatically with the inclusion of shape noise when the smoothing scale is set to 6.6 arcmin. This indicates that the higher of the two smoothing scales is better suited to SLICS and by extension the data.

An additional test of whether the chosen $(\kappa^{\rm c}, \sigma_{\rm s})$ combination is suitable comes from inspection of the clipped and unclipped correlation functions. The optimal choices for these parameters will facilitate clipping of the non-linear regions

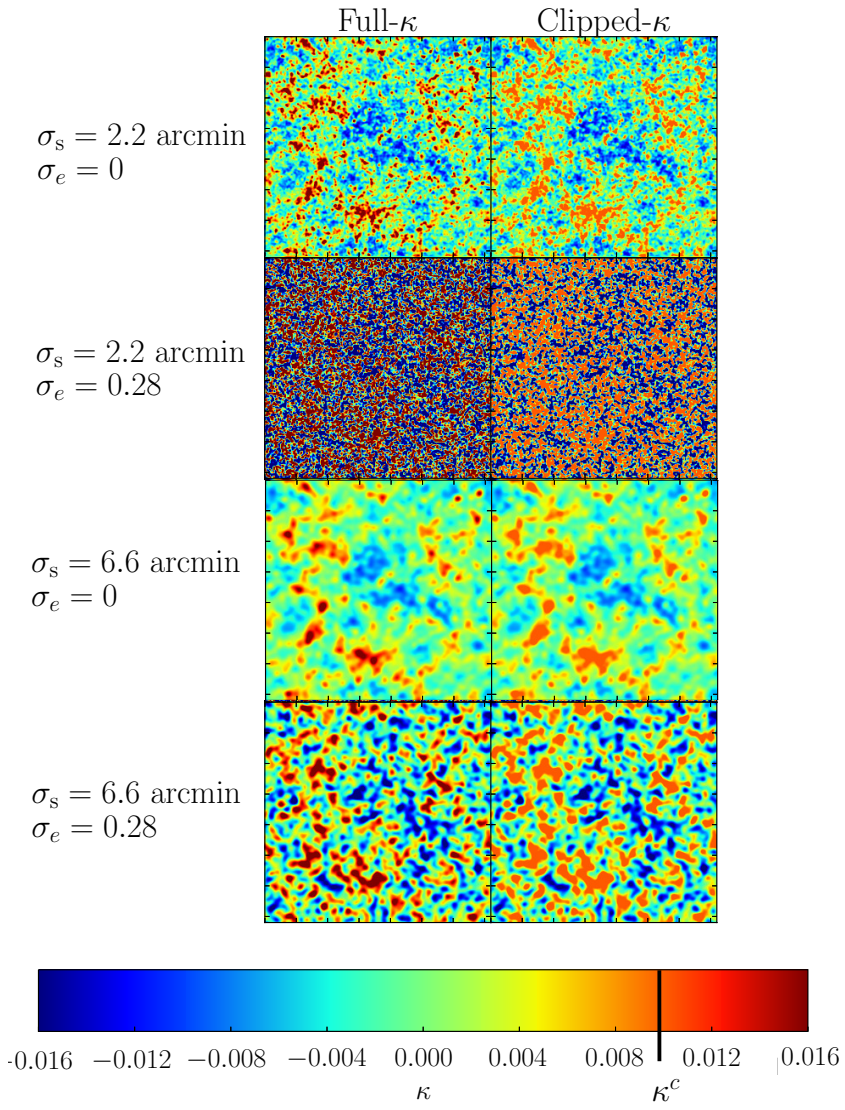

Figure 3.3 *Unclipped (left hand panels) and clipped (right hand panels; $\kappa^c = 0.010$) convergence maps for a single 100 deg^2 SLICS. For the upper two panels, the smoothing scale, σ_s, is equal to 2.2 arcmin. Comparison of these panels shows that the features in both the clipped and unclipped convergence maps for a noise-free field ($\sigma_e = 0$) change dramatically with the inclusion of KiDS-450 level shape noise (Gaussian distributed with width $\sigma_e = 0.28$). The lower two panels however have $\sigma_s = 6.6$ arcmin. Comparison of these panels shows that the clipped/unclipped maps change less dramatically with the inclusion of shape noise if the smoothing scale is set to the higher level. This suggests that using $\sigma_s = 2.2$ arcmin results in the clipping of mainly pure noise features, and that $\sigma_s = 6.6$ arcmin is a more appropriate level of smoothing for clipping the KiDS-450 cosmological signal.*

exclusively, leaving the linear signal untouched. In this case, the unclipped and clipped ξ_+ should converge on the larger, linear angular scales. In Figure 3.4, I present how the $\xi_+^{\rm clip}$ measured from the SLICS are affected by variations in the clipping threshold, smoothing scale and the galaxy shape noise. Similar trends are seen for the $\xi_-^{\rm clip}$ statistic at higher angular scales (I refer the reader to Section 3.4). The left hand panels in this Figure display $\theta\xi_+$, where ξ_+ is the mean unclipped (in solid grey) or clipped (other colours) correlation function measured from the SLICS realisations. The right hand panels display the various correlation functions normalised to that of the unclipped. In calculating the error on the ratios, I take into account the cross-covariance between the clipped and unclipped statistics. The magenta line on all panels is the same and corresponds to $\kappa^{\rm c} = 0.010, \sigma_{\rm s} = 6.6$ arcmin with KiDS-450 level shape noise.

The upper panel of Figure 3.4 illustrates the effect of increasing the clipping threshold from $\kappa^{\rm c} = 0.005$ to 0.010 to 0.015, whilst the smoothing scale is fixed to 6.6 arcmin and the shape noise is fixed to the KiDS-450 level. On average, $26\pm3\%$ of the area of the field is clipped in the case of the most aggressive clipping threshold, $\kappa^{\rm c} = 0.005$, and $3 \pm 1\%$ is clipped in the case of the least aggressive, $\kappa^{\rm c} = 0.015$. We see that when adopting $\kappa^{\rm c} = 0.005$, the clipped signal exhibits a large reduction in power at angular scales around 6 arcmin and a failure to converge with the unclipped at the larger angular scales. The power deprecation is caused by overly aggressive clipping; subtracting too much of the shear signal engenders anticorrelations in the $\xi_+^{\rm clip}$. The excess power at large θ is caused by the smoothing transferring small-scale power to larger scales. This effect is illustrated by considering the convolution of a single δ-function with a Gaussian smoothing kernel; the signal is spread by an extent given by the width of the Gaussian. This panel suggests that $\kappa^{\rm c} = 0.010$ and 0.015 are more appropriate thresholds as they better recover the large scale behaviour of the $\xi_+^{\rm unclip}$.

The variations in the $\xi_+^{\rm clip}$ when the smoothing scale is altered, whilst $\kappa^{\rm c}$ is fixed to 0.010 and the shape noise is fixed to KiDS-450 level, are shown in the middle panel of Figure 3.4. Note the lack of convergence between the unclipped and the clipped signal with $\sigma_{\rm s} = 4.4$ arcmin, indicating over-clipping of the convergence field. We also see that the angular scale at which the loss of power in the $\xi_+^{\rm clip}$ is maximised translates right with increasing smoothing scale. This is due to the loss of signal incurred from smoothing over features of this angular size. The upper and middle panels of Figure 3.4 illustrate the importance of identifying a clipping threshold and smoothing scale which are high enough to diminish the clipping

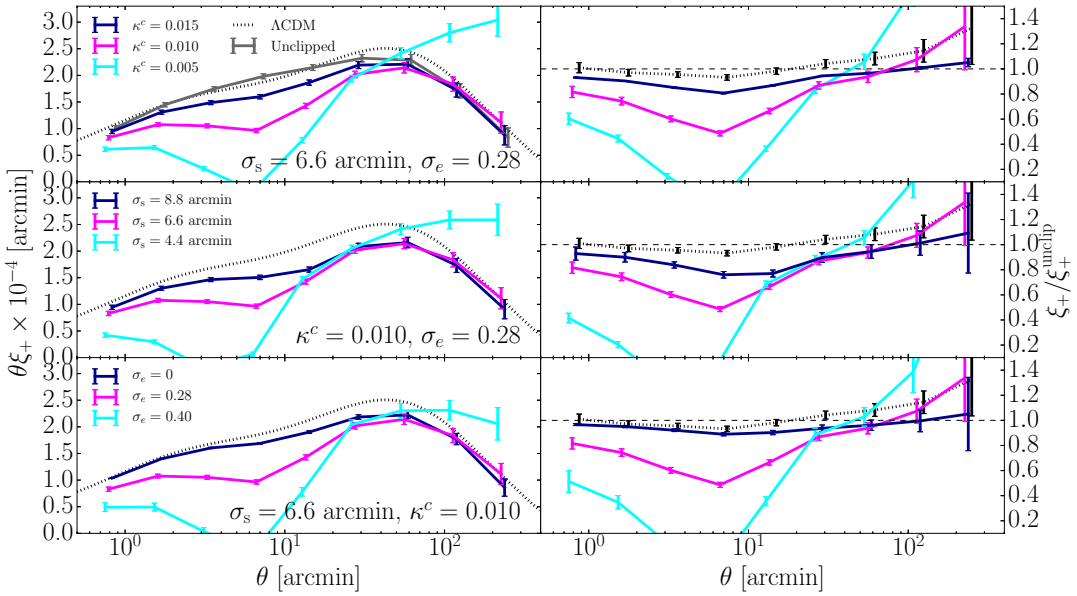

Figure 3.4 *The mean unclipped (solid grey) and clipped (other solid colours) ξ_+ correlation functions measured from the SLICS realisations. The dashed black line is the theoretical unclipped prediction from equation 3.13. The left hand panels display $\theta\xi_+$, the right hand the measurements normalised to the unclipped statistic from SLICS. The annotation in the lower right hand corner of each panel specifies which of the parameters are held constant in the calculations. The upper panel is concerned with variations in the clipping threshold, $\kappa^{\rm c}$, with fixed smoothing scale, $\sigma_{\rm s}$, and shape noise characteristics, σ_e. The middle and lower panels present variations in the smoothing scale and shape noise respectively. The magenta line in all cases depicts the measurement for the fiducial parameters: $\kappa^{\rm c} = 0.010, \sigma_{\rm s} = 6.6$ arcmin and $\sigma_e = 0.28$. The error bars are the error on the mean measurement.*

of pure noise features, but low enough to avoid smoothing out the cosmological content in the clipped statistic.

The lower panel of Figure 3.4 illustrates the sensitivity of the $\xi_+^{\rm clip}$ to the shape noise, whilst $\kappa^{\rm c}$ and $\sigma_{\rm s}$ are fixed to 0.010 and 6.6 arcmin respectively. Where $\sigma_e > 0$ the shape noise is sampled from a Gaussian distribution with width equal to σ_e, whereas $\sigma_e = 0$ refers to a measurement made in the absence of shape noise. Shape noise sampled from the broader Gaussian with $\sigma_e = 0.4$, causes greater proportions of the convergence map to exceed the clipping threshold and hence we see a greater reduction in the power after clipping. This demonstrates the importance of matching the shape noise properties of galaxies in the mocks to the data in order to get a simulated model of the clipped correlation functions. I also note that only a small reduction in the power is observed in the shape-noise-free clipped relative to the unclipped, suggesting that most of the clipped content is shape noise rather than non-linear regions. Nevertheless, I find that this small amount of clipping of non-linear cosmological signal, is sufficient for informing the parameter inference with some independent information, as evidenced by the constraints obtained in Section 3.4 and the cross-correlation coefficient matrices in Appendix 3.A.

Having quantified the effect of different choices of the clipping threshold and smoothing scale with the SLICS, in clipping the KiDS-450 data I adopt the most aggressive clipping parameters that satisfy my requirement that the clipped and unclipped ξ_+ converge within $1\sigma_{\rm mean}$, where $\sigma_{\rm mean}$ is the error on the mean measurement, on large angular scales. This is in order to maximise the difference between the clipped and unclipped statistics and thus enhance the cosmological parameter constraints. Henceforth I set $\kappa^{\rm c} = 0.010$ and $\sigma_{\rm s} = 6.6$ arcmin, and conduct clipping with these parameters on the KiDS-450 data and all simulations.

3.3.4 Calibration Corrections

In this section I discuss various calibration corrections which are necessary in order to use clipping to constrain cosmological parameters in this proof-of-concept study. These corrections, necessitated by the imperfect mass reconstruction in the presence of masks, as well as the the finite box size and low level bias in the simulations, are not intrinsic to the clipping methodology.

Mask bias

Real data is subjected to masking, which complicates all methods seeking to transform the density field. This is because it is unclear how to interpret regions where the density field is unknown. In order to investigate how masking affects the clipped correlation functions, I take a 5×10 deg^2 section of the G9 mask (H17) and concatenate it with a copy of itself, in order to fit the 10×10 deg^2 field of view of SLICS. I apply the resultant mask to each of the realisations.

As expected, the change in the $\xi_{\pm}^{\rm unclip}$ from SLICS when a mask is applied is small, in line with the sampling variance on the measurement. However I find considerable deviations between the measurements of $\xi_{\pm}^{\rm clip}$ from the masked and unmasked SLICS. Figure 3.5 shows the fractional difference between the masked and unmasked clipped and unclipped $\xi_{\pm}$ measured from the SLICS with $\kappa^{\rm c} = 0.010$ and smoothing scale of 6.6 arcmin. The fractional difference for the $\xi_{\pm}^{\rm unclip}$ (in grey) differs from zero by less than 5% across all angular scales whereas that of the clipped (magenta) features considerable deviations at angular scales below 20 arcmin. Deviations of similar magnitude and shape arise when I use masks which have different geometry but reduce the field area by similar amounts. I refer to the influence which the mask has on the clipped measurements as the *mask bias*.

The mask bias arises from the way I handle masks and edge effects in mass reconstruction. I follow the methodology of Van Waerbeke et al. (2013) by setting the convergence to zero in regions where more than 50% of the volume of the Gaussian smoothing window is centred on masked pixels. Where masked regions coincide with high convergence regions, this process causes the convergence surrounding the masked regions to be underestimated, and the overall power in the $\xi_{\pm}^{\rm clip}$ statistics to be diminished. This does not affect the $\xi_{\pm}^{\rm unclip}$ since no mass reconstruction is performed in arriving at these measurements. This issue is not a problem intrinsic to clipping, so much as it is a general issue with mass reconstruction methodology in the presence of masks. This is an active topic of research (see for example VanderPlas et al., 2012; Liu et al., 2014; Jullo et al., 2014; Chang et al., 2017) and rigorously solving this problem is beyond the scope of this Chapter. I instead opt to numerically calibrate and correct for the effect of the mask on my clipped statistics.

I find that for $\kappa^{\rm c} = 0.010$ and $\sigma_{\rm s} = 6.6$ arcmin, mask bias is negligible in

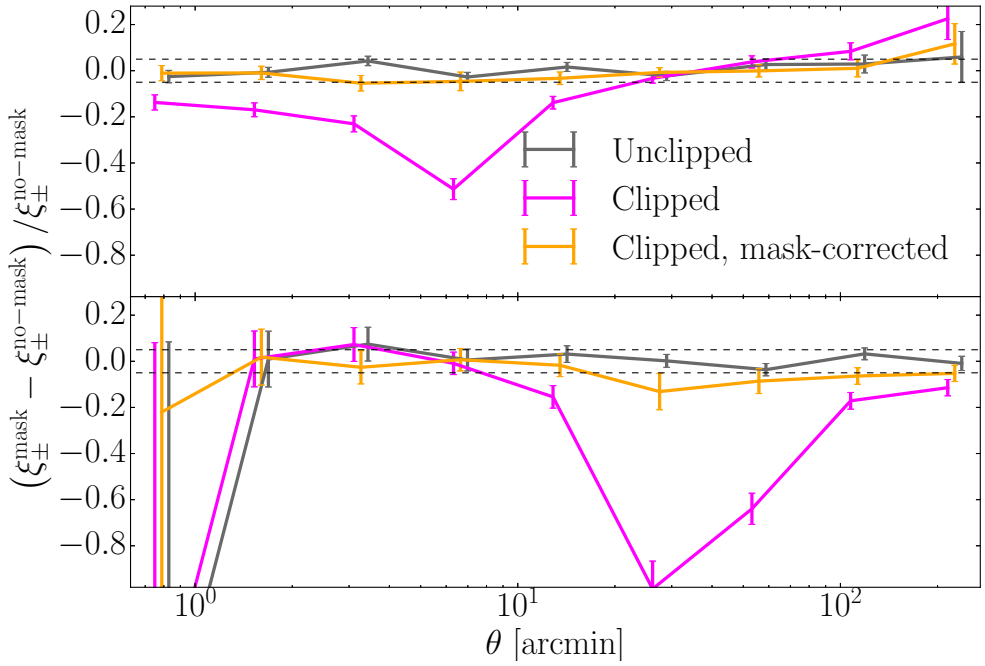

Figure 3.5 *The effect of the mask bias for the clipped and unclipped ξ_+ (upper) and ξ_- (lower) from SLICS. The grey curve shows the fractional difference between the masked and the unmasked $\xi_\pm^{\rm unclip}$ – the fact that this curve has a 5% consistency with zero across all angular scales illustrates that the $\xi_\pm^{\rm unclip}$ is fairly unaffected by masking. The magenta curve shows the fractional difference between the masked and unmasked $\xi_\pm^{\rm clip}$ – the significant deviation from zero illustrates the biasing caused by the mask. The orange curve displays the fractional difference between the masked $\xi_\pm^{\rm clip}$, once corrected for the bias with 100 noise realisations via the methodology discussed in the text, and the unmasked $\xi_\pm^{\rm clip}$. The correction reduces the mask bias to $\lesssim 5\%$ in the case of the $\xi_+^{\rm clip}$; the $\xi_-^{\rm clip}$ however still suffers residual mask-bias at a level of $\sim$10% between 20 and 50 arcmin after I employ my masking correction. The clipped measurements were made with $\kappa^{\rm c} = 0.010$ and $\sigma_{\rm s}$ =6.6 arcmin, and the error bars are measured from the SLICS realisations.*

the absence of galaxy shape noise. Consequently, I assume that for my chosen clipping threshold and smoothing scale, the mask bias is dependent on the level of shape noise and the mask geometry, and independent of the cosmology. This is a reasonable assumption given the statistical power of KiDS-450. This assumption prompts me to investigate the effect of the mask on fields consisting of pure galaxy shape noise and zero lensing. I model the mask bias correction to the clipped correlation function as

$$\text{mask bias} = {\xi_\pm}^{\text{mask, noise}} - {\xi_\pm}^{\text{no-mask, noise}}, \tag{3.17}$$

where ${\xi_\pm}^{\text{mask, noise}}$ and ${\xi_\pm}^{\text{no-mask, noise}}$ are the average measurements from fields of pure Gaussian shape noise, with mean zero and $\sigma_e = 0.28$, which are masked/unmasked respectively. By subtracting the mask bias correction from the clipped correlation functions calculated with a mask applied, I find that the influence of the mask can be mostly corrected for.

Figure 3.5 displays the $\xi_\pm^{\text{clip}}$ corrected for the mask bias (modelled in equation 3.17), using 100 noise fields, in orange. The mask applied here is that of the G9 patch reformatted to fit the SLICS lightcone, but I verify that I obtain the same results for the corrected $\xi_\pm^{\text{clip}}$ if I apply a different mask to the SLICS and recompute the correction specific to said mask. I find that the corrected ξ_+^{clip} is consistent with the measurement made in the absence of masking to within 5%. Although the corrected ξ_-^{clip} is much closer to the unmasked than the masked measurement, I find that the mask bias remains present at a $\sim$10% level at angular scales of $\sim$30 arcmin. A larger number of noise realisations does not reduce the mask bias further, implying that a more sophisticated treatment of the masks is critical if clipping is to be used in future cosmological analyses. The residual mask bias affecting the ξ_-^{clip} measurement, combined with the fact that ξ_-^{unclip} is the least powerful shear correlation function in terms of constraining cosmological parameters, motivates me to continue in this analysis using the ξ_+^{clip} and ξ_+^{unclip} statistics only.

I proceed to compute and correct the mask bias for each of the KiDS-450 patches individually. The corrections for each of the patches are similar, which is expected given the masks cause a similar reduction in effective area per patch. All clipped correlation functions from KiDS-450, presented in this Chapter and used in the likelihood analysis in Section 3.4, have been corrected for mask bias, whereas

all those from the simulations were computed without masks applied. As this is a proof-of-concept, I do not propagate the error on the mask bias through to the cosmological constraints with KiDS-450, as I want to see the improvement obtained through clipping in a scenario where the mask bias is under control.

Finite box effects

The DH10 simulations span a broad range in the $\Omega_{\rm m}$-σ_8 parameter space at the cost of having a small number of realisations per cosmology and a small box size relative to the SLICS (see Figure 3.1 and Table 3.1 for details). In simulations, the finite size of the box means that the matter power spectrum $P_\delta(k, \chi)$, appearing in equation 3.14, is limited by two scales: $k_{\rm min} = 2\pi/L_{\rm box}$, where $L_{\rm box}$ is the size of the simulation box, and $k_{\rm max} = 2\pi/L_{\rm res}$, where $L_{\rm res}$ is the smallest scale which can be resolved in the simulation. The missing modes with $k < k_{\rm min}$ cause the unclipped shear correlation functions expressed in equation 3.13 to lose power at large angular scales (see for example Harnois-Déraps & van Waerbeke, 2015). Similarly, the missing modes with $k > k_{\rm max}$ engender a loss of power at small θ. The effect of the resolution of DH10 is not prominent at the angular scales probed by my measured shear correlation functions, as is evidenced by the consistency between the theoretical and mock $\xi_+^{\rm unclip}$ at angular scales < 10 arcmin, shown in Figure 3.6. On the other hand, the k-modes absent due to the box size do cause the DH10 $\xi_+^{\rm unclip}$ to be underestimated on angular scales > 10 arcmin. I therefore need to correct for the effect of the finite box in order to constrain the cosmology of the real Universe using the DH10 $\xi_+^{\rm clip}$ measurements.

I obtain cosmology- and angular-scale-dependent corrections for the finite box effect on $\xi_+^{\rm unclip}$ by measuring the difference between the theoretical prediction from equation 3.13 in a non-truncated box, and the prediction within a box of size $L_{\rm box}$. For these predictions I use the NICAEA code from Kilbinger et al. (2009) with the HALOFIT model from Smith et al. (2003), since it is a better match to the DH10 $\xi_+^{\rm unclip}$ than that of Takahashi et al. (2012). The correction for the loss of power at large angular separations, due to the finite box, is robust to the choice of HALOFIT model however, since Smith et al. (2003) and Takahashi et al. (2012) converge at these scales.

The obvious choice for the size of the truncated box used in calibrating the finite box effect is that in which the DH10 were created, 140 h^{-1}Mpc. The theoretical

$\xi_+^{\rm unclip}$ from a box of this size however overestimates the loss of power at large θ seen in DH10. This is because the simulations are constructed as a lightcone through the box, resulting in a smooth decay in the power, in contrast to a sharp cut-off at the $L_{\rm box}$ scale. I follow Sellentin et al. (2017), by modelling the finite box effect with an effective cut-off, performing a χ^2 fitting of the theoretical $\xi_+^{\rm unclip}$ for different values of the box size to the shape-noise-free mean measurement from the fiducial DH10 cosmology. I fit the box size for the fiducial cosmology only, on account of there being the largest number of realisations and thus the lowest sampling variance overall (though I stress that the corrections I apply are specific to each cosmology). I use the covariance matrix measured from the 175 realisations for the fiducial DH10 cosmology, rather than the one from SLICS, since the former will better describe the uncertainty on DH10. Furthermore, I use only the five angular separation points > 10 arcmin in the fitting as I am most interested in finding the effective box size that best describes the large-scale behaviour of the mocks where the effect of the finite box size becomes relevant. I find that the $\xi_+^{\rm unclip}$ for the fiducial cosmology of DH10 is best described by the prediction in an *effective* box size of 250 h^{-1}Mpc. This prediction, shown by the magenta curve in Figure 3.6, fits the DH10 measurement well, with a χ^2 of 4.99 for the 4 degrees of freedom. The correction for the finite box size for this cosmology is the difference between the theoretical prediction from the non-truncated box (shown in dark blue), and the truncated box prediction.

The lack of a theoretical prediction for $\xi_+^{\rm clip}$ limits my inference of the finite box effect for this statistic. I assume therefore that the loss of power in the clipped correlation functions due to the finite box effect is equal to that of the unclipped, and so the calibration correction I derive for the unclipped correlation functions per cosmology, is applicable also to the clipped. This assumption is likely to be valid since the effect of the finite box is most prominent on scales where $\xi_+^{\rm unclip}$ and $\xi_+^{\rm clip}$ converge. I also test how much the marginalised means and 68% confidence intervals on the cosmological parameters change when the finite box correction is included/omitted and find that the effect is small and does not change my conclusions. This approach is suitable for this proof-of-concept analysis and the correction can easily be circumvented in the future with the use of larger simulations such as the Mira Titan suite (Heitmann et al., 2016), or the *cosmo*-SLICS suite (Harnois-Deraps, Giblin & Joachimi, 2019) presented in the next Chapter.

I compute individual calibration corrections for each of the 158 DH10 cosmologies,

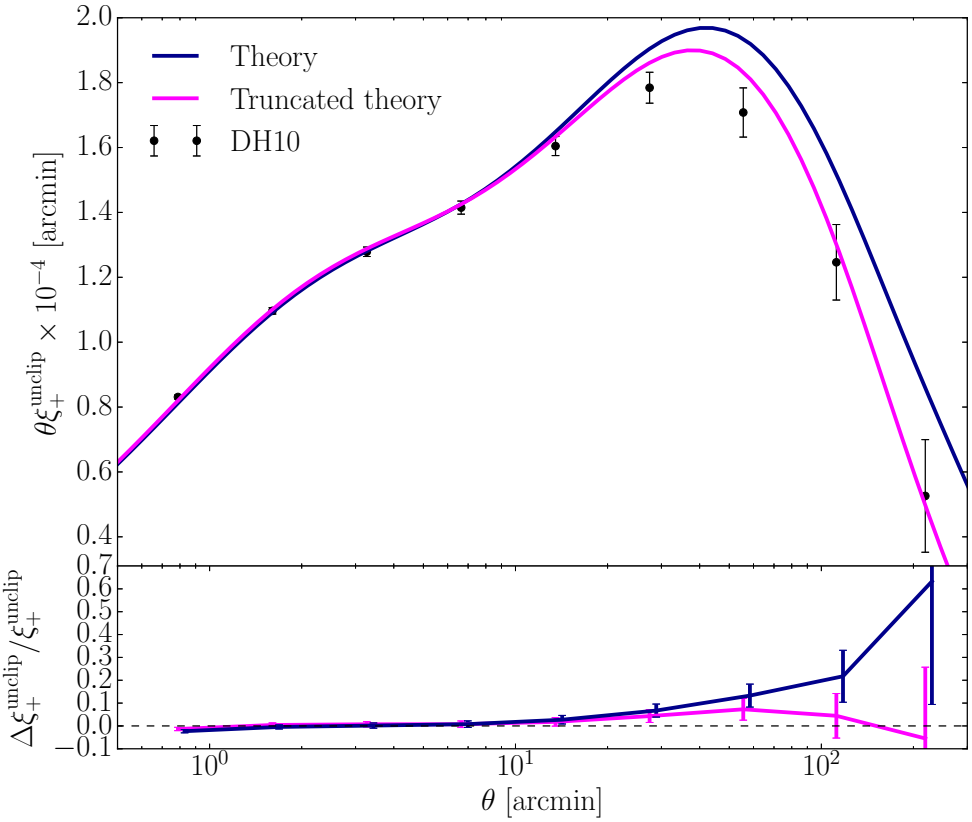

Figure 3.6 *Upper: the noise-free $\xi_+^{\rm unclip}$ measured for the fiducial cosmology of the DH10 simulations (data points), the theoretical prediction from a non-truncated box (dark blue; equation 3.13), and the theoretical prediction from a truncated box of size $L_{\rm box} = 250\,h^{-1}Mpc$ (magenta). The error bars on the data points come from the dispersion across the 175 realisations (35 simulations × 5 ray-tracing angles) for this cosmology. The difference between the dark blue and magenta lines is the finite box correction I apply to the DH10 measurement. Lower: the fractional difference between the theoretical $\xi_+^{\rm unclip}$ predictions and the DH10 measurement.*

using the box size fit to the fiducial cosmology. I then additively scale up the whole angular separation range of the clipped and unclipped ξ_+ from the simulations (the small scales remaining practically unchanged by the calibration). An additive, rather than a multiplicative, correction is appropriate for accounting for the missing k-modes in the integration over $P_\delta(k, \chi)$ in equation 3.14. The correction I apply also has the benefit of not inflating the noise in the DH10 predictions.

The SLICS are also affected by the limitations of a finite box, though the box size is larger than that of DH10, engendering a loss of power at the largest angular scales that is of order 10%-30% (I refer the reader to the ratio of the theoretical and SLICS $\xi_+^{\rm unclip}$ shown in Figure 3.4). In general the covariance that I calculate from SLICS will be affected by the loss of power in the correlation functions, but since the correction for the finite box in DH10 has a very small impact on the cosmological parameter constraints, and this effect is much smaller for SLICS, I therefore treat the SLICS covariance matrices as unbiased by the box size. I note however that returning to the effect of the finite box on covariance estimation is an important topic for future work.

Cosmological bias

In Figure 3.6 I show that the fiducial DH10 cosmology reproduces the expected $\xi_+^{\rm unclip}$, modulo a small correction for the finite box effect on large scales. In the upper panel of Figure 3.7 I compare noise-free measurements of $\xi_+^{\rm unclip}$, corrected for the box size, with theoretical predictions (equation 3.13), now for the full range of 158 cosmologies spanned by the DH10 simulations. Binning the relative difference by the input cosmology S_8 (see colour-bar), we see a trend where the low S_8 simulations tend to underestimate $\xi_+^{\rm unclip}$ by $\sim$7% between angular separations of 1 and 110 arcmin, whereas high S_8 simulations overestimate by $\sim$10% in this range. The cause of this cosmological bias, which is present irrespective of whether the finite box correction is applied, is currently unknown. I do not devote time to uncovering its origin, owing to the development of *cosmo*-SLICS (Harnois-Deraps, Giblin & Joachimi, 2019), a suite of numerical simulations[3] presented in Chapter 4, which supersede DH10. The bias is less than the level of uncertainty due to shot noise and sampling variation in the DH10 $\xi_+^{\rm clip}$ predictions (see Section 3.3.5) that increases from 5 to $\sim$100% over the full angular range, shown by the grey

[3]Completed shortly after the publication of this work.

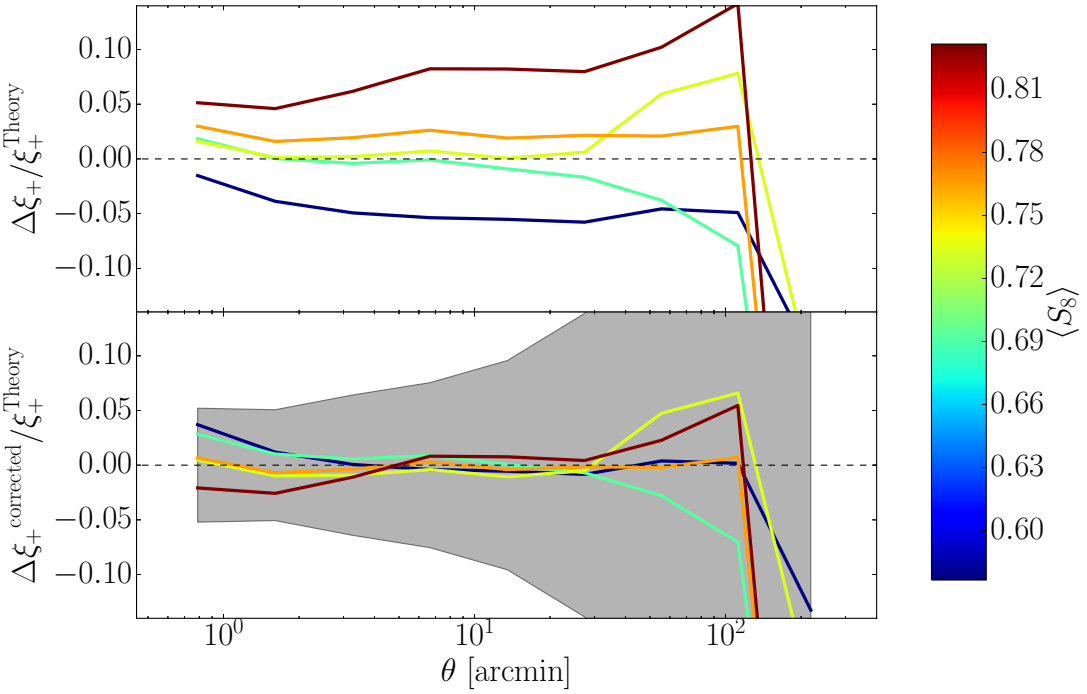

Figure 3.7 *Upper panel: the fractional difference between the 158 shape-noise-free DH10 $\xi_+^{\rm unclip}$ measurements and the theoretical predictions (equation 3.13) binned in terms of $S_8 = \sigma_8(\Omega_{\rm m}/0.3)^{0.5}$, with the colours designating the mean S_8 in each bin. We see that the low S_8 measurements underestimate $\xi_+^{\rm unclip}$, whereas the high S_8 measurements overestimate. Lower panel: the same measurements but corrected for the cosmological bias via the methodology discussed in the text. Any remaining bias can be compared to the uncertainty on the clipped predictions (shaded grey) that is included in my analysis when using the DH10 simulations.*

shaded region in the lower panel. It is therefore accounted for, to some extent, in my clipped analysis that includes an error budget to account for this level of uncertainty in the DH10 predictions. Nevertheless, I employ a correction scheme to ensure that this systematic does not artificially contribute to the improvements yielded by the combined clipped-and-unclipped analysis.

I determine a cosmological bias correction by averaging the relative difference between the shape-noise-free DH10 and theoretical $\xi_+^{\rm unclip}$ between 1 and 60 arcmin, where the bias varies slowly, in each of the five S_8 bins shown in Figure 3.7. This produces a smooth mean-bias function which monotonically increases from -5% in the lowest S_8 bin to $+8\%$ in the highest. I obtain the mean-bias for each of the 158 DH10 cosmologies by linearly interpolating/extrapolating from

this function for the simulation S_8 values. The corrected $\xi_+^{\rm unclip}$ is obtained by multiplicatively scaling the DH10 measurements by $1/[1 + b(S_8)]$, where $b(S_8)$ is the mean-bias corresponding to the S_8 value of a given simulation. The relative differences between the corrected DH10 $\xi_+^{\rm unclip}$ and the theoretical measurements are shown in the lower panel of Figure 3.7 for the five S_8 bins, and can be compared to the uncertainty included in the clipped predictions $\xi_+^{\rm clip}$ (shaded grey), which is incorporated in my cosmological parameter constraints.

As was the case with the finite box effect (see Figure 3.6), it is not possible to ascertain the extent to which the clipped predictions are affected by the cosmological bias in DH10, owing to the lack of a theoretical clipped statistic. Hence I again assume that the $\xi_+^{\rm clip}$ is biased in the same way as the corresponding $\xi_+^{\rm unclip}$ measurement. I find that my conclusions are not significantly changed however if the cosmological bias is unaccounted for; the combined clipped-and-unclipped analysis increases the constraining power by 20%, instead of 15% when the bias is corrected. I note that this bias was unaccounted for in the peak statistics analyses of DH10, Kacprzak et al. (2016) and Martinet et al. (2018), and their results will likely be affected.

3.3.5 Cosmological dependence of clipping

Although there is a large number of shear catalogues for the fiducial DH10 cosmology (35 independent simulations $\times$ 5 pseudo-independent catalogues corresponding to 5 different ray-tracing angles), there exist only 5 catalogues for the remaining 157 cosmologies. The average $\xi_\pm^{\rm clip}$, measured across each set of non-fiducial DH10 cosmologies, is therefore more significantly impacted by shot noise in comparison to the fiducial set. In the case of the unclipped correlation functions one can simply turn off the noisy galaxy ellipticities. However, as is discussed in Section 3.3.3, I find that the clipped correlation functions are critically dependent on the shape noise. This necessitates the inclusion of shape noise such that the noise properties of the mocks match the data.

In order to reduce the impact of the shot noise whilst still including the effects of the galaxy shape noise, I determine the clipped correlation functions from DH10 with different realisations of the shape noise. I find that averaging $\xi_+^{\rm clip}$ across 75 or more noise realisations is sufficient for the measurement from each of the individual catalogues of the fiducial DH10 cosmology to stabilise. This averages

away any bias in the measurement caused by a single realisation of the shape noise. I proceed to compute 75 noise realisations per catalogue for all of the DH10 cosmologies; the $\xi_+^{\rm clip}$ for each cosmology appearing in the likelihood analysis is the average over these. The remaining source of noise in the DH10 mocks is then the sampling variance across different catalogues of a given cosmology. In order to include this source of uncertainty in my likelihood analysis, I measure the covariance across the 175 clipped and unclipped ξ_+ from the fiducial DH10 cosmology, each of which is averaged across 75 noise realisations, via equation 3.16. These covariance matrices, which are at the level of 5% in the first angular separation bin (0.8 arcmin), increasing to $\sim$100% in the last bin (220 arcmin; see Figure 3.7), encompass my uncertainty on the model, both in terms of sampling variance and cosmological bias (see Figure 3.7). I add this error in quadrature to the error measured from the SLICS which describes the uncertainty in the data itself. This is discussed in more detail in Section 3.4.1.

In the upper panel of Figure 3.8, I present the clipped (upper), unclipped (middle), and the ratio (lower) for all of the DH10 cosmologies, each of which is averaged over the 75 realisations of the shape noise, with $\kappa^{\rm c} = 0.010$ and $\sigma_{\rm s} = 6.6$ arcmin. All measurements have been corrected for the finite box effect and the cosmological bias (discussed in the previous Section). In general the power in the $\xi_+^{\rm clip}$ increases with S_8 in a similar capacity to the $\xi_+^{\rm unclip}$. The prominent reduction in power at angular scales $\sim$5 arcmin is also a common feature for all of the cosmologies. I observe a number of the low S_8 cosmologies with small or negative ratios at small angular separations. This effect is not caused by these cosmologies experiencing a greater degree of clipping; indeed we see that in general less of the field is clipped for lower S_8 cosmologies as expected. Rather, this is the result of these fields being dominated by shape noise. Smoothing these fields correlates the shape noise, and clipping then leads to a reduction in power and even anticorrelations to be seen in the $\xi_+^{\rm clip}$ for these low S_8 cosmologies. This is not observed in the higher S_8 measurements which have higher signal to noise, and consequently maintain larger power in the correlations throughout the clipping pipeline. In the case of the low S_8 cosmologies, smaller values of $\sigma_{\rm s}$ and $\kappa^{\rm c}$ would have been more suitable for the clipped analysis.

The lower panel of Figure 3.8 shows these measurements in the absence of shape noise, to verify that in this case the low S_8 cosmologies experience no clipping, and the $\xi_+^{\rm clip}$ and $\xi_+^{\rm unclip}$ converge at all scales. Intuitively we see lower clipped-to-unclipped ratios as S_8 increases, due to the greater degree of clipping of the

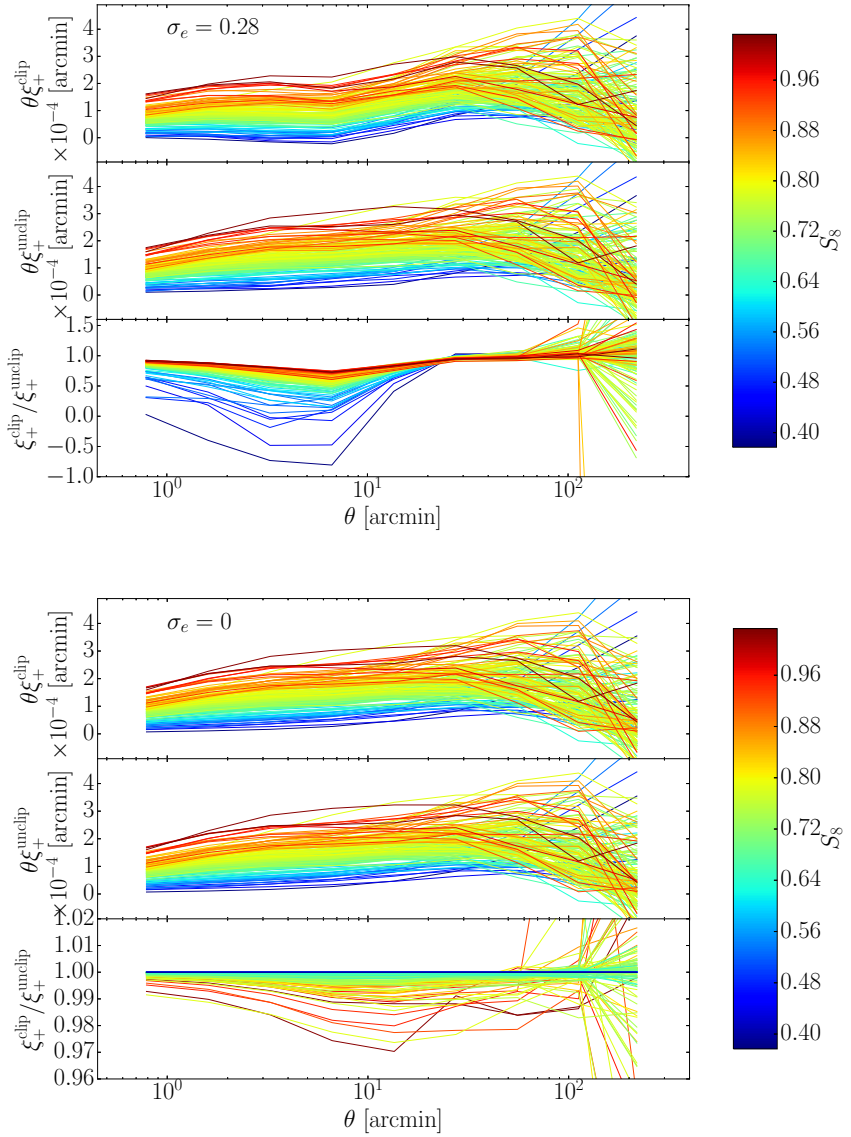

Figure 3.8 *Upper panels:* $\theta\xi_+^{\rm clip}$, $\theta\xi_+^{\rm unclip}$ *and the ratio for all of the DH10 cosmologies, each of which is averaged over 75 shape noise realisations, colour-coded by* S_8. *The clipping threshold and smoothing scale are* $\kappa^{\rm c} = 0.010, \sigma_{\rm s} = 6.6$ *arcmin respectively, selected via the methodology in Section 3.3.3. The low clipped-to-unclipped ratios seen at* < 10 *arcmin for low* S_8 *cosmologies are brought about by clipping shape noise only. Lower panels: the same measurements but with zero shape noise. The low* S_8 *cosmologies are not subject to clipping in this case, such that the clipped and unclipped* ξ_+ *converge at all angular scales. All measurements have been corrected for the finite box effect and cosmological bias (see Section 3.3.4).*

cosmological signal. As with the shape-noise-free $\xi_+^{\rm clip}$ from SLICS presented in Section 3.3.3, we see once again that clipping the non-linear signal causes only a small change in the correlation functions relative to the unclipped, but this small effect is ample for considerably informing the parameter inference (see Section 3.4). This highlights the importance of selecting a smoothing scale and clipping threshold which are well suited to the properties of the data, in order to clip the cosmological signal rather than just the noise. One need not alter $\kappa^{\rm c}$ and $\sigma_{\rm s}$ in this analysis however; the cosmologies in the extreme S_8 tail, are flagged as ill-fitting cosmologies to the data in the likelihood analysis, by virtue of the low power correlations/anticorrelations brought about by clipping noise only.

3.4 Results

In Figure 3.9 I present the $\xi_\pm^{\rm clip}$ measured from KiDS-450 and SLICS produced with a clipping threshold of $\kappa^{\rm c} = 0.010$ and smoothing scale $\sigma_{\rm s} = 6.6$ arcmin (see Section 3.3.3). The left hand panels of this Figure display $\theta\xi_\pm$, the right displays the measurements normalised to the unclipped. The error bars come from the SLICS covariance (rescaled to the effective area of KiDS-450 in the case of the data), and I include the cross-covariance between the clipped and unclipped in the error on the ratios. We see similar trends in the clipped measurements between the mock and the data, which is expected given that SLICS are tailored to reflect KiDS-450.

3.4.1 Likelihood analysis

I proceed to the likelihood analysis to constrain $S_8 = \sigma_8(\Omega_{\rm m}/0.3)^{0.5}$, with the other parameters fixed to the DH10 fiducial values, $\Omega_{\rm b} = 0.04, n_{\rm s} = 1.0, h = 0.7$ and $w_0 = -1$ (investigation of the dependence of clipping on these latter two parameters is conducted in Chapter 4). I use only the clipped and unclipped ξ_+, omitting the ξ_- for the reasons argued in Section 3.3.4, and all nine θ-bins, logarithmically spaced from $\sim$0.8 to $\sim$220 arcmin. The products required to constrain these cosmological parameters are the clipped and unclipped auto- and cross-covariance matrices from SLICS, which describe the uncertainty on the data, those measured from the fiducial cosmology of DH10, which describe the uncertainty on the predictions themselves, and the $\xi_+^{\rm clip}$ predictions from

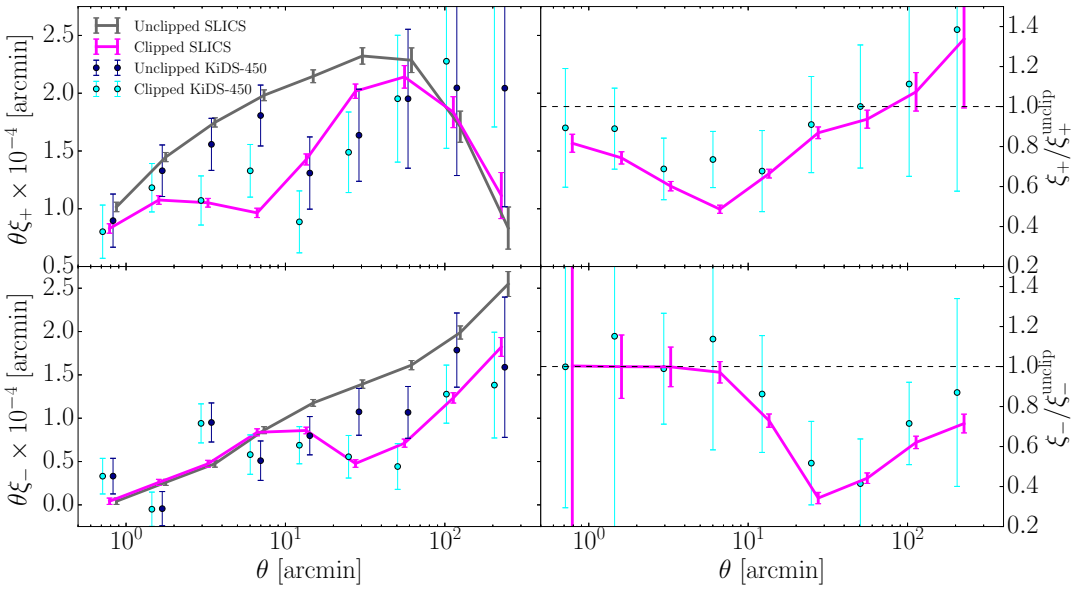

Figure 3.9 *The clipped and unclipped ξ_+ (upper) and ξ_- (lower) for KiDS-450 (data points) relative to those from the fiducial cosmology of SLICS. The left hand panels display $\theta\xi_\pm$, the right hand the ratio of the clipped and unclipped measurements. The errors plotted for SLICS are that of the mean measurement. The error bars on KiDS-450 are equal to those of SLICS scaled by the ratio of the effective unmasked areas. The mock and the data were clipped with the fiducial parameters $\kappa^{\rm c} = 0.010$ and $\sigma_{\rm s} = 6.6$ arcmin.*

DH10. Rather than use the $\xi_+^{\rm unclip}$ from DH10 in the likelihood analysis, I use the more accurate theoretical predictions (see equation 3.13) evaluated at the DH10 cosmologies, from NICAEA, which are free of the noise and low-level cosmological bias (Section 3.3.4) present in the simulations. When constraining the cosmology of a test data set from DH10 of known cosmology, I use the HALOFIT model from Smith et al. (2003), as this matches these simulations more closely than the HALOFIT model from Takahashi et al. (2012). However, when constraining the cosmology of KiDS-450 I use the latter model, since it better describes the $\xi_+^{\rm unclip}$ on small, non-linear angular scales. I find that the combined clipped-and-unclipped analyses improve cosmological parameter constraints over the unclipped alone, irrespective of whether I use the simulated or theoretical $\xi_+^{\rm unclip}$. This is discussed further in Appendix 3.B.1. I also find that the combined constraints are an improvement upon the unclipped irrespective of which θ-scales are used in the likelihood analysis. The improvements do however tend to zero when the angular scales are restricted to the range where the clipped and unclipped converge.

The Bayesian posterior probability distribution for a particular set of cosmological parameters $\boldsymbol{\pi}$ given a data vector $\mathbf{d}$ is given by

$$p(\boldsymbol{\pi}|\boldsymbol{d}) = \frac{\mathcal{L}(\boldsymbol{d}|\boldsymbol{\pi})p(\boldsymbol{\pi})}{E}\,, \tag{3.18}$$

where $\mathcal{L}(\boldsymbol{d}|\boldsymbol{\pi})$ is the likelihood, $p(\boldsymbol{\pi})$ is the prior probability of the cosmological parameter configuration $\boldsymbol{\pi}$ and E is the evidence, which normalises the integral of the posterior over all possible values of $\boldsymbol{\pi}$ to unity. I adopt a wide tophat prior over $\boldsymbol{\pi}$ which goes to zero where the likelihood becomes very small. Hence, the posterior probability in this case is simply proportional to the likelihood, given by

$$\mathcal{L}(\boldsymbol{d}|\boldsymbol{\pi}) \propto \exp\left(-\frac{1}{2}\left[\boldsymbol{d}-\boldsymbol{m}(\boldsymbol{\pi})\right]^\intercal \Sigma^{-1}\left[\boldsymbol{d}-\boldsymbol{m}(\boldsymbol{\pi})\right]\right)\,, \tag{3.19}$$

where the model prediction $\boldsymbol{m}(\boldsymbol{\pi})$ represents either the theoretical $\xi_+^{\rm unclip}$ from equation 3.13, or the $\xi_+^{\rm clip}$ from DH10. The data vector $\boldsymbol{d}$ of course takes the form of the clipped and unclipped ξ_+ from the data. Σ is the true covariance matrix describing the uncertainties affecting statistical inference. When computing the combined clipped-and-unclipped constraints, Σ is built out of the auto-covariance

matrices for the unclipped and clipped ξ_+, as well as the cross-covariance between them.

Typically, uncertainties arise from the sampling variance in the data; here, I approximate this with the covariance matrix, $C_{\rm data}$, measured from the SLICS and rescaled to the effective area of the data. However, in this analysis I also have uncertainty on the clipped model predictions owing to the noise in the DH10 simulations. I incorporate these two independent sources of error by assuming $\Sigma \simeq C = C_{\rm data} + C_{\rm model}$, where $C_{\rm model}$ describes the covariance of the predictions $\boldsymbol{m}(\boldsymbol{\pi})$. The clipped auto-covariance component of $C_{\rm model}$ is measured across the various noise realisations for each of the catalogues for the fiducial DH10 cosmology, as is discussed in Section 3.3.5. Using the theoretical predictions from equation 3.13 for $\xi_+^{\rm unclip}$, which are free of noise, causes the unclipped auto-covariance, as well as the clipped-unclipped cross-covariance components within $C_{\rm model}$ to be zero. If I were to use the DH10 unclipped predictions instead, these elements are non-zero, and are again measured from these mocks (see Appendix 3.B.1). In this case, comparison of the diagonal elements of the clipped and unclipped parts of $C_{\rm data}$ and $C_{\rm model}$, reveals that $C_{\rm data}$, and hence the survey size of KiDS-450, is the dominant source of uncertainty, by a factor of $\sim$20 in the lowest angular separation bin, decreasing to $\sim$2 in the largest bin.

Although the approximated covariance, C, is assumed to be an unbiased estimate of the true covariance, Σ, since it is calculated from simulations featuring noise, its inverse, C^{-1}, is a biased estimate of Σ^{-1} which appears in equation 3.19. This means that one cannot readily substitute C^{-1} into this expression. Hartlap et al. (2007) advocate a correction whereby the inverse covariance is rescaled[4] according to,

$$\widehat{C^{-1}} = \frac{N - D - 2}{N - 2} C^{-1}. \tag{3.20}$$

Here N is the number of simulations employed in estimating the covariance matrix C containing $D \times D$ elements. In my analysis C is the summation of $C_{\rm data}$ and $C_{\rm model}$, each of which have different Hartlap correction factors. This complicates efforts to obtain an unbiased estimate of the inverse covariance. However, the number of realisations, N, used to calculate the two matrices (906 for the data[5],

[4] Although see Sellentin & Heavens (2016) for a more rigorous correction scheme.

[5] After my clipping pipeline was run on these 906 SLICS realisations, 26 more where added to the ensemble presented in Harnois-Déraps et al. (2018). Given the negligible impact this

175 for the clipped model) greatly exceeds D, the number of θ bins in my correlation functions, (equal to 9 in the case of the separate clipped and unclipped analyses, and 18 for the combined). Thus $C_{\rm data}$ and $C_{\rm model}$ are sufficiently well estimated for me to safely neglect the Hartlap correction in my likelihood analysis.

My cosmological constraints derive from an evaluation on a fine grid within the parameter space. In the case of the clipped analysis, I obtain 2D likelihood surfaces by interpolating from the DH10 cosmologies onto $\Omega_{\rm m}$-σ_8 and $\Omega_{\rm m}$-S_8 grids. My 1D constraints on S_8 are then obtained by marginalising in the $\Omega_{\rm m}$-S_8 space. Although I have a theoretical prescription for the $\xi_+^{\rm unclip}$ as a function of cosmology (equations 3.13–3.14), I chose to also interpolate the theoretical unclipped model from the DH10 cosmologies in order to facilitate a direct comparison between the clipped and unclipped results.

An open question is whether this interpolation should be performed at the level of the clipped and unclipped ξ_+ or at the level of the likelihoods. If one interpolates the model, the cosmological parameter constraints are dependent on the square of any systematic bias which could potentially reside in the interpolation, whereas the dependence is only linear if one interpolates the likelihoods. I try both methods and find that extrapolating the likelihoods outside of the range of the DH10 cosmologies, is more reliable than extrapolating the model. Thus in this Section, I present the results having interpolated the DH10 likelihoods. I demonstrate in Appendix 3.B.2 however, that overall my conclusions are unchanged for a range of different interpolation schemes.

I follow Martinet et al. (2018) and interpolate from the DH10 cosmologies using radial basis functions, employing the SCIPY.INTERPOLATE.RBF Python function set to the multiquadratic model[6]. Whereas the unclipped predictions are noiseless and come from theory, the clipped predictions, from DH10 have added uncertainty. Consequently, the interpolated clipped and combined likelihoods are fairly noisy, featuring spurious spikes which fracture the 68% and 95% clipped and combined contours. I apply a small amount of smoothing in the interpolation to reduce this effect and obtain cohesive contours. So that the clipped and combined contours can be directly compared to the unclipped, I apply the same level of smoothing when interpolating the unclipped predictions also. I verify with the unclipped statistic that the interpolation does not considerably affect the

would have on my analysis, I did not include them.

[6] `https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.Rbf.html`

recovered cosmology relative to a standard grid-based likelihood method without interpolation.

3.4.2 Cosmological constraints

DH10 constraints

Before constraining the cosmology of the KiDS-450 data, I investigate the power of combining the clipped and unclipped ξ_+ statistics for a case where the cosmology is already known. Since I only have clipped measurements at the cosmologies of the various simulations at my disposal, the natural choice for the "data" in this test is the clipped and unclipped ξ_+ corresponding to the fiducial DH10 cosmology. Specifically, I take a subset of the simulations with this cosmology spanning 360 deg^2, the unmasked area of KiDS-450, as the data. I also omit the $\xi_+^{\rm clip}$ and $\xi_+^{\rm unclip}$ with the fiducial cosmology from the predictions, such that there is no "perfect match" between the predictions and the data I am constraining the cosmology of, as is the case when working with real data. All cosmological constraints presented hereafter have the corrections for the finite box size and cosmological bias applied, not only to the predictions, but also to the data from DH10. I have verified that I better recover the known input cosmology with these calibrations included.

The upper panel of Figure 3.10 shows the clipped (magenta), unclipped (orange) and combined (black) 68% and 95% constraints on the fiducial DH10 cosmology, in the $\Omega_{\rm m}$-σ_8 parameter space. The lower panel of this Figure shows the constraints in the $\Omega_{\rm m}$-S_8 parameter space, where $S_8 = \sigma_8(\Omega_{\rm m}/0.3)^{0.5}$. I note first of all that I do not obtain a clockwise rotation of the clipped contours relative to the unclipped, as predicted by S16. In answer to this, I remind the reader that this prediction was for a Euclid-like 5000 deg^2 survey (see Section 6.2 for further discussion), whereas my constraints correspond to a 360 deg^2 survey. It is possible that a rotation becomes evident given smaller error bars. If I were to scale the covariance on the data, $C_{\rm data}$, so as to correspond to a survey of Euclid-like proportions, the covariance on the clipped predictions from DH10, $C_{\rm model}$, becomes the dominant source of uncertainty. This prevents me from making a meaningful prediction for the cosmological constraints for a survey of this size. In the future, larger simulation suites will facilitate interesting predictions for larger size surveys.

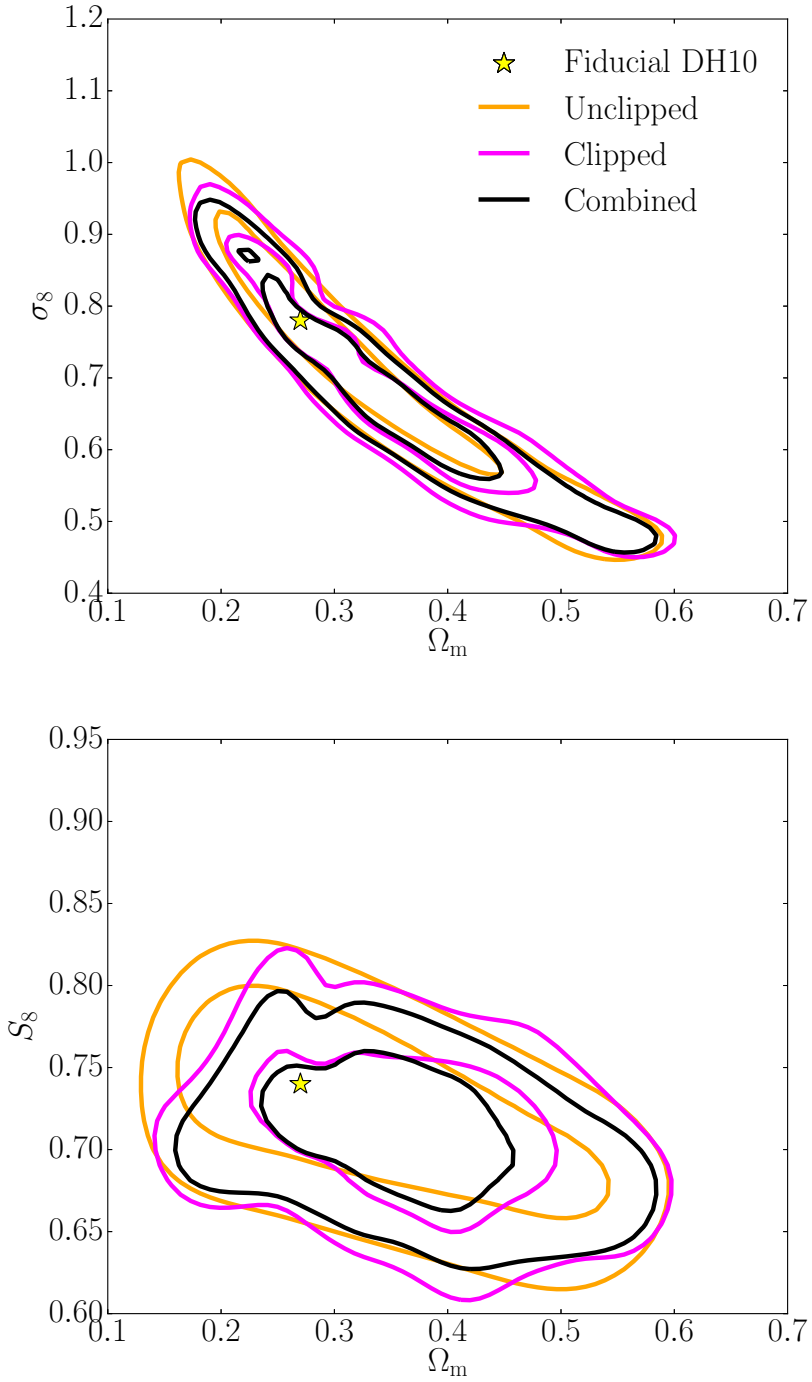

Figure 3.10 *The unclipped (orange), clipped (magenta) and combined (black) 68% and 95% confidence intervals for the fiducial cosmology from DH10 (shown by the yellow star) in the Ω_m-σ_8 and Ω_m-S_8 parameter spaces. I use only a subset of the fiducial cosmology simulations for the data vector in this test, corresponding to a KiDS-450-like survey of 360 deg^2. The unclipped contours are smooth as their theoretical expectation value is noise-free. In contrast the clipped likelihood is interpolated across sparse measurements from DH10. The resulting clipped and combined contours are therefore noisy in comparison to the unclipped constraints.*

Table 3.2 *The marginalised means and 68% confidence intervals on* $S_8 = \sigma_8(\Omega_m/0.3)^{0.5}$ *for a subset of independent DH10 simulations with the fiducial cosmology spanning 360 deg*^{2}*. The improvements in the constraint over the unclipped are presented in bold to the nearest percentage.*

Input $S_8 = 0.740$	
Unclipped	0.725 ± 0.042
Clipped	0.710 ± 0.037 **(11%)**
Combined	0.710 ± 0.033 **(22%)**

The combined constraints shown in Figure 3.10 recover the input cosmology, offering a significant improvement on the unclipped constraints. For example, the combined 95% confidence intervals are 18% and 29% smaller in area than those of the unclipped, in the Ω_m-σ_8 and Ω_m-S_8 parameter spaces, respectively. In comparison, the clipped contours are of comparable size to the unclipped in either parameter space.

Table 3.2 displays the marginalised mean and 68% confidence intervals on S_8 from the clipped, unclipped and combined contours in the Ω_m-S_8 plane. The improvement in the size of the confidence intervals offered by the combined analysis relative to the unclipped is 22%. This improvement, which is not changed considerably by the corrections for the finite box effect and cosmological bias in DH10, is indicative of the independent information in the clipped and unclipped statistics. Indeed, this is evidenced by the cross-correlation coefficient matrices presented in Appendix 3.A.

The clipped analysis alone yields S_8 constraints which are 11% tighter than the unclipped. For the clipped analysis to outperform the unclipped, the loss of information in clipping must be outweighed, by either the gain of phase information on the peaks, or the improvement in the clipped statistic for probing the more linear, clipped field. In this test however, I find that the success of the clipped analysis relative to the unclipped depends on the details of my interpolation scheme (see Appendix 3.B.2). The combined analysis consistently outperforms the unclipped in constraining the cosmology of the DH10 data set however, with all interpolations considered.

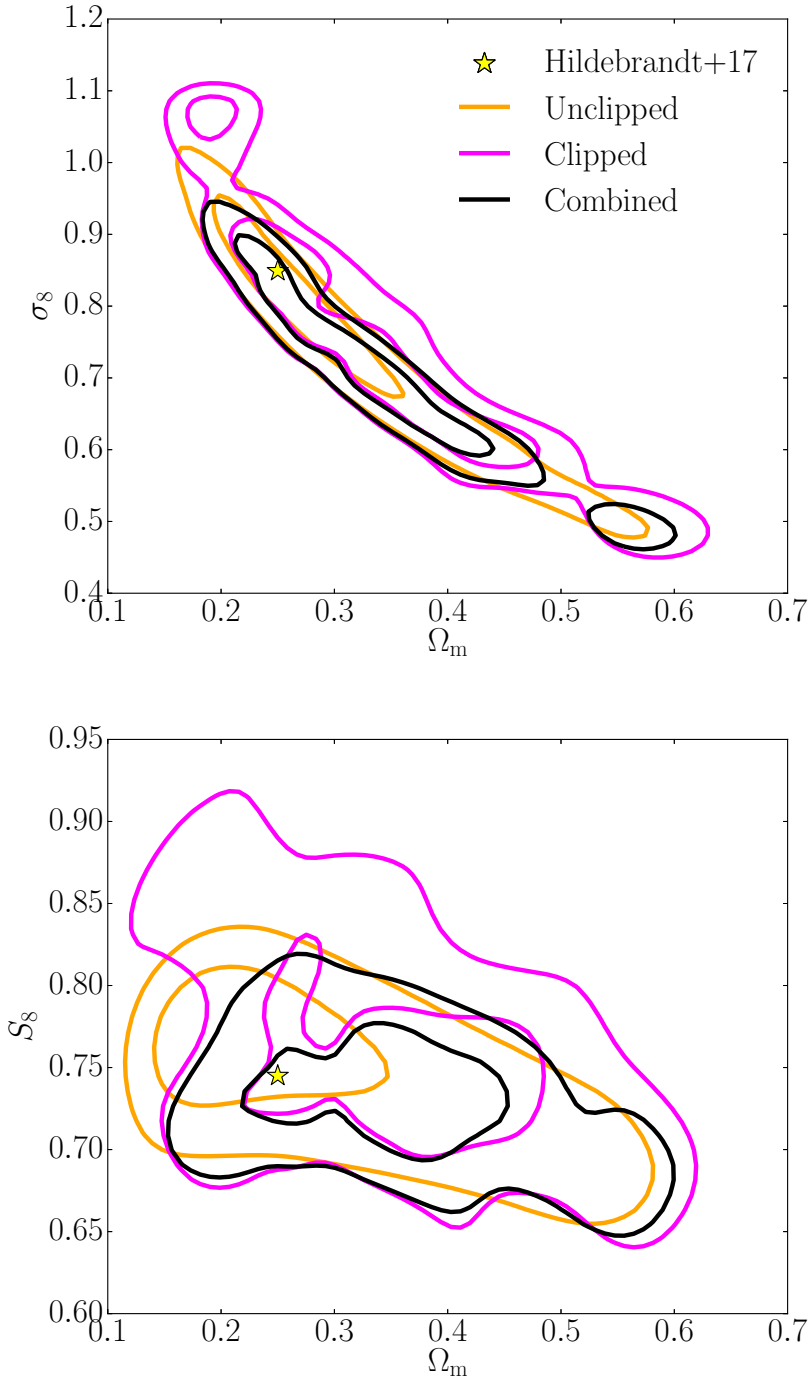

Figure 3.11 *The unclipped (orange), clipped (magenta) and combined (black) 68% and 95% confidence intervals for the KiDS-450 data in the Ω_m-σ_8 and Ω_m-S_8 parameter spaces. The yellow star depicts the best-fit cosmological parameters from the H17 cosmic shear analysis. The unclipped contours are smooth as their theoretical expectation value is noise-free. In contrast the clipped likelihood is interpolated across sparse measurements from DH10. The resulting clipped and combined contours are therefore noisy in comparison to the unclipped constraints.*

KiDS-450 constraints

After verifying that the combined clipped-and-unclipped analyses improve cosmological parameter constraints with a mock data set, I proceed to constrain the cosmology of the KiDS-450 data. Figure 3.11 displays the 68% and 95% confidence regions in the $\Omega_{\rm m}$-σ_8 and $\Omega_{\rm m}$-S_8 parameter spaces for this data set. The best-fit cosmology from the H17 cosmic shear analysis is designated by the yellow star. Once again I have applied the finite box and cosmological bias calibration corrections to the clipped predictions from DH10. I have also interpolated from the DH10 cosmologies with radial basis functions, and applied the same degree of smoothing as in Figure 3.10. The slight discontinuities in the tails of the clipped and combined contours in the $\Omega_{\rm m}$-σ_8 space, seen also by Martinet et al. (2018) in their analysis involving the DH10 mocks, are a product of the sparsity of the simulated cosmologies, and disappear if I apply a greater degree of smoothing.

In both the $\Omega_{\rm m}$-σ_8 and $\Omega_{\rm m}$-S_8 parameter spaces, we see that the clipped and combined contours are consistent with the best fit cosmological parameters from H17, despite the differences in the analyses. In addition to accounting for more systematics, the H17 result was obtained using four tomographic bins in the 0.1–0.9 photometric redshift range, as opposed to the single 0.5–0.9 bin. H17 also used both the $\xi_+^{\rm unclip}$ and $\xi_-^{\rm unclip}$, but omitted the largest two and smallest three θ bins for these statistics, respectively. In the $\Omega_{\rm m}$-σ_8 and $\Omega_{\rm m}$-S_8 parameter spaces shown in Figure 3.11, the 95% confidence intervals from the combined analysis are about 13% and 10% smaller than the unclipped, respectively, whereas those of the clipped analysis are considerably larger. There are a number of extra sources of noise when working with the KiDS-450 data, which could cause the clipped contours to inflate relative to the unclipped, in contrast to what was observed when working with the DH10 data vector. These include galaxy shape measurement, baryonic effects and $n(z)$ uncertainties; this is discussed further in Appendix 3.B.2.

The marginalised constraints on S_8 from the $\Omega_{\rm m}$-S_8 plane are shown in Table 3.3; again bold percentages detail improvements in the confidence intervals relative to the unclipped. As we saw with the DH10 data vector in earlier in this Section, the combined analysis offers improvements on the unclipped constraint, by 17%. This is comparable to the $\sim$20% improvement in S_8 found by Martinet et al. (2018) when constraining the KiDS-450 cosmology with combined peak statistics and standard shear correlation functions.

Table 3.3 *The marginalised means and 68% confidence intervals on $S_8 = \sigma_8(\Omega_\mathrm{m}/0.3)^{0.5}$ for KiDS-450. The improvement in the constraint over the unclipped are presented in bold to the nearest percentage. I remind the reader that the results of this work are not directly comparable to the H17 result, owing to the differences in the analyses discussed in the text.*

H17 $S_8 = 0.745^{+0.038}_{-0.038}$	
Unclipped	0.754 ± 0.036
Clipped	0.760 ± 0.051
Combined	0.734 ± 0.030 **(17%)**

3.5 Conclusions

In this Chapter I have performed a proof-of-concept analysis demonstrating that clipping transformations, which suppress the contribution from overdense regions to the weak lensing signal, can be used alongside a conventional "unclipped" cosmic shear analysis to improve cosmological parameter constraints. My pipeline reconstructs the projected surface mass density, performs clipping, determines the shear corresponding to the overdensities, and obtains "clipped" shear correlation functions. I have experimented with the threshold controlling the severity of the clipping transformation, and the smoothing employed in mass reconstruction, and found values well suited to the KiDS-450 data set.

There is currently no analytical prediction for clipped statistics as a function of cosmology, and so I calibrate the clipped shear correlation functions with numerical simulations spanning a broad range of Ω_m and σ_8. Consequently, I show that the combined clipped-and-unclipped analysis facilitates tighter constraints on $S_8 = \sigma_8(\Omega_\mathrm{m}/0.3)^{0.5}$, at fixed values of $\Omega_\mathrm{b}, n_\mathrm{s}, h$ and w_0, than the conventional unclipped analysis alone. For a mock data set with known cosmology, I find that the 68% confidence interval on S_8 is improved upon the unclipped by 22%. In the case of the KiDS-450 data, the improvement is 17%. The combined constraints from clipping could improve further given optimisation for the clipping threshold and mass reconstruction smoothing scale, though I leave this for future work on account of the computational cost.

The DH10 simulations with the calibration corrections are sufficiently accurate for modelling in this work. However, the limitations of the mocks that I have examined here do impact the improvement reported for clipping and are

likely to affect peak statistic studies also, reinforcing my conclusion that the success of these new statistics is intimately linked with the future accuracy and abundance of cosmological simulations. With new suites of simulations, the level of improvement seen in the joint analysis will increase in the future, as it will no longer be limited by the ∼7–100% uncertainty that I currently include with the DH10 predictions (this is investigated in the next Chapter). I note that a joint analysis of peak counts, cosmic shear and clipping both peaks and voids, also poses an interesting topic for further investigation.

My best-fit $S_8 = 0.734 \pm 0.030$ for the KiDS-450 data, inferred from a single photometric redshift bin in the range 0.5–0.9, is in good agreement with the tomographic cosmic shear analysis of H17, who found $S_8 = 0.745 \pm 0.038$. I note that H17 marginalised over Ω_b, n_s, and h whereas my constraints are made at fixed values of these parameters. In the future, larger suites of numerical simulations will permit investigation of how clipped statistics vary with these cosmological parameters. H17 also marginalise over photometric redshift uncertainties, the effects of baryons and intrinsic alignments, which I have not contended with here. In order for clipping to become a standard tool for constraining cosmology, work must be done to fold these extra systematic uncertainties into the clipped analysis. Finally, mass reconstruction methods with a more sophisticated handling of the masks are needed to reduce the bias imposed by this essential part of the clipping pipeline. Nevertheless, the results obtained in this work robustly demonstrate that clipping improves constraining power and should be explored in future cosmic shear analyses.

3.A SLICS covariance matrices

My likelihood analysis for cosmological parameters necessitates auto-covariance matrices for the clipped and unclipped statistics, as well as the cross-covariance between the two. Non-Gaussianity in cosmological density fields engenders correlations between the different angular scales probed by these measurements which are not well described by theory. Therefore I use the SLICS numerical simulations to model these covariance matrices. From the SLICS covariance matrices, defined in equation 3.16, I calculate correlation coefficient matrices,

defined as

$$\mathcal{CC}_{\pm}(\theta_i, \theta_j) = \frac{C_{\pm}(\theta_i, \theta_j)}{\sqrt{C_{\pm}(\theta_i, \theta_i) \times C_{\pm}(\theta_j, \theta_j)}}\,, \tag{3.21}$$

where $C_{\pm}(\theta_i, \theta_j)$ represents either the auto-covariance matrices for the $\xi_{\pm}^{\rm unclip}$ or $\xi_{\pm}^{\rm clip}$ statistics, or the cross-covariance matrix between the clipped and unclipped statistics. In the correlation coefficient matrix, the covariance is normalised to a value of unity for the strongest positive correlations on the leading diagonal, and values between -1 and 1 for all other elements.

In Figure 3.12, I display correlation coefficient matrices for the clipped and unclipped ξ_+ in the upper panel, and for the ξ_- in the lower panel. Each of these matrices are built out of the following components. The auto-correlation coefficient matrix for the unclipped statistic is in the lower-left corner, the clipped is in the upper-right corner, the matrix describing the cross-correlation coefficients between these two statistics is in the upper-left, and its transpose is in the lower-right.

The fact that many of the off-diagonal elements of these matrices are non-zero (varying in the range -0.1 to 0.8 in either panel), indicates the need for simulations in order to model the correlations not only across angular scales, but also the correlations between the clipped and unclipped statistics. The cross-correlation matrices reveal that the clipped and unclipped statistics are not perfectly correlated and thus contain some independent information. It is also interesting to note that the clipped auto-correlation matrices seem to feature slightly weaker correlations between scales around $\sim$10 arcmin and $\sim$100 arcmin in the upper and lower panels respectively, than the unclipped auto-correlation matrices. This is consistent with the clipped field being more Gaussian than the unclipped. I note that the correlation between the clipped and unclipped measurements does not tend to unity on the largest scales probed in this analysis. This is a reflection of that fact that the largest-scale clipped and unclipped measurements for my fiducial analysis also do not converge. I find that for a less aggressive clipping threshold, (see the upper panel in Figure 3.4), both the cross-correlation coefficients and the ratio between the clipped and unclipped measurements do however converge to unity as expected. For my fiducial set-up I would expect perfect correlation between the clipped and unclipped signals to occur on scales that are larger than I can currently test with the SLICS or

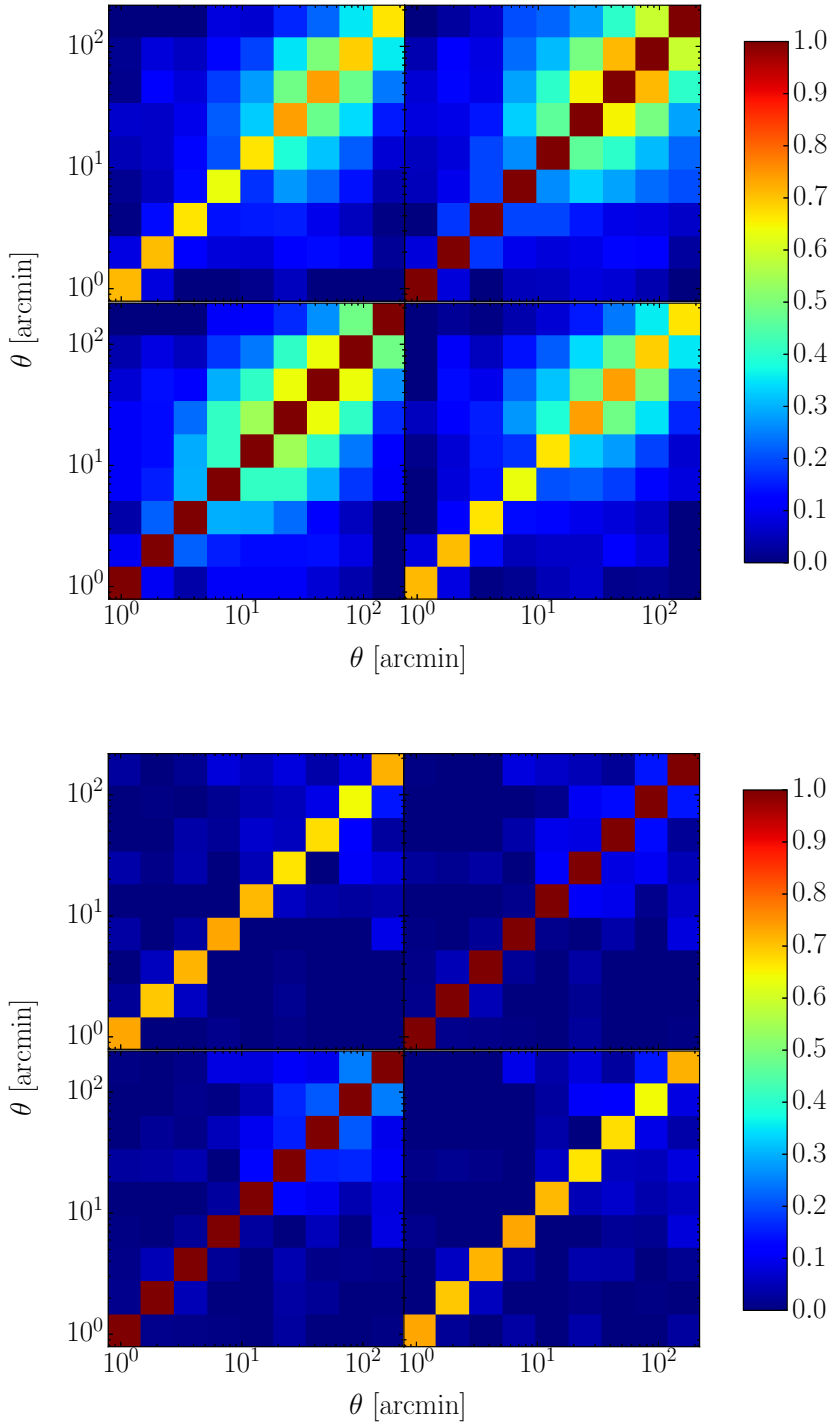

Figure 3.12 *The correlation coefficient matrices measured from SLICS (featuring shape noise typical of KiDS-450) for the clipped and unclipped ξ_+ (upper panel) and ξ_- (lower panel). Each panel consists of the following components. Lower left: the auto-correlations for the $\xi_\pm^{\rm unclip}$. Upper right: the auto-correlations for the $\xi_\pm^{\rm clip}$. Upper-left (and lower-right): the cross-correlations between $\xi_\pm^{\rm unclip}$ and $\xi_\pm^{\rm clip}$ (and its transpose).*

KiDS-450 survey area.

3.B Cosmological constraints

3.B.1 Sensitivity to the unclipped predictions

In Section 3.4.2, I use the theoretical $\xi_+^{\rm unclip}$ from equation 3.13 to constrain the cosmology of the subset of DH10 simulations with the fiducial cosmology spanning 360 deg^2. I could alternatively have used the unclipped predictions from the simulations themselves, though these predictions are subject to the finite box effect, cosmological bias and additional uncertainty, as discussed in Sections 3.3.4 and 3.3.5 respectively. The noise-free theoretical predictions (e.g. from NICAEA) are a more suitable choice for constraining cosmology where such predictions are available (which is of course not so, in the case of the clipped statistic). Nevertheless, I verify that one still obtains improved cosmological constraints in the combined analysis irrespective of whether I employ the theoretical or simulated $\xi_+^{\rm unclip}$.

Figure 3.13 compares the marginalised means and 68% confidence intervals on S_8 from the $\Omega_{\rm m}$-S_8 parameter space when I use the unclipped predictions from equation 3.13 and from DH10. These constraints are clearly consistent with one another and the input S_8, but differ in their details, as is shown in Table 3.4. The theoretical unclipped better recovers the input S_8 indicating again that they should be used over DH10 whenever possible. One consequence of this choice however, is that I find it leads to a $\sim 0.4\sigma$ difference in the mean marginalised constraints on S_8 when comparing the clipped and unclipped analyses in Table 3.4. Given the high correlation between these two statistics, shown in Figure 3.12, I would expect better agreement, which I indeed find when using the DH10 measurements for both the clipped and unclipped predictions. In this case the mean S_8 agree to within 0.05σ. When using DH10 for both the clipped and unclipped predictions, my finding that the combined clipped-and-unclipped analyses improves cosmological parameter constraints holds but in this case the level of improved constraining power decreases to 12%. I find these conclusions are robust to different realisations of the DH10 data vector.

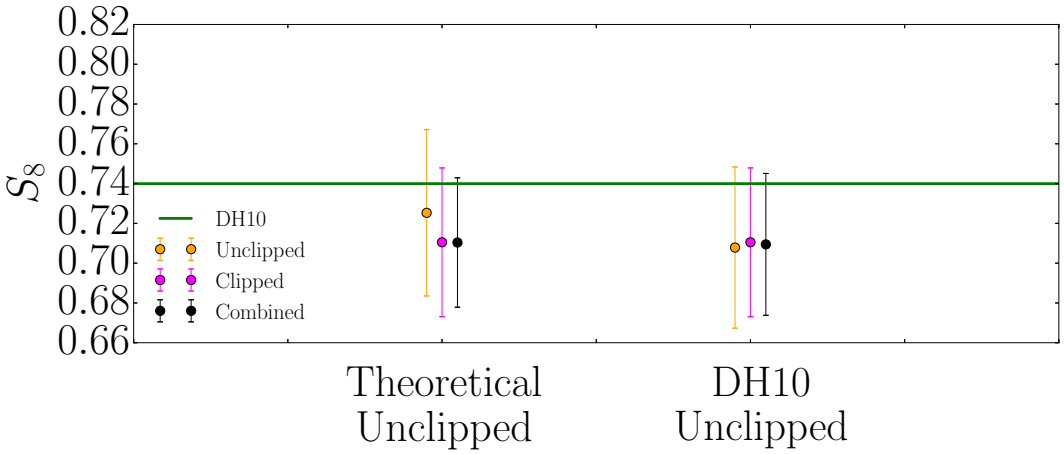

Figure 3.13 *The marginalised means and 68% confidence intervals on S_8 from the Ω_m-S_8 plane for the DH10 fiducial cosmology data vector, depending on whether the $\xi_+^{\rm unclip}$ derive from equation 3.13 or from the DH10 mocks themselves. The input S_8 is designated by the horizontal green line. The corrections for the finite box size and cosmological bias have been applied to the predictions from DH10.*

Table 3.4 *The marginalised means and 68% confidence intervals on the DH10 data vector from Figure 3.13 expressed in tabular form. Improvements over the unclipped confidence intervals are detailed in bold to the nearest percentage. The corrections for the finite box size and cosmological bias have been applied to the predictions from DH10.*

	Theoretical Unclipped	DH10 Unclipped
Input $S_8 = 0.740$		
Unclipped	0.725 ± 0.042	0.708 ± 0.041
Clipped	0.710 ± 0.037 **(11%)**	0.710 ± 0.037 **(8%)**
Combined	0.710 ± 0.033 **(22%)**	0.709 ± 0.036 **(12%)**

3.B.2 Sensitivity to the interpolation method

Qualitatively my finding that the combined clipped-and-unclipped analyses improves cosmological parameter constraints holds irrespective of how I choose to interpolate from the DH10 cosmologies onto the $\Omega_{\rm m}$-σ_8 and $\Omega_{\rm m}$-S_8 grids. Quantitatively however there is a dependence of the marginalised constraints on these choices, particularly for the highly degenerate $\Omega_{\rm m}$ and σ_8 parameters. This is to be expected given the level of noise in the predictions and the sparsity with which the predictions are sampled across the parameter space. I find that the measurement of S_8 is the least sensitive to the interpolation scheme adopted, motivating the use of this statistic to highlight the benefit of clipping throughout this Chapter. In this Appendix I compare my marginalised S_8 constraints for KiDS-450 and the DH10 mock data for four different interpolation methods.

DH10 constraints

The first method I consider for interpolating the likelihoods from the DH10 simulations, is the interpolation with radial basis functions (RBFs), smoothing the contours as described in Section 3.4. Secondly, I try the RBF interpolation with no contour smoothing. Thirdly, I consider simple 2D linear interpolation. I also obtain results from interpolating the clipped and unclipped DH10 ξ_+ statistics, for each θ bin individually, rather than the likelihoods, onto the $\Omega_{\rm m}$-S_8 plane. I use the smoothed-RBF method when interpolating the correlation functions in this comparison. By comparing the theoretical $\xi_+^{\rm unclip}$ with those extrapolated outside of the range of the DH10 cosmologies, I find that the extrapolation of the correlation functions is inaccurate. Thus I impose a prior which sets the likelihoods calculated from the extrapolated clipped and unclipped ξ_+ to zero. Since I find good agreement between theory and the mocks when I extrapolate the unclipped likelihoods instead of the unclipped correlation functions, I do not impose this prior when performing the likelihood-interpolations. Indeed, I find it does not change my results significantly when it is imposed.

Figure 3.14 and Table 3.5 present a comparison of the marginalised means and 68% confidence intervals on S_8 for the DH10 data set. Featured, are the three likelihood-interpolations methods and one ξ_+-interpolation method. Clearly all of the marginalised constraints from the different ways of interpolating are consistent

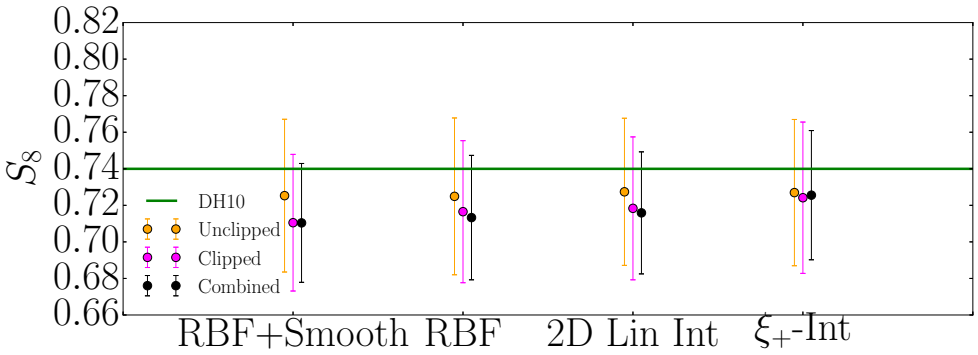

Figure 3.14 *The marginalised means and 68% confidence intervals on S_8 from the Ω_m-S_8 plane for the DH10 fiducial cosmology data vector, via different intepolation methods listed on the horizontal axis. From the left-hand side, the first three methods are likelihood-interpolations. "RBF+Smooth" refers to the likelihood-interpolation with radial basis functions and contour smoothing. "RBF" refers to this interpolation with no smoothing, and "2D Lin Int" designates simple 2D linear interpolation. "ξ_+-Int" refers to interpolating the clipped and unclipped shear correlation functions, instead of the likelihoods, again with the smoothed-RBF method. The input S_8 is designated by the horizontal green line. The corrections for the finite box size and cosmological bias have been applied to the $\xi_+^{\rm clip}$ predictions from DH10. The $\xi_+^{\rm unclip}$ predictions come from equation 3.13 and are calculated using* NICAEA.

Table 3.5 *The marginalised means and 68% confidence intervals on S_8 for the DH10 data vector from Figure 3.14 expressed in tabular form. Improvements over the unclipped confidence intervals are detailed to the nearest percentage in bold.*

	RBF+Smooth	RBF	2D Lin Int	ξ_+-Int
Input $S_8 = 0.740$				
Unclipped	0.725 ± 0.042	0.725 ± 0.043	0.727 ± 0.040	0.727 ± 0.040
Clipped	0.710 ± 0.037 **(11%)**	0.717 ± 0.039 **(9%)**	0.718 ± 0.039 **(3%)**	0.724 ± 0.041
Combined	0.710 ± 0.033 **(22%)**	0.713 ± 0.034 **(21%)**	0.716 ± 0.033 **(17%)**	0.726 ± 0.035 **(12%)**

with one another, and with the true cosmological parameters to $< 1\sigma$. We see that the combined analysis invariably is an improvement upon the unclipped, with 68% confidence intervals that are between 12% and 22% tighter. The combined analysis also yields improvements on $\Omega_{\rm m}$ and σ_8, of 28% and 24% respectively, in the standard analysis with the DH10 data set presented in Section 3.4. Though the greater sensitivity of these results to the interpolation scheme means that I ascribe more confidence in my measurement of S_8.

KiDS-450 constraints

The marginalised constraints on the KiDS-450 data fluctuate more than those on the DH10 data vector across the different interpolation schemes. This is to be expected given that KiDS-450 features extra sources of noise, such as galaxy shape measurement, baryonic effects and $n(z)$ uncertainties which have not been accounted for in this proof-of-concept analysis. These may engender spurious peaks in the interpolated likelihoods which bias some interpolation methods more than others. What is more, the nuisance cosmological parameters $\Omega_{\rm b}, n_{\rm s}$ and h are almost certainly mismatched between the data and the predictions. In principle this could affect the $\xi_+^{\rm clip}$ differently than the $\xi_+^{\rm unclip}$ predictions.

I find that the improvements over the unclipped found in the combined marginalised S_8 constraints, displayed visually in Figure 3.15 and numerically in Table 3.15, are consistent for the interpolation schemes which incorporate smoothing, "RBF+Smooth" and "ξ_+-Int", between 14% and 17%. The interpolation schemes without smoothing however, "RBF" and "2D Lin Int", yield little to no improvement in the combined constraints. This is because the interpolated clipped and combined likelihoods for the KiDS-450 data set are reasonably noisy, and the methods without smoothing are more strongly affected by this. The smoothing reduces the impact of spurious noise spikes in the likelihoods biasing the parameter constraints. Thus I regard the constraints obtained with these interpolations as more accurate, and maintain that the improvement found by combining the clipped and unclipped analyses is around the 17% level for the KiDS-450 data.

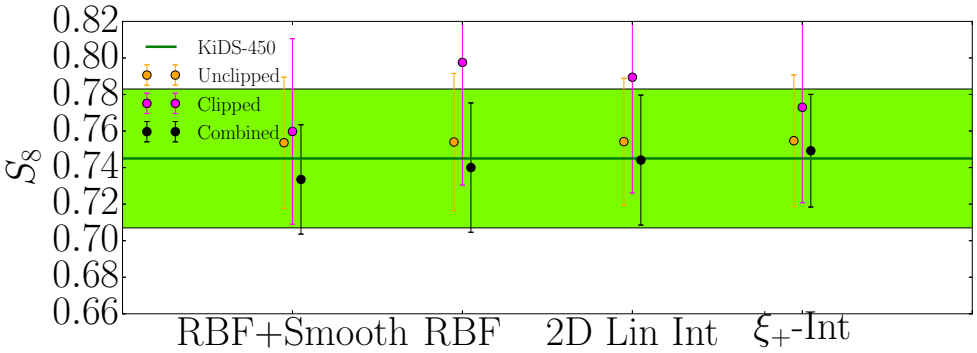

Figure 3.15 *The same as Figure 3.14 but for the KiDS-450 data. The light-green region corresponds to the 68% confidence region from the H17 cosmic shear analysis.*

Table 3.6 *The same as Table 3.5 but for the KiDS-450 data. Improvements over the unclipped confidence intervals are detailed to the nearest percentage in bold. I remind the reader that the results of this work are not directly comparable to the H17 result, owing to the differences in the analyses discussed in the text.*

	RBF+Smooth	RBF	2D Lin Int	ξ_+-Int
H17 $S_8 = 0.745^{+0.038}_{-0.038}$				
Unclipped	0.754 ± 0.036	0.754 ± 0.038	0.754 ± 0.035	0.755 ± 0.036
Clipped	0.760 ± 0.051	0.798 ± 0.067	0.789 ± 0.063	0.773 ± 0.052
Combined	0.734 ± 0.030 **(17%)**	0.740 ± 0.035 **(6%)**	0.744 ± 0.036	0.749 ± 0.031 **(14%)**

3.C KiDS-450 mass maps

In Figures 3.16 and 3.17 I present convergence maps for the North and South KiDS-450 patches respectively. In producing these maps, I follow the mass reconstruction methodology of Kaiser & Squires (1993) as detailed in Section 3.3.1. The maps are smoothed with a Gaussian filter with width $\sigma_{\rm s} = 6.6$ arcmin, and the regions exceeding the clipping threschold $\kappa^{\rm c} = 0.010$ are highlighted with the green contours. I follow Van Waerbeke et al. (2013) and set the convergence to zero in regions where more than 50% of the Gaussian smoothing window is centred on masked pixels.

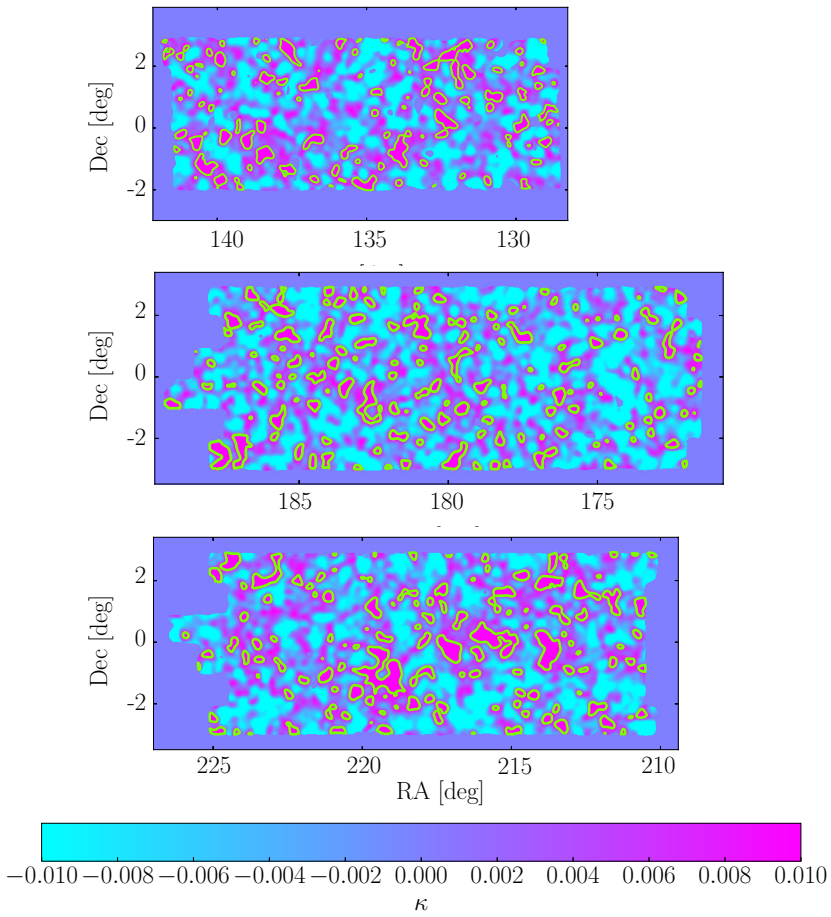

Figure 3.16 *Maps of the convergence, κ, for the three KiDS-450 North patches, G9 (upper), G12 (middle) and G15 (lower). The maps have been smoothed with a Gaussian filter with width $\sigma_s = 6.6$ arcmin. Unobserved/masked regions are given zero convergence, as is described in the text. The regions highlighted by the green contours, exceed the clipping threshold, $\kappa^c = 0.010$, and are therefore clipped in my pipeline. The clipped regions make up $12\pm1\%$ of the effective area of the five KiDS-450 patches.*

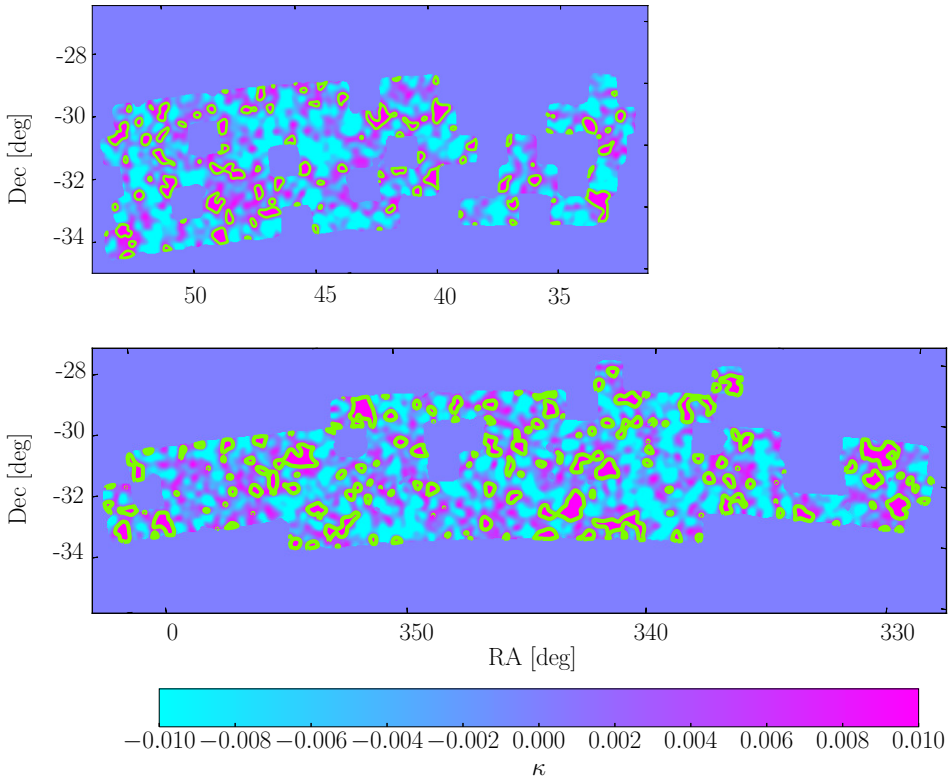

Figure 3.17 *The same as Figure 3.16 but for the two KiDS-450 South patches, GS (upper), and G23 (lower).*

4

Emulating Different Universes

In the previous Chapter, I found that novel "clipping" transformations enhance constraints within a cosmic shear analysis, but the absence of an analytical prescription means that this technique is critically dependent on numerical simulations. Moreover, the simulations used for calibration (Dietrich & Hartlap, 2010, DH10) varied only in two parameters, $\Omega_{\rm m}$ and σ_8. Future applications of clipping, and indeed all cosmological statistics lacking an analytical prediction, necessitate mocks which explore the full parameter space to which lensing is sensitive. Increasing the dimensionality in turn calls for a more sophisticated and robust method of interpolating a given measurement from the calibration simulations, over the simple 2D radial basis function interpolation discussed in Section 3.4.1 and Appendix 3.B.2.

In this Chapter, I develop methodology in answer to these requirements, by tailoring a *Gaussian process regression emulator* to train on the *cosmo*-SLICS numerical simulation suite. In Section 4.1 I outline the theoretical framework underlying the emulator methodology. In Section 4.2 I introduce the *cosmo*-SLICS, and explain why this suite is an excellent resource for exploring the cosmological dependence of novel cosmic shear statistics. In Section 4.3 I present

my methodology for training and testing the accuracy of the *cosmo*-SLICS emulator. I apply clipping to these new simulations in Section 4.4, and test how the increased precision in the clipped shear signal, facilitated by the *cosmo*-SLICS emulator, benefits inference of cosmological parameters. Finally, I present my conclusions regarding the future applications of the *cosmo*-SLICS suite and emulator, in Section 4.5. Section 4.3 of this Chapter comprises a significant part of the paper introducing these simulations, Harnois-Deraps, Giblin & Joachimi (2019), currently under review.

4.1 Gaussian process regression theory

We are tasked with approximating unknown relationships between variable input and measurable output frequently in our lives. For example, the dependence of hangover severity (output) on the volume and strength of alcohol consumed the previous evening, along with additional negating input parameters such as the amount of food eaten and sleep obtained, is a confounding inverse-problem familiar to many. It is unlikely that the conditions of any previous night-out are ever recreated perfectly. That is to say, one probably will never experience the *exact* same combination of input variables (alcohol, food, sleep etc.) on two given occasions. Therefore we are tasked with using our prior experiences and observations, to *train* our intuition to predict how much suffering tomorrow morning will bring, based on previously untried combinations of the input variables.

When computers are tasked with using empirical data to constrain the mapping relation between input and output, thereby facilitating predictions of the output at previously unseen input values, we enter the realm of *emulators* employing supervised machine learning. It is prudent to distinguish between cases of *regression*, in which the output are continuous quantities, and those of *classification*, in which the output are discrete, due to differences in the formalism used to tackle these two types of problems. Astronomical examples of these include predicting non-linear power spectra or clipped correlation functions (see Chapter 3) as a function of input cosmological parameters (regression), and star/galaxy separation based on photometric data (classification). Indeed our hangover example could in theory be cast as either, depending on if the output is duration of headache experienced the next day (a continuous quantity; regression)

or if the output is whether one feels too unwell to go for a morning cycle ("yes" or "no", discrete quantities; classification).

In this Chapter, I implement machine learning in *regression* problems to demystify the dependence of cosmological statistics on input parameters. Specifically I develop *Gaussian process* (GP) emulators which interpolate the statistics sampled at an ensemble of input coordinates (the training set) to new values of the input (the trial coordinates). GP regression describes the distribution of all functions which can possibly map between an input coordinate, $\boldsymbol{x} = \{x_1, \cdots x_d\}$ where d is the dimensionality of the input, and output observable, y, via

$$y = f(\boldsymbol{x}) + \epsilon\,, \tag{4.1}$$

where ϵ is an additive noise term preventing perfect reconstruction of the output by the function, $f(\boldsymbol{x})$. The noise is sampled from a mean-zero Gaussian distribution with variance, σ^2, determined by the uncertainty on the measurement of y,

$$\epsilon \sim \mathcal{N}(0, \sigma^2). \tag{4.2}$$

GP regression earns the prefix "Bayesian" on account of the fact that the n functions, $\boldsymbol{f} \equiv \boldsymbol{f}(X)$, relating the output $\boldsymbol{y}$ with the input coordinates contained in the $n \times d$ *design matrix* of the form,

$$X = \begin{bmatrix} \boldsymbol{x_1} \\ \vdots \\ \boldsymbol{x_n} \end{bmatrix} = \begin{bmatrix} x_{11} & \cdots & x_{1d} \\ \vdots & & \vdots \\ x_{n1} & ... & x_{nd} \end{bmatrix}, \tag{4.3}$$

are sampled from a posterior probability distribution,

$$p(\boldsymbol{f}|\boldsymbol{y}, X) = \frac{p(\boldsymbol{y}|\boldsymbol{f}, X)p(\boldsymbol{f})}{p(\boldsymbol{y}|X)}. \tag{4.4}$$

Here $p(\boldsymbol{y}|\boldsymbol{f}, X)$ denotes the likelihood, or the probability density of the observations over the functions and $p(\boldsymbol{f})$ denotes the prior distribution of functions deemed to reasonably map between the input and output before any data is

observed. The denominator on the right-hand side is the evidence, or marginal likelihood, given by

$$p(\boldsymbol{y}|X) = \int p(\boldsymbol{y}|\boldsymbol{f}, X)p(\boldsymbol{f})d\boldsymbol{f}\,, \tag{4.5}$$

which ensures that the posterior integrated over all possible functions is unity.

First let us consider the construction of the prior, a mean-zero Gaussian distribution, the behaviour of which is governed by its covariance function, or 'kernel', which defines the covariance between the function values at coordinates $\boldsymbol{x}$ and $\boldsymbol{x'}$,

$$k(\boldsymbol{x}, \boldsymbol{x'}|\boldsymbol{h}) \equiv \mathrm{cov}\left(f(\boldsymbol{x}), f(\boldsymbol{x'})|\boldsymbol{h}\right)\,, \tag{4.6}$$

where $\boldsymbol{h}$ is a vector of free hyperparameters. One implements prior belief on the behaviours of functions which could map between coordinate $\boldsymbol{x}$ and output y via the choice of the functional form for the kernel, of which many are possible, and the values of said kernel's hyperparameters. A common choice of kernel (Rasmussen & Williams, 2006; Habib et al., 2007; Heitmann et al., 2009) is the squared exponential[1] given by

$$k(\boldsymbol{x}, \boldsymbol{x'}|\boldsymbol{h}) = A\prod_{l=1}^{d}\exp\left[\frac{(x_l - x'_l)^2}{p_l^2}\right]\,, \tag{4.7}$$

for which the hyperparameters are A, the positive amplitude determining the overall variance of the sampled functions, and $\boldsymbol{p} = \{p_1, \cdots, p_d\}$, which govern how quickly the function fluctuates in each of the d dimensions of the input coordinate, $\boldsymbol{x}$. This kernel has the following properties: (1) the covariance varies smoothly within the parameter space; (2) it depends only on the Euclidean distance between points, such that $k\left[f(\boldsymbol{x}), f(\boldsymbol{x'})|\boldsymbol{h}\right] = k\left[f(\boldsymbol{x'}), f(\boldsymbol{x})|\boldsymbol{h}\right]$; (3) predictions become maximally correlated when $\boldsymbol{x} = \boldsymbol{x'}$; and (4) the correlation is large for points in relative proximity and small for largely separated points.

Equation 4.7 describes the covariance between the function values at a pair of coordinates. Extending this mathematical formalism to n parameter values, I define the $n \times n$ covariance *matrix*, $K(X, X)$, the elements of which are populated

[1] Also known as the radial basis function or Gaussian kernel.

by the kernel values, $k(\boldsymbol{x}, \boldsymbol{x'}|\boldsymbol{h})$. In the absence of observed data, trial functions are sampled directly from the prior,

$$\boldsymbol{f^*} \sim \mathcal{N}\left(0, K(X^*, X^*)\right) , \tag{4.8}$$

where $\boldsymbol{f^*} \equiv \boldsymbol{f}(X^*)$ is an array of functions, denoted by the superscript '$*$' to distinguish from the training set, corresponding to a design matrix of trial coordinates X^*.

Each of the three left hand panels of Figure 4.1 depict five functions sampled directly from the prior with the squared-exponential kernel given by Equation 4.7, with dimensionality $d = 1$ and the length scale, p, varying from 0.1 (upper), to 1 (middle) to 10 (lower). The grey region illustrates plus and minus the standard deviation, determined by the fixed amplitude, $A = 1$, about the mean of an infinite number of sampled functions. We see that smaller values of p engender rapidly changing functions whereas larger values produce functions which vary more slowly and smoothly.

Now let us now consider conditioning the GP on a data set, $\mathcal{D}$, consisting of n measurements of the observable at different values of the input parameters, $\mathcal{D} = \{X, \boldsymbol{y}\} \equiv \{(\boldsymbol{x}_i, y_i)|i = 1, ..., n\}$, such that all functions need be consistent with the mapping between $\boldsymbol{x}$ and $\boldsymbol{y}$ in this training set. Assuming all y_i feature the same variance, σ^2, the joint prior distribution of the training set output and the functions, $\boldsymbol{f^*}$, corresponding to n^* trial coordinates in the design matrix X^*, is then given by

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{f^*} \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(X, X) + \sigma^2 I & K(X, X^*) \\ K(X^*, X) & K(X^*, X^*) \end{bmatrix}\right) \tag{4.9}$$

where $K(X, X^*)$ denotes the $n \times n^*$ matrix of covariances between all pairs of training and trial coordinates, and I refers to the $n \times n$ identity matrix. To construct the posterior distribution over functions, the prior distribution must be restricted to only those which are statistically consistent with the training set. One could consider sampling multiple functions from the prior and rejecting those which disagree with the data, although this would prove computationally inefficient. Instead, the GP is conditioned on the training set via the product of

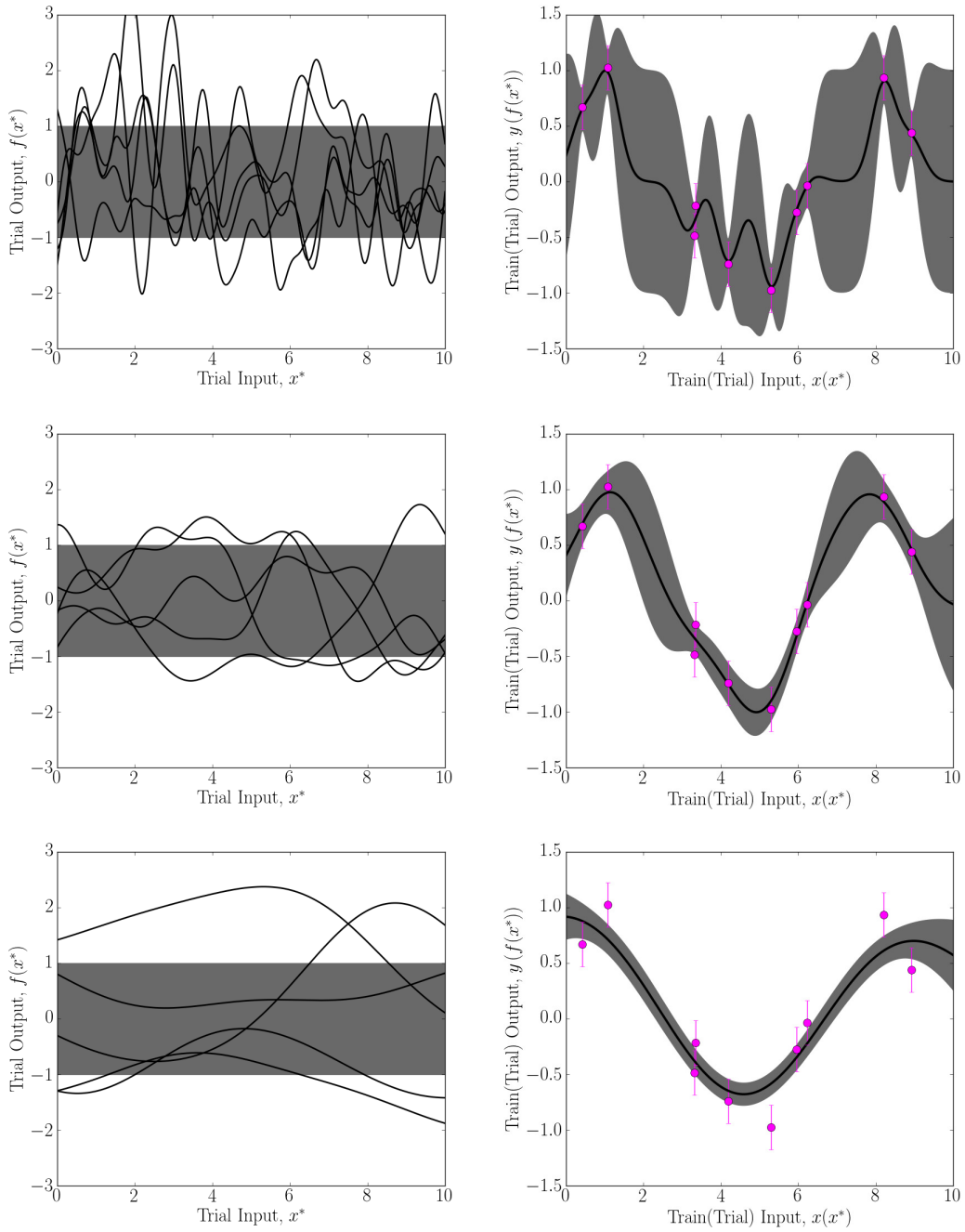

Figure 4.1 *A 1-dimensional GP example featuring the prior with the squared exponential kernel (see equation 4.7) with amplitude $A = 1$ and $p = 0.1$ (upper), 1 (middle) and 10 (lower). Left panels: 5 trial functions, f^*, sampled directly from the prior at 500 linearly spaced input coordinates, $x^* \in [0, 10]$. Right panels: the mean trial function (f^*; black line) sampled at 500 x^* points from the posterior conditioned on the noisy sinusoid training set ($y(x)$; magenta points). The grey region in all cases indicates plus and minus the standard deviations about the mean function.*

the prior with the likelihood given by

$$p(\boldsymbol{y}|\boldsymbol{f}, X) = \prod_{i=1}^{n} p(y_i|\boldsymbol{x}_i, f_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(\mathrm{y_i} - \mathrm{f_i})^2}{2\sigma_\mathrm{i}^2}\right). \tag{4.10}$$

If I once again assume all training predictions feature the same variance ($\sigma_i^2 \rightarrow \sigma^2$), this simplifies further, giving

$$p(\boldsymbol{y}|X, \boldsymbol{f}) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}|\boldsymbol{y} - \boldsymbol{f}|^2\right) = \mathcal{N}(\boldsymbol{f}, \sigma^2 I)\,, \tag{4.11}$$

where $|\boldsymbol{z}|$ denotes the Euclidean length of a vector $\boldsymbol{z}$. The product of the Gaussian likelihood and prior generate a Gaussian posterior, from which consistent trial functions can be sampled,

$$\begin{aligned} \boldsymbol{f^*}|X^*, X, \boldsymbol{f} \sim \mathcal{N}(&K(X^*, X)K(X, X)^{-1}\boldsymbol{f}, \\ &K(X^*, X^*) - K(X^*, X)K(X, X)^{-1}K(X, X^*))\,, \end{aligned} \tag{4.12}$$

where $\boldsymbol{f}$ are the function values sampled at the training input coordinates.

The black line in each of the right hand panels of Figure 4.1 show the average of many functions sampled from a 1-dimensional GP posterior distribution conditioned on a data set represented by the magenta points. The data in question is a sinusoid with Gaussian noise, sampled at 10 points, x, randomly sampled from a uniform distribution. Again, the grey region indicates plus and minus the standard deviation around the mean function.

As the length scale, p, increases logarithmically from 0.1 (upper) to 10 (lower), the contribution to $\boldsymbol{y}$ from the function signal, $\boldsymbol{f}$, relative to the noise, $\boldsymbol{\epsilon}$ (see equation 4.1), decreases. This can be seen by considering the two data points closely clustered at $x \simeq 3.5$. In the upper panel where $p = 0.1$, the sampled functions vary rapidly enough for the difference in the y values of these two points to be explained by a sharp up-turn in the mean signal, with a correspondingly small standard deviation. The large level of variation allowed by the sampled functions however means that the error bars inflate away from the training data. In contrast, in the middle and lower panels where $p = 1$ and 10 respectively, the two points at $x \simeq 3.5$ are explained by roughly the same function value but

with a larger level of noise than in the $p = 0.1$ case. As the sampled functions vary slowly however, the noise remains mostly uniform across the input range bracketed by the data.

4.1.1 Training the emulator

The previous example demonstrates the principle task of developing a useful GP emulator – training the GP to make accurate predictions for trial coordinates absent from the training set, means optimising for the values of the hyperparameters, and potentially also the functional form of the kernel. Figure 4.1 suggests an apparent simplicity to this task; for this 1D problem, one can see by eye that the more slowly varying functions, and hence larger values of the length scale p, are more appropriate for explaining the trends seen in the data, than the complex, rapidly fluctuating functions. However, this task becomes more complicated for multi-dimensional input.

Consider the posterior distribution over the functions expressed in equation 4.4. Expanding this to include the dependence on the functional form of the kernel, say $\mathcal{H}_i$, and associated hyperparameters, $\boldsymbol{h}$, gives

$$p(\boldsymbol{f}|\boldsymbol{y}, X, \boldsymbol{h}, \mathcal{H}_i) = \frac{p(\boldsymbol{y}|X, \boldsymbol{f}, \mathcal{H}_i)p(\boldsymbol{f}|\boldsymbol{h}, \mathcal{H}_i)}{p(\boldsymbol{y}|X, \boldsymbol{h}, \mathcal{H}_i)}\,, \tag{4.13}$$

and correspondingly, the evidence (marginal likelihood) takes the form,

$$p(\boldsymbol{y}|X, \boldsymbol{h}, \mathcal{H}_i) = \int p(\boldsymbol{y}|X, \boldsymbol{f}, \mathcal{H}_i)p(\boldsymbol{f}|\boldsymbol{h}, \mathcal{H}_i)d\boldsymbol{f}. \tag{4.14}$$

This posterior in fact forms the lowest level of a hierarchical Bayesian model, wherein inference is made at several tiers (see Alsing et al., 2016; Leistedt et al., 2016; Simpson et al., 2017, for recent examples in cosmology). At the level above this, we have the posterior distribution over the hyperparameters, given by

$$p(\boldsymbol{h}|\boldsymbol{y}, X, \mathcal{H}_i) = \frac{p(\boldsymbol{y}|X, \boldsymbol{h}, \mathcal{H}_i)p(\boldsymbol{h}|\mathcal{H}_i)}{p(\boldsymbol{y}|X, \mathcal{H}_i)}\,, \tag{4.15}$$

where we notice that the marginal likelihood from the first level (equation 4.14) plays the role of the likelihood at this level. $p(\boldsymbol{h}|\mathcal{H}_i)$ is known as the *hyper-prior*,

and the normalising constant at this tier is given by

$$p(\boldsymbol{y}|X, \mathcal{H}_i) = \int p(\boldsymbol{y}|X, \boldsymbol{h}, \mathcal{H}_i) p(\boldsymbol{h}|\mathcal{H}_i) d\boldsymbol{h}. \tag{4.16}$$

Finally, at the top tier we compute the posterior over the kernel model,

$$p(\mathcal{H}_i|\boldsymbol{y}, X) = \frac{p(\boldsymbol{y}|X, \mathcal{H}_i) p(\mathcal{H}_i)}{p(\boldsymbol{y}|X)}\,, \tag{4.17}$$

where again we see the marginal likelihood from the previous level appear in the place of the likelihood here, and $p(\boldsymbol{y}|X) = \sum_i p(\boldsymbol{y}|X, \mathcal{H}_i) p(\mathcal{H}_i)$. Typically one would expect the prior probability of any given kernel model to be flat.

We see that optimising for the choice of kernel and hyperparameter values apparently calls for the calculation of several integrals, which can be computationally expensive. In particular, the marginal likelihood in equation 4.16 is likely to be challenging for multi-dimensional input (Rasmussen & Williams, 2006, Chapter 5). One finds this is typically circumvented in the literature by choosing the kernel to be the squared-exponential (Rasmussen & Williams, 2006; Habib et al., 2007; Heitmann et al., 2009, see equation 4.7). Bayesian inference of the hyperparameters can then be conducted at the second level, by maximising the marginal likelihood in equation 4.14 with respect to the hyperparameters. This can then be made tractable by, for example, Markov Chain Monte Carlo (MCMC) algorithms, in which 'walkers' explore the parameter space by sampling the likelihood at trial coordinates, relocating there should it prove more probabilistically favourable than the current coordinate, or else doing so with a given probability. Recent applications in cosmology can be found in Troxel et al. (2017); Salvador et al. (2019); Prat et al. (2019); Joudaki et al. (2019), for example.

4.2 The *cosmo*-SLICS

The numerical simulation suite, *cosmo*-SLICS (Harnois-Deraps, Giblin & Joachimi, 2019), are an extension of the SLICS suite (Harnois-Déraps et al., 2018) used for covariance estimation in Chapter 3. As with its predecessor, *cosmo*-SLICS was created using the gravity solver CUBEP^{3}M (Harnois-Déraps et al., 2013), with many of the same specifications, such as the lightcone area, $A_{\rm lc} = 100\,{\rm deg}^2$, box size, $L_{\rm box} = 505\,{\rm Mpc}/h$ per side, and the number of particles, $n_{\rm p} = 1536^3$.

Whereas SLICS featured only a single cosmology however, its successor hosts 26 varying in the mass energy density, Ω_m, the clustering parameter, $S_8 = \sigma_8/\sqrt{\Omega_\mathrm{m}/0.3}$, the dimensionless Hubble constant h and the present-day dark energy equation of state parameter, w_0. These simulations therefore lend greater sensitivity to testing the cosmological dependence of statistics such as clipping over the DH10 mocks, which vary in only Ω_m and the amplitude of density fluctuations, σ_8. The aforementioned *cosmo*-SLICS specifications are also an improvement upon DH10 (see Table 3.1), although they are surpassed in this regard by the Mira Titan simulations ($A_\mathrm{lc} = 5160\,\mathrm{deg}^2$, $L_\mathrm{box} = 2100\,\mathrm{Mpc}/h$, $n_\mathrm{p} = 3072^3$; Heitmann et al., 2016). Nevertheless, *cosmo*-SLICS are the more appropriate choice for use in constraining current lensing data sets as the input cosmologies, detailed in Table 4.1, span the full 95% confidence region of the most up-to-date cosmic shear constraints (see, for example, Figure 2.3), whereas the Mira Titan simulations cover a parameter space too narrow for this purpose.

Identifying the *optimal* distribution of input cosmologies for sampling the parameter space is a non-trivial problem. Depending on the statistic considered, the sensitivity to various combinations of cosmological parameters may change, which in turn has implications for how densely different regions of the input parameter space need be sampled. Various authors have addressed the puzzle of parameter space sampling optimisation, using machine learning (Caron et al., 2019) and Bayesian methods (Rogers et al., 2019). A good first-guess however, used extensively with previous sampling designs (Sacks et al., 1989; Currin et al., 1991; Heitmann et al., 2006; Schneider et al., 2008; Agarwal et al., 2012; Liu et al., 2015; Garrison et al., 2018), is to distribute the input parameters on a *Latin hypercube* (LH; McKay et al., 1979). Such designs have the property that none of the points overlap when projected onto any axis. This takes an agnostic stance on the distribution of the "nodes", treating no dimension as any more important than any other in terms of the required sampling density. An LH configuration was hence adopted for the 25 cosmologies labelled 00–24 in Table 4.1. The remaining cosmology, 'FID', shares the base cosmology of SLICS, and was specifically added such that the two simulation suites could be compared, and so the standard model, ΛCDM, is represented in *cosmo*-SLICS. I demonstrate in this Chapter that the chosen design facilitates per cent level interpolation accuracy across most of the parameter space with my *cosmo*-SLICS emulator.

A key difference in the design of SLICS and *cosmo*-SLICS is the number of realisations per cosmology. Whereas the former has more than 900 realisations for

Table 4.1 *Cosmological parameters in the 25+1 cosmo-SLICS models, with S_8 defined as $\sigma_8/\sqrt{\Omega_\mathrm{m}/0.3}$. In all runs, the baryon density, primordial tilt and neutrino density have been fixed to $\Omega_\mathrm{b} = 0.0473$, $n_\mathrm{s} = 0.969$ and $\Omega_\nu = 0$. Two matched-seed N-body simulations are evolved at each of these nodes, as discussed in the text.*

ID	Ω_m	S_8	h	w_0	σ_8	Ω_c	Ω_Λ
FID	0.2905	0.8231	0.6898	-1.0000	0.8364	0.2432	0.7095
00	0.3282	0.6984	0.6766	-1.2376	0.6677	0.2809	0.6718
01	0.1019	0.7826	0.7104	-1.6154	1.3428	0.0546	0.8981
02	0.2536	0.6133	0.6238	-1.7698	0.6670	0.2063	0.7464
03	0.1734	0.7284	0.6584	-0.5223	0.9581	0.1261	0.8266
04	0.3759	0.8986	0.6034	-0.9741	0.8028	0.3286	0.6241
05	0.4758	0.7618	0.7459	-1.3046	0.6049	0.4285	0.5242
06	0.1458	0.7680	0.8031	-1.4498	1.1017	0.0985	0.8542
07	0.3099	0.7861	0.6940	-1.8784	0.7734	0.2626	0.6901
08	0.4815	0.6804	0.6374	-0.7737	0.5371	0.4342	0.5185
09	0.3425	0.7054	0.8006	-1.5010	0.6602	0.2952	0.6575
10	0.5482	0.6375	0.7645	-1.9127	0.4716	0.5009	0.4518
11	0.2898	0.7218	0.6505	-0.6649	0.7344	0.2425	0.7102
12	0.4247	0.7511	0.6819	-1.1986	0.6313	0.3774	0.5753
13	0.3979	0.8476	0.7833	-1.1088	0.7360	0.3506	0.6021
14	0.1691	0.8618	0.7890	-1.6903	1.1479	0.1218	0.8309
15	0.1255	0.6131	0.7567	-0.9878	0.9479	0.0782	0.8745
16	0.5148	0.8178	0.6691	-1.3812	0.6243	0.4675	0.4852
17	0.1928	0.8862	0.6285	-0.8564	1.1055	0.1455	0.8072
18	0.2784	0.6500	0.7151	-1.0673	0.6747	0.2311	0.7216
19	0.2106	0.8759	0.7388	-0.5667	1.0454	0.1633	0.7894
20	0.4430	0.8356	0.6161	-1.7037	0.6876	0.3957	0.5570
21	0.4062	0.6620	0.8129	-1.9866	0.5689	0.3589	0.5938
22	0.2294	0.8226	0.7706	-0.8602	0.9407	0.1821	0.7706
23	0.5095	0.7366	0.6988	-0.7164	0.5652	0.4622	0.4905
24	0.3652	0.6574	0.7271	-1.5414	0.5958	0.3179	0.6348

its single set of input parameters, the latter features only two simulations for each of the 26 cosmologies. Typically, having such a small number of realisations will lead to a noisy estimate of the mean for a given cosmological statistic. In order to reduce this effect, the initial conditions of the pair of simulations ran for each cosmology, are specifically chosen such that mean of the initial power spectra is within 5% of the theoretical prediction. This is similar to the approach advocated by Angulo & Pontzen (2016), but to avoid the non-Gaussianity induced in the field by their means of identifying a suitable pair of initial conditions, for *cosmo*-SLICS a large number of initial conditions were generated using NICAEA[2], and the pair that offered the greatest noise cancellation was selected (Harnois-Deraps, Giblin & Joachimi, 2019). In this analysis, I use the results of ray-tracing with 25 different random rotations and shifts of the lightcone for each cosmology and matched simulation pair, thereby probing each node in the design 50 times. In Harnois-Deraps, Giblin & Joachimi (2019) we demonstrate that the covariance matrix measured from these realisations for the 'FID' cosmology, closely matches one calculated from more than 800 truly independent SLICS, thereby demonstrating the power of pair-fixing methodology.

One final notable difference between the SLICS and *cosmo*-SLICS, is that the source galaxies in the former simulations followed the DIR-calibrated redshift distribution from the Hildebrandt et al. (2017) cosmic shear analysis of the KiDS-450 data. In contrast, the *cosmo*-SLICS sources trace the more accurate "KV450" redshift distribution, i.e. a recalibration of the former distribution using the five near-infrared bands from the VISTA Kilo-Degree Infrared Galaxy Survey (VIKING), in addition to the four optical bands from KiDS (Hildebrandt et al., 2018).

As in Chapter 3, I calculate the shear correlation functions from the simulations using equation 3.12 in 9 bins of angular separation, θ, logarithmically spaced between 0.5 and 300 arcmin. For this I employ the public TREECORR software (Jarvis, 2015). The fractional difference between the shear correlation functions[3], $\xi_{\pm}$, from the simulations (see equation 3.12) and the theoretical predictions (equation 3.13), using the non-linear matter power spectrum calibration provided by HALOFIT (Takahashi et al., 2012, see Section 1.3.1), are presented for all 26

[2]The use of NICAEA for this purpose presents another deviation from the SLICS design; the initial condition generator used for the former suite, from CUBEP^{3}M, is applicable only to ΛCDM cosmologies.

[3]Here, and for most of this Chapter, $\xi_{\pm}$ specifically denotes "unclipped" shear correlation functions. In Section 4.4, where I apply clipping to *cosmo*-SLICS, clarity is provided by the superscripts 'unclip' and 'clip', as in Chapter 3.

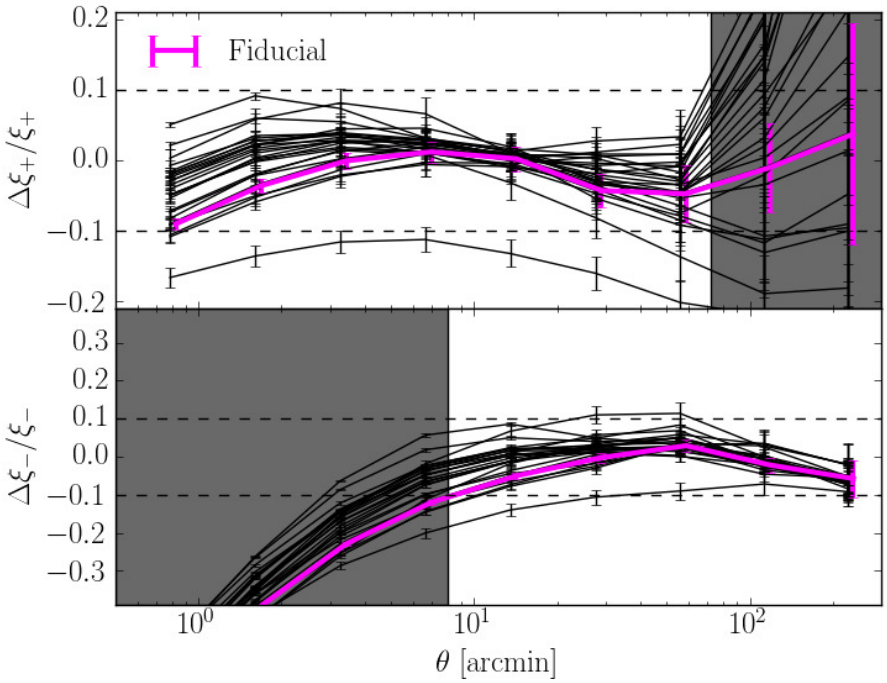

Figure 4.2 *The fractional differences between the cosmo-SLICS $\xi_\pm$ for all models, measured using equation 3.12 and averaged across the 50 lightcones, and the corresponding theoretical predictions (see equation 3.13) (with the* HALOFIT *calibration from Takahashi et al., 2012). The magenta line corresponds to the measurements from the fiducial cosmology, and the grey bands indicate angular scales I exclude in evaluating the emulator accuracy. I plot here the error on the mean.*

cosmologies in Figure 4.2.

4.3 The *cosmo*-SLICS emulator

In this Section, I describe how I tailor a Gaussian process regression emulator to the *cosmo*-SLICS suite and thus predict weak lensing statistics for wCDM cosmologies. I present the accuracy of the emulator's predictions of the shear correlation functions, $\xi_{\pm}$, as a function of the galaxy angular separation and cosmological parameters, by comparing to theoretical predictions from NICAEA, ran with the recalibrated HALOFIT model (Takahashi et al., 2012), and assume these results are representative of those which would be obtained for an arbitrary cosmological statistic measured from these simulations. I further show to what extent the accuracy of the emulator depends on the distribution of the cosmological parameters, $\boldsymbol{\pi} = \{\Omega_{\rm m}, S_8, h, w\}$, rather than the noise on the training set predictions, by replacing the simulated $\xi_{\pm}$ from *cosmo*-SLICS with the noise-free theoretical $\xi_{\pm}$. I use the public SCIKIT LEARN Gaussian process regression code[4] for all analyses in this Section.

4.3.1 Emulation strategy

Emulators are trained by finding values for the hyperparameters which define a distribution of functions that are optimally consistent with all realisations in the training set. In this work, I fit for these using the method built-in to SCIKIT LEARN, which employs a gradient ascent optimisation of the marginal likelihood expressed in equation 4.14. Emulator accuracy is also strongly affected by the shape of the observable being predicted, performing best for smooth monotonic functions with narrow dynamic ranges. Since the $\xi_{\pm}(\theta)$ statistics vary over orders of magnitude, $\ln \xi_{\pm}(\theta)$ presents a wiser choice of quantity to emulate. I find that emulation performance is further improved by decomposing the $\ln \xi_{\pm}(\theta)$ observable into a linear sum of n_{Φ} orthogonal basis vectors, $\phi^i_{\pm}(\theta)$ where $i \in [1, n_{\Phi}]$, using a principal component analysis (PCA),

[4] `https://scikit-learn.org/stable/modules/gaussian_process.html`

$$\ln \xi_{\pm}(\theta; \boldsymbol{\pi}) = \mu_{\pm}(\theta) + \sum_{i=1}^{n_\phi} \phi_{\pm}^{i}(\theta) w_{\pm}^{i}(\boldsymbol{\pi}) + \epsilon_{\pm}^{i}(\boldsymbol{\pi}) + \epsilon_{\pm}^{\mathrm{PCA}}(\boldsymbol{\pi}) \,, \tag{4.18}$$

where $\mu_{\pm}(\theta)$ is the mean across the training set $\ln \xi_{\pm}(\theta; \boldsymbol{\pi})$ predictions, and the orthogonal basis functions, $\phi_{\pm}^{i}(\theta)$, are calculated from a PCA of the mean-subtracted training set. In this formalism, the weight parameters, $w_{\pm}^{i}(\boldsymbol{\pi})$, specifying how much each basis function contributes to the $\ln \xi_{\pm}(\theta; \boldsymbol{\pi})$ recipe for a given $\boldsymbol{\pi}$, now become the target of my emulator's predictions, taking the place of $y(\boldsymbol{\pi})$ in equation 4.1, rather than $\ln \xi_{\pm}(\theta; \boldsymbol{\pi})$ itself. The $\epsilon_{\pm}^{\mathrm{PCA}}$ and $\epsilon_{\pm}^{i}$ are terms arising from two different sources of error, that vary slightly between the *cosmo*-SLICS cosmologies.

$\epsilon_{\pm}^{\mathrm{PCA}}$ arises if one uses an insufficient number of basis functions to reconstruct the emulated statistic. PCA decomposition is a standard procedure (see for example Habib et al., 2007; Schneider et al., 2008; Heitmann et al., 2016), facilitating improvements in emulation time where n_Φ is less than the length of the statistic of interest, in this case determined by the number of θ bins. Computational expense is not a problem for my $\xi_{\pm}(\theta)$ measured from *cosmo*-SLICS however, consisting of only 9 bins in angular separation. Hence I simply set n_ϕ equal to the maximum of 9 basis functions, for perfect PCA reconstruction of the $\ln \xi_{\pm}(\theta; \boldsymbol{\pi})$. I verified however that this number is sufficient to reconstruct more than 99.99% of the variance in theoretical $\ln \xi_{\pm}$ sampled in 70 bins and that using more basis functions has minimal effect on the emulator accuracy. Hence, with 9 basis functions the error induced from the PCA reconstruction is negligible.

The remaining error term, $\epsilon_{\pm}^{i}(\boldsymbol{\pi})$, comes from the Gaussian noise, denoted by $\epsilon_{\mathrm{n}}(\boldsymbol{\pi})$ in equation 4.1, arising from uncertainties on the training set. To inform the emulator of the error on the *cosmo*-SLICS predictions, I first calculate the standard deviation of the $\ln \xi_{\pm}(\theta; \boldsymbol{\pi})$ across the 25 lightcones and 2 seeds for each cosmology, $\sigma_{\pm}(\theta; \boldsymbol{\pi})$. I translate this into uncertainties on the PCA weights by computing the upper and lower bounds, given by

$$w_{\pm}^{i,\mathrm{upper}} = \sum_{m=1}^{9} \Phi_{\pm}^{i}(\theta_m) \left[\ln \xi_{\pm}(\theta_m) + \left(\sigma_{\pm}(\theta_m) / \sqrt{50} \right) \right] ,$$

$$w_{\pm}^{i,\mathrm{lower}} = \sum_{m=1}^{9} \Phi_{\pm}^{i}(\theta_m) \left[\ln \xi_{\pm}(\theta_m) - \left(\sigma_{\pm}(\theta_m)/\sqrt{50} \right) \right]. \tag{4.19}$$

Here, the $\xi_{\pm}$ is the average of the measurements for the different lightcones and seeds per cosmology, the factor $\sqrt{50}$ is included to scale the standard deviation to an error on the mean, and for simplicity I have dropped the dependence on the cosmological parameters. The error on the PCA weight, approximated as

$$\epsilon_{\pm}^{i} = \frac{1}{2} \left(w_{\pm}^{i,\mathrm{upper}} - w_{\pm}^{i,\mathrm{lower}} \right), \tag{4.20}$$

serves as the standard deviation of the Gaussian distribution from which the $\epsilon_{\mathrm{n}}(\boldsymbol{\pi})$ is sampled. In this work I also emulate noise-free HALOFIT predictions to test the accuracy facilitated purely by the distribution of *cosmo*-SLICS nodes. In these cases I set the ϵ_{n} for all $\boldsymbol{\pi}$ to the arbitrarily-small constant default value in SCIKIT LEARN[5].

All results presented in this work demonstrating the emulator performance correspond to accuracies in the inferred $\xi_{\pm}$, and not the logarithmic transforms of these statistics nor the weight vectors, $\boldsymbol{w}_{\pm}(\boldsymbol{\pi})$.

4.3.2 Emulator results

Having established my emulation strategy, I now seek to test how accurately I can predict the $\xi_{\pm}(\theta; \boldsymbol{\pi}^*)$ corresponding to an ensemble of trial cosmologies, $\boldsymbol{\pi}^*$. It is too computationally expensive to produce a fine grid of trial predictions covering the entire 4D parameter space, against which emulator accuracy can be tested. Instead I generate two separate ensembles of trial coordinates. The first, which I refer to as the "grid" ensemble, $\boldsymbol{\pi}^*_{\mathbf{g}}$, seeks to illuminate how accurately I can reproduce the predictions for different regions of the emulation space. This ensemble consists of six cosmological parameter grids, with dimensions 50×50, for the six different 2D projections of the 4D space. For each grid in which

[5]One cannot set $\epsilon_{\mathrm{n}} = 0$ or the marginal likelihood, entering into the posterior from which predictions are sampled, becomes singular.

two parameters vary, the remaining two are fixed to the corresponding fiducial values from $\{\Omega_\mathrm{m} = 0.3251, S_8 = 0.7524, h = 0.7082, w_0 = -1.254\}$, selected on account of being the centre of the *cosmo*-SLICS training set. This ensemble is useful for identifying for which combinations of cosmological parameters my emulator will perform best and where there is room for improvement. The second, "bulk", ensemble, $\boldsymbol{\pi}^*_\mathbf{b}$, consists of 300 cosmologies which probe the bulk accuracy of the emulator throughout the emulation space by varying in all 4 parameters simultaneously. I sample these cosmologies from an independent 4-dimensional Latin hypercube with dimensions equal to that of the *cosmo*-SLICS training set.

A crucial ingredient in evaluating the emulator's accuracy is a theoretical prediction with which to compare the emulator's. However, the fact that the *cosmo*-SLICS $\xi_\pm(\theta;\boldsymbol{\pi})$ differ from the corresponding theoretical predictions, as shown by Figure 4.2, means that the emulator will not recover the theoretical predictions used to gauge accuracy, even at the nodes. The disagreement between the two arises not only because of residual noise and small, non-linear angular scales that are not fully resolved in *cosmo*-SLICS, but also because of inaccuracies in the HALOFIT model prescription. These are caused by resolution limitations also present in the simulations used to calibrate the Takahashi et al. (2012) fitting function methodology outlined in Section 1.3.1, and also the fact that the range of input cosmologies for these mocks does not cover the full range of the *cosmo*-SLICS input parameters, especially in the w_0 dimension. This is shown by the distribution of black stars (Takahashi et al., 2012 simulation nodes) relative to the magenta circles (*cosmo*-SLICS nodes) in the upper-left panel of Figures 4.3 and 4.4. The effect of the imperfections in the *cosmo*-SLICS (training) and HALOFIT (trial) predictions on the emulator performance cannot be *completely* disentangled. Therefore, my results for the accuracy of the *cosmo*-SLICS emulator should be regarded as a conservative, "worst case scenario"; performance would likely improve with perfect trial predictions to compare with.

To suppress the contribution of inaccuracies on non-linear scales, I consider only the $0.5 < \theta < 72$ arcmin angular range for ξ_+ and $8.0 < \theta < 300$ arcmin for ξ_- in evaluating the emulator accuracy. This roughly corresponds to the scales used in the Hildebrandt et al. (2017) cosmic shear analysis, but with a slightly higher lower limit for ξ_-, to select an angular range with good agreement between *cosmo*-SLICS and NICAEA predictions for this statistic (see Figure 4.2). In addition to testing the emulator with the *cosmo*-SLICS training set, I also test with noise-free NICAEA $\xi_\pm(\theta;\boldsymbol{\pi})$ training sets of various sizes. Whereas training with *cosmo*-

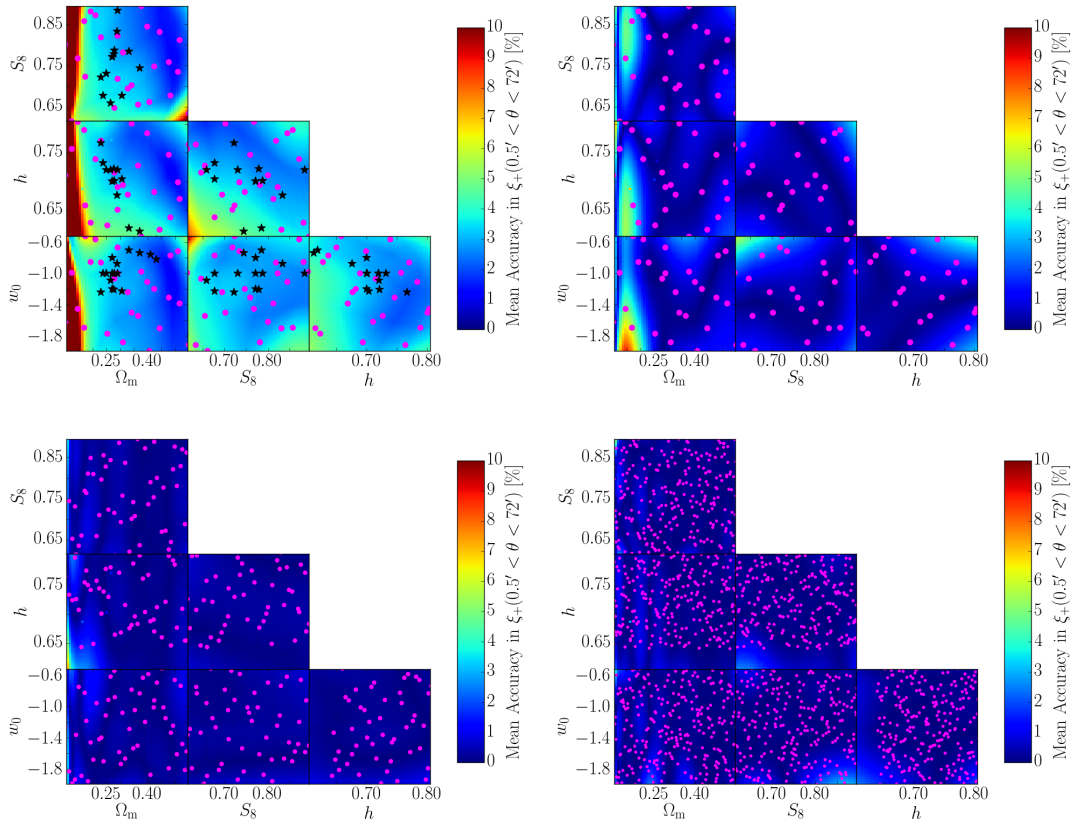

Figure 4.3 *The observed emulator accuracies for ξ_+, averaged between 0.5 and 72 arcmin, with the grid ensemble of trial cosmologies, $\boldsymbol{\pi}^*_{\mathbf{g}}$, shown by the colour maps, when trained on the 26 cosmo-SLICS predictions (upper-left) and 26, 50 and 250 noise-free* NICAEA *predictions (upper-right, lower-left, lower-right respectively). The training nodes are shown by the magenta circles. The black stars in the upper-left panel show the input parameters of the Takahashi et al. (2012) simulations over the cosmo-SLICS parameter volume (their two highest Ω_m nodes have h and S_8 values that exceed the boundaries). For each grid in which two cosmological parameters vary, the remaining two are fixed to the corresponding fiducial values from $\{\Omega_m = 0.3251, S_8 = 0.7524, h = 0.7082, w_0 = -1.254\}$. The contrast between the upper panels, for which the training cosmologies are the same, indicates the extent to which simulation noise and inaccuracies in both the simulations and theoretical predictions degrade the apparent emulation accuracy.*

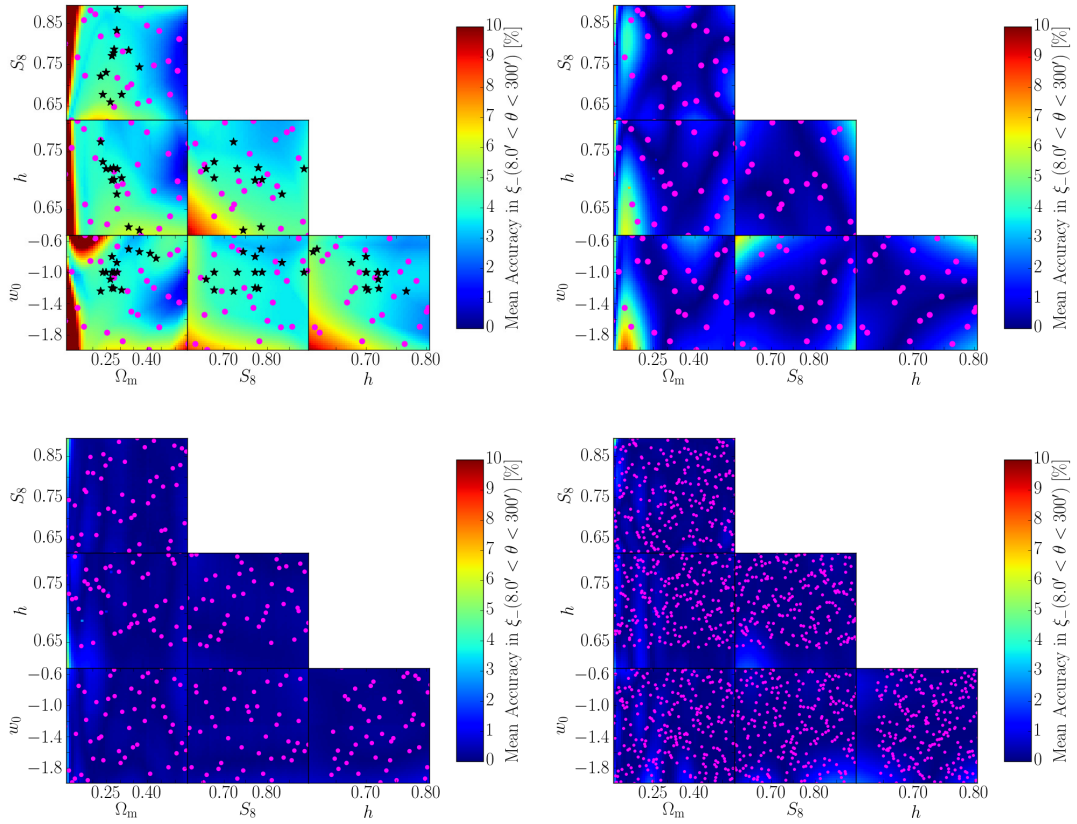

Figure 4.4 *The same as Figure 4.3 but for ξ_- with accuracies averaged between 8 and 300 arcmin.*

SLICS probes how emulator accuracy is affected by the limitations of both the simulations and the trial HALOFIT predictions, the latter isolates how well the emulator is able to interpolate $\xi_\pm$ statistics from finite distributions of points.

The accuracies for the emulated ξ_+ and ξ_-, averaged across the aforementioned θ ranges, for the grid ensemble are shown in Figures 4.3 and 4.4 respectively. The upper-left panel in either figure shows the accuracies when training on *cosmo*-SLICS. The remaining panels correspond to the noise-free NICAEA sets, increasing in size from that of our simulation suite, to 50 and finally 250 training predictions.

When training on the *cosmo*-SLICS mocks themselves, I observe emulation accuracies $\leq 5\%$ in both ξ_+ and ξ_- across much of the emulation space, suggesting that the *cosmo*-SLICS nodes are well-placed to sample the cosmological dependence on these parameters. Noticeably worse accuracies of 5–10% manifest at low Ω_m values however. Features such as this are expected at the edges of the training set, where there is a lower concentration of nodes from which to interpolate. I also note that this region is not sampled at all by the HALOFIT training set, hence the predictions completely rely on extrapolation. Similarly, I find edge-effects at some corners in the other projections, but again most of these were not part of the model calibration. The high dependence of the $\xi_\pm$ statistics on Ω_m is perhaps the reason why the feature is strongest in the 2D planes with this parameter. Comparison of the upper-left panel to the upper-right, where the training predictions are replaced by noise-free theoretical $\xi_\pm$, reveals how much of the inaccuracy seen when training on *cosmo*-SLICS can be attributed to noise in the simulations and differences between *cosmo*-SLICS and the HALOFIT prescription. The average observed accuracy reduces to $\leq 2\%$ although worse performance continues to be observed at $\Omega_\mathrm{m} < 0.2$.

The lower two panels of Figures 4.3–4.4 show the emulation accuracy when the training sets consist of 50 and 250 noise-free theoretical predictions respectively, with nodes indicated by the magenta points[6]. I find that these numbers of training points are sufficient to achieve accuracies around the level of 1% across all of the explored parameter space, and that the improvement between 50 and 250 nodes

[6]The h-range for these training nodes, $\in [0.65, 0.8]$, reflects that of a previous experimental design for the *cosmo*-SLICS suite, before the lower limit of $h = 0.6$ was chosen to better represent observational constraints. The cosmologies of the grid ensemble were selected to cover the range of the *present cosmo*-SLICS suite, hence why the 50 and 250 magenta points do not cover the full grid size in projections featuring h. It is not necessary to adjust the distribution of 50 and 250 training points however, since these training sets already permit very accurate extrapolation to these low h values.

is negligible, suggesting the former already samples the cosmological dependence of the $\xi_{\pm}$ very well. The noticeable improvement increasing from the 26 to 50 training nodes could be considered argument for running *cosmo*-SLICS at 50 distinct cosmologies. However, I remind the reader that given an amount of computing resources fixed to 50 runs, opting for running all different cosmologies would lack the benefits of the matched-pair simulation strategy, which facilitate significant reduction in the noise of the estimate of the true $P(k)$ and $\xi_{\pm}(\theta)$ (see Section 4.2 and Harnois-Deraps, Giblin & Joachimi, 2019, for discussion). I interpret these results instead as evidence that augmenting *cosmo*-SLICS with an additional 24 cosmologies each having the matched-pair simulations, would be quite beneficial to emulation performance, especially at low $\Omega_{\rm m}$ values, but going beyond this sized suite is unnecessary. Also worth considering is that in this parameter space, baryons contribute to up to 50% of the total matter density, hence will likely have a different and stronger feedback on the lensing signal.

Figure 4.5 presents the results of exploring the bulk accuracy of the emulator, where all 4 cosmological parameters are varied simultaneously in the 300 trial ensemble. Here I show the fraction of trial cosmologies for which the mean accuracy across the fiducial angular separation range is better than the threshold, A_c, plotted on the horizontal axis. I see that when training on the $N = 26$ noise-free theoretical $\xi_{\pm}$, the emulator recovers more than 90% of the trial predictions to better than 5% accuracy (solid magenta and grey curves). Further inspection reveals that the trial cosmologies with mean accuracies worse than 5% all reside on the edges of the hypercube defined by the training set, where emulation is expected to perform less well. In particular, I find cosmologies with $\Omega_{\rm m} < 0.2$ over-represented, by factors of 3 (considering ξ_{-} predictions) and 5 (considering ξ_{+}), in the set of trials which failed to achieve this mean accuracy. This is consistent with my accuracy tests involving the grid ensemble, further pointing to a necessity for extra training nodes to improve the emulation for this part of the parameter space.

The dashed lines in Figure 4.5 demonstrate the cumulative mean accuracy when I instead train on the *cosmo*-SLICS predictions. We see a decrement in performance relative to the noise-free training set results as expected; for 25%(33%) of the trial cosmologies, the mean emulator accuracies for the $\xi_{+}(\xi_{-})$ statistics are worse than 5%. The slight asymmetry in performance for these two statistics is also consistent with grid ensemble tests, where accuracy for emulating ξ_{+} (Figure 4.3) when training on the *cosmo*-SLICS predictions was slightly better than emulations of

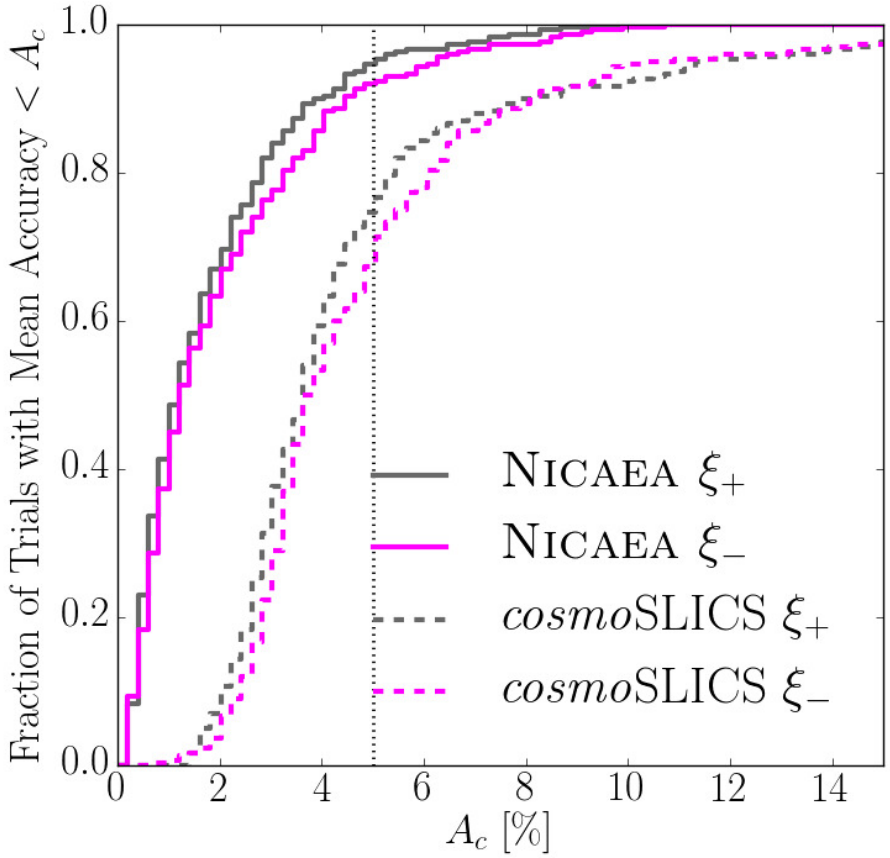

Figure 4.5 *The fraction of the trial cosmologies from the bulk ensemble, $\boldsymbol{\pi}^*_{\mathbf{b}}$, with accuracies, averaged over a range of angular scales (0.5–72 arcmin for ξ_+, 8.0–300 arcmin for ξ_-), better than the value, A_c, plotted on the horizontal axis. The grey curves correspond to ξ_+ predictions, magenta to ξ_-. The solid curves result from training the emulator on the noise-free theoretical predictions from* NICAEA*, whereas the dashed result from training on cosmo-SLICS itself. The decrement in performance when training on cosmo-SLICS is expected due to the added noise in the training set and inaccuracies in the theoretical predictions.*

ξ_- (Figure 4.4). I emphasise once again that these results represent a conservative view of emulation accuracy given *cosmo*-SLICS as a training set, owing to the imperfections of the theoretical predictions used for comparison. I hence conclude that the simulation suite permits emulated predictions with accuracies at the level of $\simeq 5\%$ or better.

Future work will investigate alternative interpolation strategies, such as sparse polynomial chaos expansion, as exercised by Euclid Collaboration et al. (2018). This method models the unknown dependence of output statistics on input variables, as a summation of multivariate polynomials weighted by vector coefficients which are unknown *a priori*. The coefficients are sought as the solution of a 'least angle regression', similar in principle to a straight-forward least squares regression, but with the inclusion of a term which adjusts the trade-off between the goodness-of-fit and complexity of the solution (Blatman & Sudret, 2011). This has the effect of minimising the global error at the expense of imperfectly recovering the training predictions at the nodes.

4.4 Clipping *cosmo*-SLICS

In this Chapter, I have tailored an emulation strategy to *cosmo*-SLICS and tested the accuracy of its predictions for a statistic for which theoretical predictions are at hand. I now proceed to use these simulations for their intended purpose: exploring the cosmological dependence of statistics lacking an analytical prediction. Obvious choices of statistic are the clipped shear correlation functions from Chapter 3.

I apply clipping to the 26 *cosmo*-SLICS cosmologies using the pipeline outlined in Section 3.3.2. This consists of adding galaxy shape noise to the shear, constructing and clipping smoothed mass maps, before inferring the clipped shear signal. I assume the clipping threshold, $\kappa^{\rm c} = 0.010$, and smoothing scale, $\sigma_{\rm s} = 6.6$ arcmin, identified as appropriate for isolating non-linear regions of the SLICS in Section 3.3.3, to be fairly unaffected by the slight difference in the redshift distribution to that of *cosmo*-SLICS. Thus I once again adopt these values, and the corresponding photometric redshift cut $z_{\rm B} \in [0.5, 0.9]$, in performing the clipping here. Furthermore, to minimise the shot noise introduced by the inclusion of galaxy shape noise, on which the clipped signal is critically dependent, following the methodology outlined in Section 3.3.5, I process 25 different shape

noise realisations for each of the 50 lightcone rotations per cosmology. This corresponds to one-third the number required for the DH10 simulations, owing to *cosmo*-SLICS having $\sim 3\times$ larger lightcone area.

Figure 4.6 shows the clipped and unclipped shear correlation functions for the 26 *cosmo*-SLICS cosmologies, colour-coded by S_8. The upper panel of each plot corresponds to $\xi_{\pm}^{\rm clip}$, the lower to $\xi_{\pm}^{\rm unclip}$. We see that there is a fairly strong correlation in the amplitude of the measurements with S_8 as expected, although visually this appears weaker than that which was observed for the DH10 simulations in Figure 3.8. The dilution of this effect here is caused by the changing values of h and w_0 across the simulations, which are fixed in the case of DH10. The precise cosmology corresponding to each clipped/unclipped correlation function in this Figure can be identified from the data point with the corresponding colour on the projections map shown in Figure 4.7.

4.4.1 Emulating clipped *cosmo*-SLICS

In training the *cosmo*-SLICS emulator on the clipped $\xi_{\pm}$ measurements presented in Figure 4.6, I simply follow the procedure outlined in Section 4.3, replacing the unclipped with its clipped counterpart. Prior to using the emulator's predictions in obtaining cosmology constraints, however, it is prudent to gauge the emulation accuracy for the clipped statistics. With no analytical predictions to hand, one cannot simply repeat the thorough tests performed for the $\xi_{\pm}^{\rm unclip}$ in Section 4.3.2. Instead I conduct a leave-one-out cross-validation (CV) test. This consists of iteratively omitting each of the simulation nodes from the training, and checking how well the emulator can recover the prediction at the missing node.

The range of accuracies obtained in this test, as a function of angular separation, θ, are shown by the pink and grey bands, respectively for the clipped and unclipped ξ_{+}, in Figure 4.8. From this Figure I omit the results obtained for three nodes residing at the edges of the $\Omega_{\rm m} - S_8$ plane shown in Figure 4.7 (specifically, the two blue-coloured and one yellow-coloured nodes at the lower and left boundaries respectively). As expected, the CV reveals particularly poor recovery for these cosmologies, due to the lower density of points from which to interpolate, and the decrement in interpolation accuracy observed in this corner of the parameter space in Figures 4.3-4.4. The CV results for these cosmologies are excluded from this Figure, as they are unrepresentative of the bulk emulation accuracy I wish to infer with this test. Note, that unlike the results presented

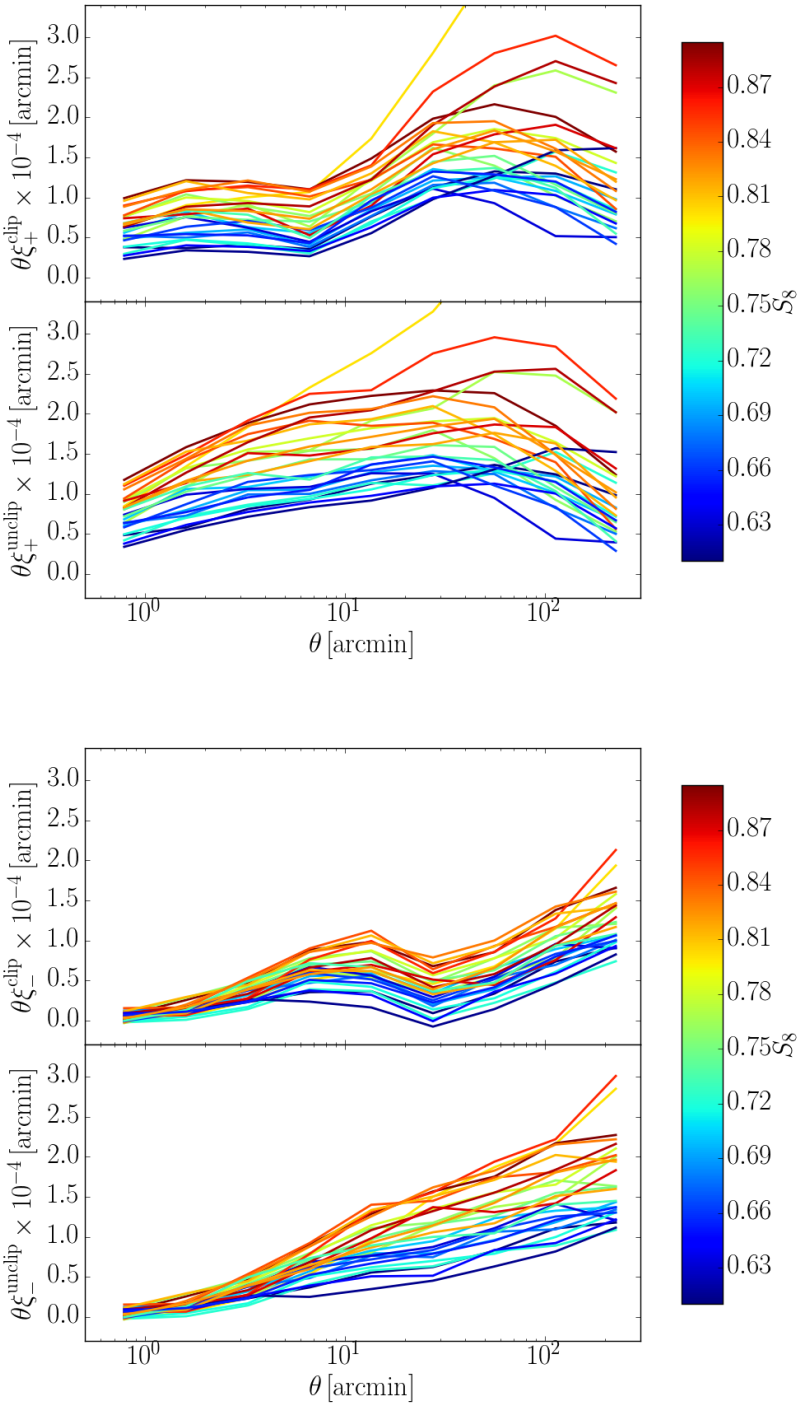

Figure 4.6 *The clipped (upper panel of each plot) and unclipped (lower panel of each plot) shear correlation functions from cosmo-SLICS, weighted by θ and colour-coded by S_8. The upper and lower plot show ξ_+ and ξ_- respectively. The values of the clipping threshold and smoothing scale are those identified as appropriate for SLICS and the KiDS-450 data in Section 3.3.3, $\kappa^{\rm c} = 0.010, \sigma_s = 6.6$ arcmin. The cosmological parameters for each measurement on these plots can be identified from the corresponding coloured data point on the projections map shown in Figure 4.7.*

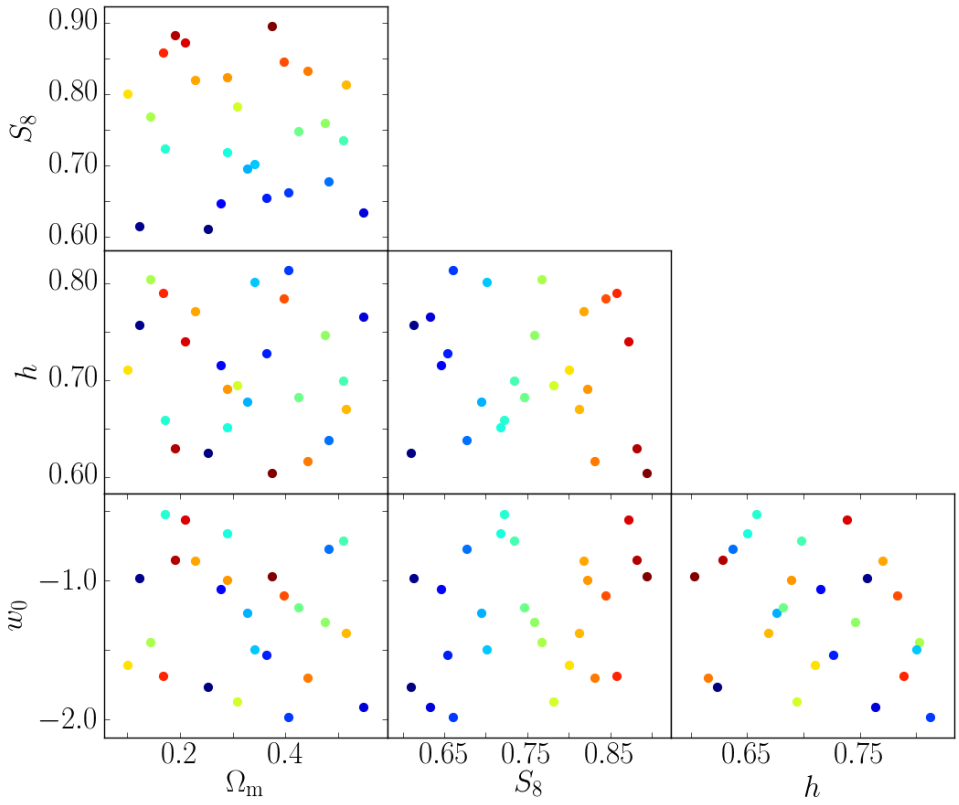

Figure 4.7 *The cosmo-SLICS cosmologies colour-coded by S_8.*

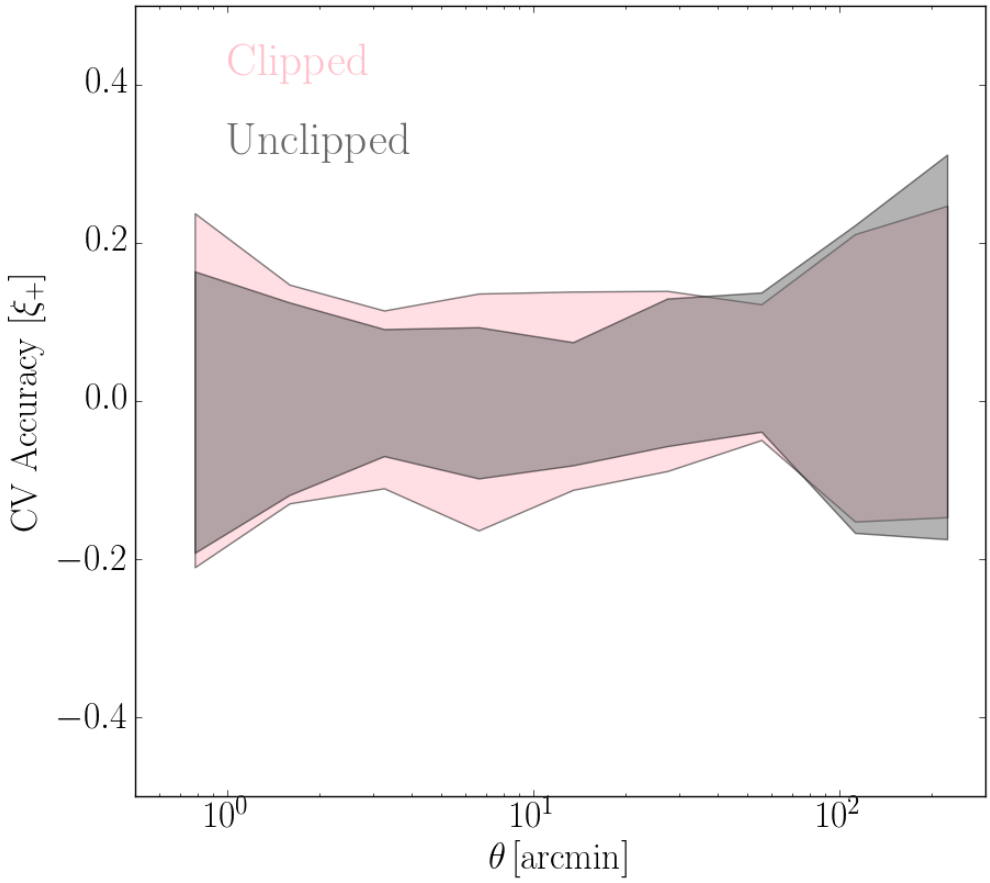

Figure 4.8 *The range of emulation accuracies achieved in the cross-validation (CV) tests with the clipped (pink) and unclipped (grey) ξ_+ from cosmo-SLICS. The results for the three nodes at the lower and left edges of the $\Omega_\mathrm{m} - S_8$ plane (see Figure 4.7) are omitted for the reasons explained in the text.*

in Section 4.3.2, I now include shape noise in the calculation of the $\xi_\pm^\mathrm{unclip}$ from *cosmo*-SLICS, via equation 3.12, to aid in the comparison of emulation accuracy between this and the clipped.

Figure 4.8 reveals comparable accuracy in the emulation of the clipped and unclipped ξ_+, but with the former performing $\sim 5\%$ worse for $1 \lesssim \theta \lesssim 60$ arcmin. Quantitatively similar results are also obtained in the CV with the clipped and unclipped ξ_-. This is to be expected given the slightly more complex shape of the $\xi_\pm^\mathrm{clip}$ at intermediate angular scales compared to the $\xi_\pm^\mathrm{unclip}$ (see Figure 4.6). The accuracy for the clipped correlation functions may be improved, should the emulator be tasked with predicting an alternative transform of these statistics than the natural logarithm, found to be effective for the unclipped in Section 4.3.

This I leave for future investigation however.

4.4.2 Cosmic shear with clipped *cosmo*-SLICS

Having inferred that the emulator performs comparably in predicting the clipped and unclipped $\xi_\pm$, I now apply its predictions to obtaining combined clipped and unclipped cosmological constraints with data. This follows the general likelihood analysis of Section 3.4.1, with a few a key differences, highlighted in this Section.

First of all, it is not valid to use the clipped and unclipped $\xi_\pm$ measured from KiDS-450 in the previous Chapter (see Figure 3.9), as the data vectors in the likelihood analysis here. This is because the differences in the redshift distributions inferred in the KiDS-450 (Hildebrandt et al., 2017) and KV450 (Hildebrandt et al., 2018) analyses change the photometric redshift binning of the observed source galaxies. Hence, the clipped/unclipped $\xi_\pm$ obtained with KV450-based redshift cuts, as is the case with *cosmo*-SLICS, cannot be readily compared to those with KiDS-450-based cuts, as is the case with the data[7]. Remeasuring these quantities from the KV450 data set is beyond the scope of this analysis, owing to the computational expense in calibrating the mask bias (see Section 3.3.4), which differs between KiDS-450 and KV450 with the survey footprints. I therefore use the emulator to produce mock clipped and unclipped data vectors corresponding to the fiducial cosmology of the grid ensemble from Section 4.3.2, $\{\Omega_\mathrm{m} = 0.3251, S_8 = 0.7524, h = 0.7082, w_0 = -1.254\}$, selected as being the approximate centre of the *cosmo*-SLICS parameter space.

Regarding the covariance, Harnois-Deraps, Giblin & Joachimi (2019) demonstrate with Fisher forecasts, that there are significant changes in contour area, depending on whether the cosmologies of the covariance and data are well- or ill-matched. Hence, I simply use the covariance (clipped auto-, unclipped auto-, and clipped-unclipped cross-covariance; see Figure 3.12) from the simulation with the smallest Euclidean distance from the data cosmology. Following the procedure enacted in Section 3.3.2, I also once again scale the covariance, this time to correspond to 18,000 deg^2, approximately the sky coverage of the next-generation survey, the Large Synoptic Survey Telescope (LSST Dark Energy Science Collaboration, 2012, see Section 6.2 for further discussion), after 10 years of observations[8]. I

[7]The change in the redshift distributions also prevents direct comparison between the simulated $\xi_\pm^\mathrm{clip}$ from *cosmo*-SLICS and those from SLICS/DH10

[8]It is important to note, however, that LSST will feature different noise properties than

reiterate that this scaling, argued by Schneider et al. (2002c), is an approximation, and neglects the impact of the survey geometry (Troxel et al., 2018).

As mentioned in Section 3.4.2, the cosmological bias (see Section 3.3.4) and irreducible uncertainty in the DH10 model predictions, prevents me from reaching meaningful combined constraints for a next-generation survey size, using these simulations. I find, however, that the statistical uncertainty on the emulated clipped and unclipped predictions, given by the width of the Gaussian posterior from which trial functions are sampled (equation 4.12), is subdominant to the sampling variance of the survey, even with its large sky coverage. For the following likelihood analyses, I therefore present the results without propagation of the statistical error on the model predictions, which I have verified does not alter my conclusions. There still remains a systematic offset in the emulator's predictions, conservatively estimated to be of order $\sim 5\%$, which require additional simulation nodes to reduce, as demonstrated by Figures 4.3-4.5, and would be necessary for applying these resources to the next-generation lensing data. Lacking a theoretical prescription for the clipped measurements however, I am unable to calibrate this offset in the current test. Nevertheless, since the data and model predictions are all produced by the emulator, this bias should approximately cancel out in the calculation of the likelihood.

As a further caveat, I refrain from including the Hartlap et al. (2007, see equation 3.20) correction for the bias in the inverted covariance imposed by noise. The advocated correction is for simulations which lack the pair-fixing of initial conditions to achieve noise cancellation, as featured in *cosmo*-SLICS. Hence, applying this factor will overestimate the bias present in my inverted simulated covariance matrices. As mentioned in Section 4.2, the *cosmo*-SLICS covariance are consistent with those obtained from more than 800 independent SLICS realisations, for which the Hartlap correction is negligible. It is therefore reasonable to omit this ingredient from the likelihood.

A further departure of this likelihood analysis from the one discussed in Section 3.4.1, is the subject and method of the interpolation from the simulations. In Chapter 3, I calculate the likelihood at the DH10 simulation nodes, and use the 2D radial basis function interpolation to get the likelihood at arbitrary cosmologies in the $\Omega_{\rm m} - \sigma_8$ and $\Omega_{\rm m} - S_8$ planes. This route is taken on account of the

KiDS in its imaged galaxies. In turn this affects the covariance and the clipping parameters (see Section 3.3.3), both of which are tailored to the latter survey. I neglect these details here, this again serving as a proof-of-concept for the effect of clipping on cosmic shear constraints.

noise associated with interpolating the clipped/unclipped ξ_+ statistics themselves via this method (explored in Appendix 3.B.2). I find, however, that the GP interpolation of these statistics from *cosmo*-SLICS, is far less subject to noise, such that the unclipped contours obtained via this method, are consistent with those obtained using the theoretical predictions (see equation 3.13). In this test, I therefore use the *cosmo*-SLICS emulator to interpolate the clipped/unclipped ξ_+ directly, from the simulation nodes to the arbitrary trial cosmologies. For simplicity, I use only the ξ_+ measurements in this test, and all angular scales as in Section 3.4.1, although it is straight-forward to extend this to also include the ξ_-. The likelihood of each cosmology is calculated with equation 3.19 as before.

One final difference from the likelihood analysis of Section 3.4.1 is the means of sampling the parameter space. Whereas previously, the likelihood was computed on 2-dimensional $\Omega_\mathrm{m} - \sigma_8$ and $\Omega_\mathrm{m} - S_8$ grids, in this case the likelihood occupies a 4-dimensional parameter space, and so it is more computationally efficient to sample the probability via MCMC methods, described briefly in Section 4.1.1. For this I use the publicly available EMCEE code (Foreman-Mackey et al., 2013). I execute separate chains for the unclipped, clipped and combined analyses, with the only difference being the covariance and shear correlation functions entering into the likelihood, expressed in equation 3.19, calculated at each step in the chains. In all cases, I apply wide tophat priors over the $\{\Omega_\mathrm{m}, S_8, h, w_0\}$ parameter spaces, equivalent in extent to the grids shown in Figures 4.3-4.4.

Figure 4.9 presents the 68% and 95% confidence intervals projected onto each of the 2-dimensional planes in the parameter space, as well as the 1-dimensional marginalised probability distributions of each parameter, obtained from the chains in the unclipped (orange), clipped (magenta) and combined (black) analyses.

Whilst still consistent with the cosmology of the data (red data points and lines), the combined analysis clearly offers improved constraints upon the unclipped, as indicated by the smaller area of the 2-dimensional black contours compared to the orange, and by the relative widths of the corresponding 1-dimensional distributions. Considering the marginalised constraints on Ω_m, S_8, h and w_0, I find that the combined analysis yields 1σ error bars which are $25\%, 21\%, 26\%$ and 18% tighter than those of the unclipped, respectively. I remind the reader, that the covariance used in the combined analysis includes the cross-covariance between the clipped and unclipped statistics, as described in Section 3.4.1, to avoid double-counting information. The clipped analysis therefore contains information on each

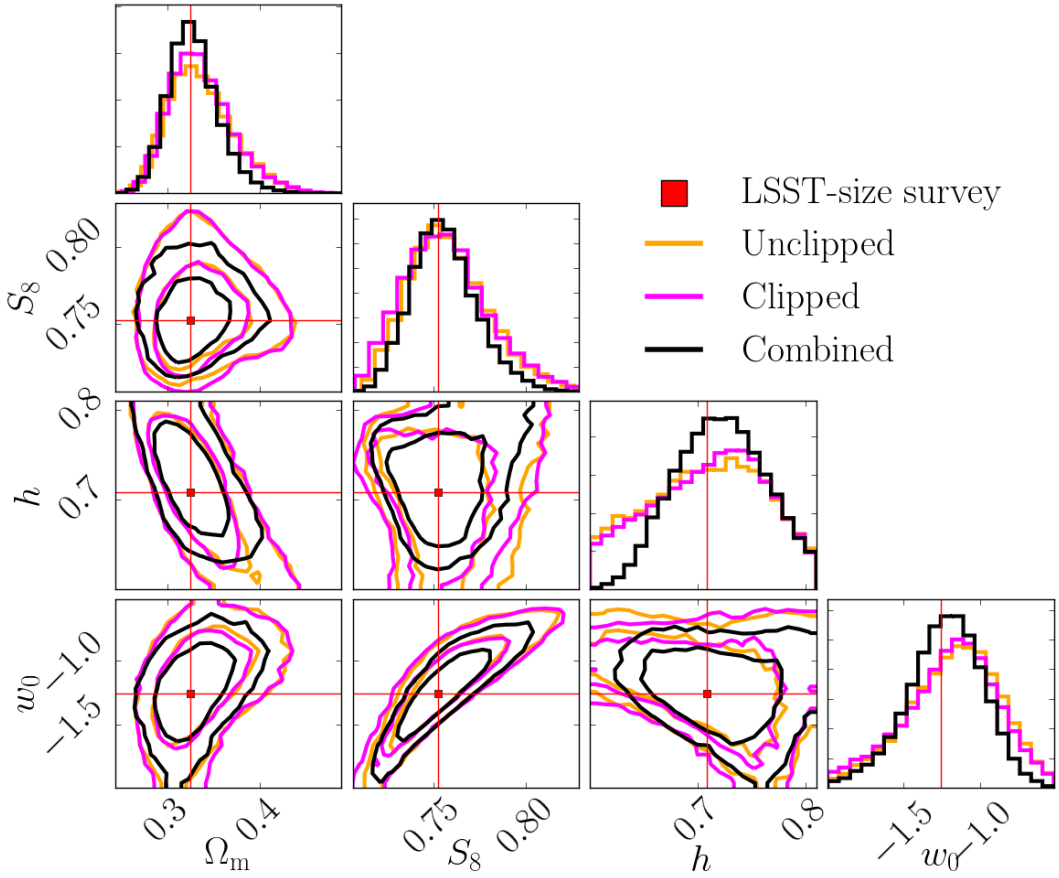

Figure 4.9 *The unclipped (orange), clipped (magenta) and combined (black) constraints on mock data from an LSST-sized survey, the cosmology of which is designated by the red data points/lines:* $\{\Omega_\mathrm{m} = 0.3251, S_8 = 0.7524, h = 0.7082, w_0 = -1.254\}$. *The contours show the 68% and 95% confidence intervals in each 2-dimensional projection of the parameter space. The histograms indicate the marginalised posterior distributions over each parameter. The data and model clipped and unclipped* ξ_+ *entering into the likelihoods in this MCMC analysis, were produced with the cosmo-SLICS emulator after training on the simulations. The covariance is taken from the cosmo-SLICS with the closest cosmology to that of the mock data.*

of these four cosmological parameters which is independent of the unclipped. It is particularly interesting to note the considerable improvement on h and w_0 - a finding which is of course, only made possible by the enhanced design specifications of *cosmo*-SLICS over DH10, featuring variation of these two extra cosmological parameters across the simulations.

4.5 Conclusions

The implementation of many novel statistics in constraining cosmological parameters hinges on a) numerical simulations for use in exploring their cosmological dependence, and b) robust means to interpolate the simulation measurements to arbitrary coordinates in the parameter space. In this Chapter I have addressed these points, presenting the *cosmo*-SLICS numerical simulation suite (Harnois-Deraps, Giblin & Joachimi, 2019) and a Gaussian process regression emulator tailored to this resource.

The *cosmo*-SLICS, described in Section 4.2, are composed of 26 individual cosmologies varying in $\{\Omega_{\rm m}, S_8, h, w_0\}$, across a range bounded by present cosmic shear constraints. The initial conditions of the two realisations at each node are designed such that the expectation of the matter power spectrum is a fairly unbiased estimate of the theoretical prediction. I show that the *cosmo*-SLICS emulator (Section 4.3), when trained on noise-free theoretical shear correlation functions for the 26 nodes, achieves per cent level accuracy in its predictions across most of the parameter space relevant to the constraining power of present lensing surveys. When trained instead on the simulations, including the noise present across the realisations and inaccuracies in the HALOFIT predictions used for comparison, the measured accuracy is estimated at the level of $\lesssim 5\%$ for this statistic, i.e., still competitive for concurrent lensing analyses (see Figures 4.3-4.5).

After performing the necessary validation tests for the accuracy of the *cosmo*-SLICS emulator, I employ these simulations for their primary purpose - determining the cosmological dependence of novel lensing statistics. To this end, I subject the *cosmo*-SLICS suite to the clipping transforms explored in Chapter 3, subsequently measuring the clipped shear correlation functions presented in Figure 4.6. I demonstrate that emulation performance for these statistics is comparable to that of their unclipped counterparts (Figure 4.8). Finally, I employ

the *cosmo*-SLICS emulator in constraining the cosmology of mock data from an LSST-sized survey, producing clipped and unclipped ξ_+ model predictions at each step in an MCMC sampling of the likelihood surface. I find that the combination of these probes yields marginalised constraints on $\Omega_\mathrm{m}, S_8, h$ and w_0, displayed in Figure 4.9, which are 18%-26% tighter than the conventional, unclipped analysis alone. I also find that the statistical uncertainties in the emulated clipped and unclipped predictions, are subdominant to those of the next-generation surveys, for all of the parameter space relevant to the next generation of lensing surveys. However, the systematic bias of $\lesssim 5\%$, quantified in Section 4.3.2, necessitates additional simulation nodes to offset and to facilitate use of these resources with future lensing data.

There are a considerable number of applications of the *cosmo*-SLICS suite and emulator combination. For example, future work will include combined clipped and unclipped cosmic shear constraints on the new release of 1000 deg^2 of deep multi-band imaging from KiDS. The simulations and emulator are also currently being implemented as part of a project exploring the constraining power of peak statistics with this data set (Martinet et al. in prep.).

These resources may also prove invaluable to future projects featuring alternative estimators such as higher-order statistics (Fu et al., 2014), Minkowski functionals (Petri et al., 2015) and density-split lensing (Gruen et al., 2018; Brouwer et al., 2018). Indeed, a highly impactful project advocated by this author, would be a comprehensive investigation of the *combined* constraining power of these various probes, such that the optimal 'weapon of choice' for use with the next generation of lensing data is identified. I hope that these suggestions for applications of the publicly available emulator[9] and *cosmo*-SLICS suite, prove to be but the tip of the iceberg in terms of their usefulness to the cosmology community.

[9] `https://github.com/benjamingiblin/GPR_Emulator`

5

Emulating the Non-Linear Matter Power Spectrum for Arbitrary Cosmologies

In this Chapter, I introduce an emulator approach to predict the non-linear matter power spectrum for broad classes of beyond-ΛCDM cosmologies, using only a suite of ΛCDM N-body simulations. By including a range of suitably modified initial conditions in the simulations, and rescaling the resulting emulator predictions with analytical 'halo model reactions', accurate non-linear matter power spectra for general extensions to the standard ΛCDM model can be calculated. I optimise the emulator design by substituting the simulation suite with non-linear predictions from the standard HALOFIT tool. I review the performance of the emulator for artificially generated departures from the standard cosmology as well as for theoretically motivated models, such as $f(R)$ gravity and massive neutrinos. For the majority of cosmologies I have tested, the emulator can reproduce the matter power spectrum with errors $\lesssim$1% deep into the highly non-linear regime. This work demonstrates that with a well-designed suite of ΛCDM simulations, extensions to the standard cosmological model can be tested in the non-linear regime without any reliance on expensive beyond-ΛCDM simulations. This Chapter in its entirety has been submitted to MNRAS in Giblin et al. (2019),

and is currently under review.

5.1 Introduction

Large cosmological data sets have ushered in the era of "precision cosmology", in which parameters describing the ΛCDM model are known with uncertainties of only a few per cent. This increased precision has recently brought to light some level of discordance between early- and late-time cosmological probes. In particular, all three major weak lensing surveys (Hildebrandt et al., 2017; Hikage et al., 2018; Troxel et al., 2017) infer lower values of $\sigma_8\sqrt{\Omega_\mathrm{m}}$ compared to those obtained from observations of the temperature and polarisation of the cosmic microwave background (CMB; Planck Collaboration et al., 2018). Significant tensions have also arisen between the value of the Hubble constant derived from the CMB and its direct local measurement via the distance-redshift relation calibrated with Cepheid variables and Type Ia supernovae (Riess et al., 2016, 2018a). If not the product of an unlikely statistical fluctuation, then these discrepancies may be ascribed either to unaccounted systematic uncertainties in the analyses or to an incompleteness of the concordance cosmology (see, e.g., Verde et al., 2013; Joudaki et al., 2017; Mörtsell & Dhawan, 2018; DES Collaboration et al., 2018).

In the event that the data prefer a revision of the standard cosmological model, accurate predictions of large-scale structure statistics for alternative cosmologies to ΛCDM become very valuable. However, we currently lack a method to predict these observables for general extensions to ΛCDM and over the range of scales relevant for ongoing and future galaxy surveys (Laureijs et al., 2011; LSST Dark Energy Science Collaboration, 2012; Hildebrandt et al., 2017; DES Collaboration et al., 2017; Taylor et al., 2018). This means we cannot take advantage of the wealth of information contained in the non-linear scales without resorting to computationally expensive numerical simulations.

It is however infeasible to simulate all possible combinations of cosmological parameters, even within the ΛCDM paradigm. For current applications, fitting functions calibrated on a small number of N-body simulations are good enough to obtain unbiased constraints (Smith et al., 2003; Takahashi et al., 2012; Mead et al., 2015). On the other hand, future all-sky surveys will place more stringent requirements on our theoretical predictions, which makes emulator techniques

a preferable strategy as far as accuracy is concerned. As demonstrated in the previous Chapter, the basic idea is to interpolate between the output of ∼10–100 N-body simulations of cosmologies associated with carefully-chosen input parameters (Habib et al., 2007; Schneider et al., 2008; Heitmann et al., 2009, 2016; Lawrence et al., 2017; Euclid Collaboration et al., 2018; Winther et al., 2019; Harnois-Deraps et al., 2019). Previous implementations of this methodology have been applied either to flat, massless neutrinos (i.e. vanilla) ΛCDM, or to a few specific extensions, that is models with massive neutrinos, evolving dark energy with a linearly parameterised equation of state or $f(R)$ gravity. To date, owing to the impracticalities of numerical simulations, there exists no emulator capable of making non-linear predictions accurate at the level of ∼1% across a broad range of beyond-ΛCDM cosmologies. This severely limits our capacity for constraining these alternative models with data probing the growth of structure on cosmic scales.

In this Chapter, I complement the work of Cataneo et al. (2018, referred to as 'C18' henceforth), who formulated an efficient method to compute the non-linear matter power spectrum in arbitrary extensions to the standard cosmology with the desired per cent accuracy. The C18 strategy consists of rescaling the non-linear power spectrum of a *tailored* ΛCDM-*like* model – the *pseudo* cosmology – by an analytical function encapsulating the non-linear physics beyond the vanilla cosmology – the *halo model reaction*. By construction the pseudo cosmology shares the linear power spectrum of the *real* beyond-ΛCDM cosmology.

For ΛCDM extensions exhibiting a scale-independent linear growth of structure, the corresponding pseudo non-linear power spectrum can be obtained from a ΛCDM emulator by matching the amplitude of mass fluctuations, σ_8, to that of the real cosmology at the redshift of interest. In general, however, structures in alternative cosmological scenarios grow at different rates on different scales, which makes traditional emulators unfit for this task. Here, I present a new Gaussian process emulator designed to compute the pseudo non-linear matter power spectrum of arbitrary cosmologies, including both scale-independent and scale-dependent linear growth. I show that a suite of ∼1000 ΛCDM-*like* simulations is enough to predict with per cent accuracy the pseudo non-linear power spectrum of popular cosmologies which have scale-dependent linear growth, such as $f(R)$ gravity and massive neutrino models. I also obtain similar performance for synthetic cosmological models created by rescaling the linear ΛCDM power spectra by smooth arbitrary functions. The combination of the C18 halo

model reactions with the emulator developed in this work will hence provide accurate predictions for the non-linear matter power spectrum in a broad class of cosmological models, including modified gravity, dark energy and massive neutrinos. There is also the potential to use this technique to model the impact of baryonic feedback.

This Chapter is organised as follows. In Section 5.2 I examine the expected range of deviations in the linear matter power spectrum of viable extensions to the ΛCDM cosmology, and outline the C18 reaction framework for modelling the non-linear power spectrum. In Section 5.3 I detail the emulator methodology for arbitrary input linear power spectra. I also describe the construction of my suite of surrogate numerical simulations used as a training set for the emulator, which I design to encompass a reasonable range of ΛCDM extensions. In Section 5.4 I present the performance of my emulator and determine the minimum number of simulations needed to reach the accuracy required by the high-quality data from the next generation of all-sky imaging surveys. Finally, I conclude and discuss the impact of my findings in Section 5.5.

5.2 Beyond vanilla ΛCDM

The late-time phenomenology of the concordance model – a globally flat Universe with a matter-energy content dominated by cold dark matter, baryonic matter and the cosmological constant – is entirely described by five *base* parameters, $\boldsymbol{\pi}_\Lambda \equiv \{\Omega_\mathrm{b}h^2, \Omega_\mathrm{m}h^2, h, n_\mathrm{s}, \ln(10^{10}A_\mathrm{s})\}$, where Ω_b is the current fraction of energy-density in baryonic matter, Ω_m is the present-day total matter energy-density fraction, h is the dimensionless Hubble constant, and n_s and A_s are, respectively, the slope and amplitude of the primordial scalar power spectrum. Under the assumption of the standard cosmology, measurements of the CMB constrain the radiation content (i.e. photons and three massless neutrino species) to better than one part in a thousand (Fixsen, 2009).

Changes to the dark sector and to the law of gravity have profound implications for the formation of structures in the Universe over a wide range of scales. Departures from the standard cosmology, therefore, leave distinctive features on the statistical quantities measuring the clustering of matter, such as the two-point correlation function, or its Fourier transform, the power spectrum $P(k)$. The landscape of ΛCDM extensions is exceptionally vast, and in this

work I only consider a family of $f(R)$ gravity theories (Hu & Sawicki, 2007) and massive neutrino cosmologies (Lesgourgues & Pastor, 2006), both with a ΛCDM background. However, with little or no modification this methodology (see Section 5.3) can be applied to a much broader class of models, including non-standard dark matter candidates (Marsh & Silk, 2014; Schneider, 2015; Cyr-Racine et al., 2016; Marsh, 2016; Poulin et al., 2016; Hložek et al., 2017; Dakin et al., 2019; Thomas et al., 2019), extra relativistic degrees of freedom (Baumann et al., 2018), and Horndeski theories (Zumalacárregui et al., 2017).

What makes models like $f(R)$ gravity and massive neutrino cosmologies interesting is their scale-dependent linear growth of structure. More specifically, $f(R)$ theories enhance the growth on scales comparable to, or smaller than, the Compton wavelength, which at $z = 0$ reads (Hu & Sawicki, 2007)

$$\lambda_{\rm C0} \approx 30\sqrt{\frac{(n+1)}{4-3\Omega_{\rm m}}\frac{|f_{R0}|}{10^{-4}}} \quad h^{-1}{\rm Mpc}, \tag{5.1}$$

where f_{R0} defines the extent of the departure from General Relativity, recovered for $f_{R0} = 0$. I set $n = 1$ and use $|f_{R0}| = 10^{-4}, 10^{-5}, 10^{-6}$ (referred to as F4, F5 and F6 in the following), which brackets the range of values giving rise to interesting cosmological behaviours, with F4 already strongly disfavoured by current data (Terukina et al., 2014; Lombriser, 2014; Cataneo et al., 2015; Liu et al., 2016; Alam et al., 2016).

On the other hand, the high thermal velocities of massive neutrinos prevent them from clustering on scales smaller than the free-streaming length, $\lambda_{\rm FS}$. This results in the suppression of the growth of structure on scales smaller than the maximum free-streaming length, which is defined at the time of the non-relativistic transition (after recombination for $m_\nu \lesssim 0.5\,{\rm eV}$), and is given by (Lesgourgues & Pastor, 2006)

$$\lambda_{\rm FS}^{\rm max} \approx 350\sqrt{\frac{1}{\Omega_{\rm m}}\frac{1\,{\rm eV}}{m_\nu}} \quad h^{-1}{\rm Mpc}. \tag{5.2}$$

For simplicity, I will consider cosmologies including two massless neutrinos and one massive neutrino with $m_\nu = 0.05, 0.1, 0.2, 0.4\,{\rm eV}$, where the smallest value is close to the minimum sum of neutrino masses consistent with neutrino oscillation experiments (Forero et al., 2014), and the largest value is outside the 95% confidence region found by Planck Collaboration et al. (2018). In addition, for

Table 5.1 *The massive neutrino, $f(R)$ gravity and hybrid models used in this work to design and test my emulator scheme, each labelled according to the sum of neutrino masses ($\sum m_\nu$) and/or the strength of the deviation from standard gravity ($|f_{R0}|$). I consider two base ΛCDM cosmologies: (A) $\Omega_\mathrm{b}h^2 = 0.0225, \Omega_\mathrm{m}h^2 = 0.1382, h = 0.6898, n_\mathrm{s} = 0.969, \ln(10^{10}A_\mathrm{s}) = 3.195$; and (B) $\Omega_\mathrm{b}h^2 = 0.0173, \Omega_\mathrm{m}h^2 = 0.0864, h = 0.6, n_\mathrm{s} = 0.969, \ln(10^{10}A_\mathrm{s}) = 3.807$. Note that (B) sits several standard deviations away from the Planck Collaboration et al. (2018) best fit cosmology, thus working as stress test for my method.*

model	$\sum m_\nu$	$\lvert f_{R0}\rvert$	base parameters
MNU_0.05	0.05 eV	–	(A)
MNU_0.1	0.1 eV	–	(A)
MNU_0.2	0.2 eV	–	(A)
MNU_0.4	0.4 eV	–	(A)
xMNU_0.05	0.05 eV	–	(B)
xMNU_0.1	0.1 eV	–	(B)
xMNU_0.2	0.2 eV	–	(B)
xMNU_0.4	0.4 eV	–	(B)
F4	–	10^{-4}	(A)
F5	–	10^{-5}	(A)
F6	–	10^{-6}	(A)
F4-MNU_0.4	0.4 eV	10^{-4}	(A)
F5-MNU_0.2	0.2 eV	10^{-5}	(A)
F6-MNU_0.1	0.1 eV	10^{-6}	(A)

all these extensions I keep the base cosmological parameters fixed to their ΛCDM values, and compensate for the presence of massive neutrinos by reducing the cold dark matter density as $\Omega_\mathrm{c} \to \Omega_\mathrm{m} - \Omega_\mathrm{b} - \Omega_\nu$, where

$$\Omega_\nu h^2 = \frac{\sum m_\nu}{93.14\,\mathrm{eV}}. \tag{5.3}$$

Table 5.1 lists all models, collectively referred to as *physical models,* used to test my emulation scheme described in Section 5.3, including hybrid cosmologies with massive neutrinos and modified gravity. For these extensions, the linear power spectrum deviations from the standard predictions are obtained with MGCAMB[1] (Zhao et al., 2009; Hojjati et al., 2011). Figure 5.1 shows a $z = 0$ example of beyond-ΛCDM to ΛCDM linear matter power spectrum ratios.

[1] `http://aliojjati.github.io/MGCAMB/mgcamb.html`

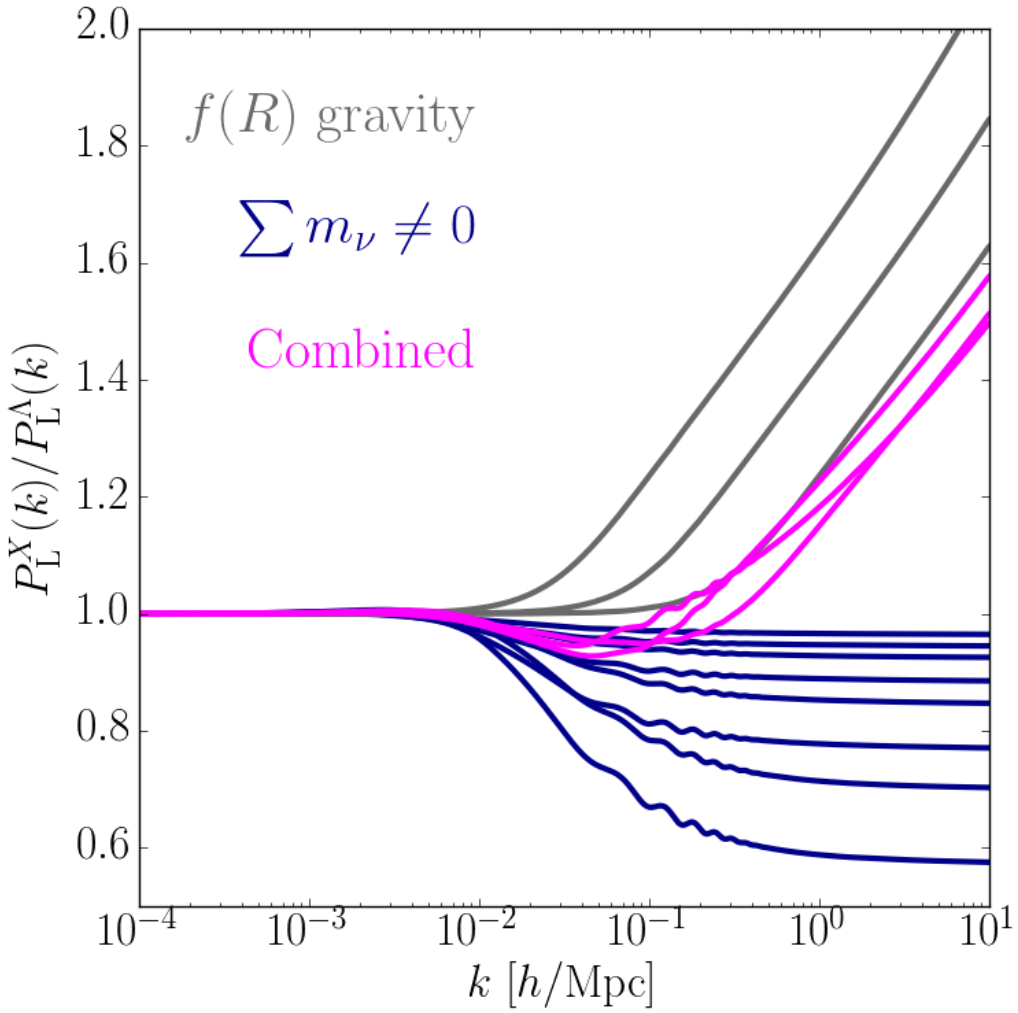

Figure 5.1 *Beyond-ΛCDM ($P_{\rm L}^X$) to ΛCDM ($P_{\rm L}^\Lambda$) linear matter power spectrum ratios for the physical models in Table 5.1 at $z = 0$. Pure $f(R)$ gravity (in grey) and massive neutrino cosmologies (in blue) alter the clustering of matter in similar but opposite ways. This degeneracy is clearly visible for the hybrid models (in magenta), where both extensions are active. These curves will be used as guidelines to design my emulator.*

5.2.1 The halo model reaction framework

By unlocking the information stored in the statistical distribution of matter on small scales we can substantially increase our sensitivity to physics beyond the vanilla cosmology, be it new particles/fluids, modifications to the theory of gravity, or astrophysics (see, e.g. Huterer et al., 2015; Copeland et al., 2018; Nori & Baldi, 2018; Heymans & Zhao, 2018; Schneider et al., 2019). Not surprisingly, then, the non-linear matter power spectrum is found at the core of all cosmological analyses hinged on galaxy survey data (Parkinson et al., 2012; Kitching et al., 2014; Gil-Marín et al., 2016; Hildebrandt et al., 2017; van Uitert et al., 2018; Joudaki et al., 2018; DES Collaboration et al., 2017; Gil-Marín et al., 2018; Hikage et al., 2018). To take full advantage of the exquisite observations from forthcoming experiments (Laureijs et al., 2011; LSST Dark Energy Science Collaboration, 2012; Levi et al., 2013), theoretical predictions must reach per cent level accuracy deep in the non-linear regime of structure formation (Taylor et al., 2018). At present, however, no method is at the same time computationally efficient, accurate and flexible enough to be employed in future analyses aimed at constraining physics beyond the standard paradigm (Lesgourgues et al., 2009; Bird et al., 2012; Heitmann et al., 2014; Brax & Valageas, 2013; Zhao, 2014; Blas et al., 2014; Massara et al., 2014; Levi & Vlah, 2016; Lawrence et al., 2017; Aviles & Cervantes-Cota, 2017; Senatore & Zaldarriaga, 2017; Bose et al., 2018; Euclid Collaboration et al., 2018; Heisenberg & Bartelmann, 2019; Winther et al., 2019).

The halo model reaction framework first proposed by C18, and based on the work of Mead (2017), stands out as a promising, practical solution to the long-standing problem of predicting the non-linear matter power spectrum with the required accuracy and speed, while also being readily applicable to a variety of cosmological scenarios. Within this approach, the non-linear power spectrum of a given *real* cosmology takes the form

$$P_{\rm NL}^{\rm real}(k,z) = \mathcal{R}(k,z) \times P_{\rm NL}^{\rm pseudo}(k,z), \tag{5.4}$$

where the *pseudo* cosmology is defined as a ΛCDM cosmology sharing the same linear matter power spectrum of the *real* cosmology, that is

$$P_{\rm L}^{\rm pseudo}(k,z) = P_{\rm L}^{\rm real}(k,z). \tag{5.5}$$

The *reaction*,

$$\mathcal{R}(k,z) \equiv \left.\frac{P_{\rm NL}^{\rm real}(k,z)}{P_{\rm NL}^{\rm pseudo}(k,z)}\right|_{\rm HM}, \tag{5.6}$$

captures the non-linearities sourced by late-time physics beyond ΛCDM, and is predicted with the halo model (HM) and perturbation theory (see Cataneo et al., 2018, for details). In taking the ratio of the two halo model predictions, Mead (2017) showed that the impact of the known inaccuracies in the individual power spectra is significantly reduced. On linear scales and for *real* ΛCDM cosmologies one has the trivial relation $\mathcal{R} = 1$.

Here, I formulate a method to provide predictions for $P_{\rm NL}^{\rm pseudo}$ that are more accurate than any semi-analytical prescription based on the halo model. It is this accurate estimate that is then used to calibrate the reaction, $\mathcal{R}$, in equation 5.4 in order to derive the non-linear matter power spectrum of the target beyond-ΛCDM cosmology, $P_{\rm NL}^{\rm real}$. To this end, I develop a novel emulation scheme where in addition to the base ΛCDM parameters I use the shape of the linear matter power spectrum as input. This work is intended as a first feasibility study, and as such I approximate the output of otherwise expensive *pseudo* N-body simulations with the HALOFIT fitting functions (Takahashi et al., 2012) implemented in a modified version of the public Einstein-Boltzmann solver CAMB[2] (Lewis et al., 2000). In practice, for a model X, CAMB reads in both $\boldsymbol{\pi}_\Lambda$ and the externally generated $P_{\rm L}^X/P_{\rm L}^\Lambda$ ratio, evaluates $P_{\rm L}^{\rm pseudo}$ and gives it to HALOFIT, which finally provides the desired non-linear quantity. The grey curves in the upper panel of Figure 5.2 show the *pseudo* non-linear power spectra for the physical models in Table 5.1, where I used the set (A) of base ΛCDM parameters also for the xMNU cosmologies. Note that despite HALOFIT predictions being approximate, this strategy is theoretically consistent, in that *pseudo* cosmologies are simply ΛCDM cosmologies with non-standard initial conditions. My approach enables a rapid design, construction and performance assessment of the emulator (Heitmann et al., 2009; Euclid Collaboration et al., 2018).

5.3 Methodology

In this Section, I very briefly summarise the basics of Gaussian process emulation of cosmological statistics. I also detail my pipeline, which starts from the linear

[2] `https://github.com/cmbant/CAMB`

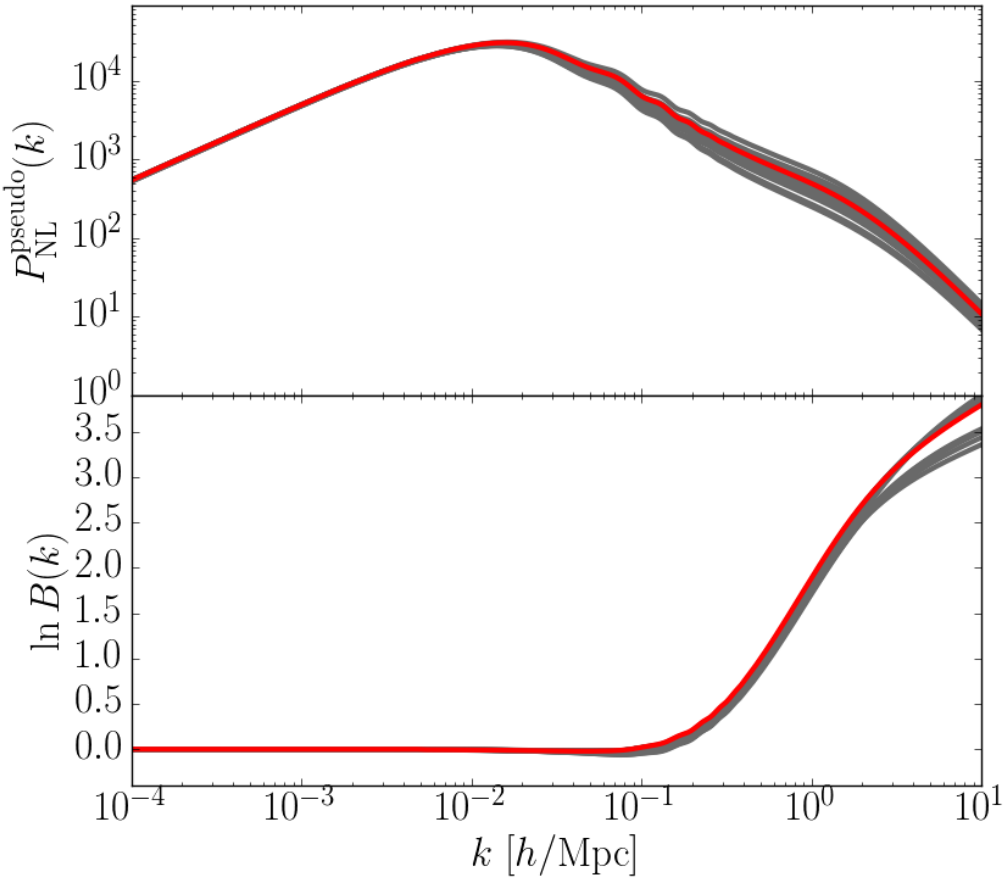

Figure 5.2 *Upper panel: in grey, the $z = 0$ pseudo non-linear matter power spectra for the extensions to ΛCDM listed in Table 5.1, computed with* HALOFIT *as discussed in Section 5.2.1. Lower panel: the natural logarithm of the boost factor obtained by taking the ratio of pseudo non-linear and linear matter power spectra (see equation 5.8). For reference, the base ΛCDM cosmology (A), given in Table 5.1, is shown in red. The monotonicity and narrow dynamic range of the $\ln B$ relative to the $P_{\rm NL}^{\rm pseudo}$ help to improve the emulator's performance significantly.*

matter power spectrum of an arbitrary cosmology and maps this onto a point in the multi-dimensional parameter space over which the emulator performs the interpolation. This gives the emulator high versatility, in that it will be able to make predictions for a broad range of non-standard cosmologies absent from the training set.

5.3.1 Emulation strategy

The mathematics behind Gaussian process (GP) regression emulators were covered in detail in Section 4.1. Here, I very briefly summarise its key points, which I once again implement using the publicly available code SCIKIT-LEARN[3] (Pedregosa et al., 2011).

GP regression is a non-parametric Bayesian machine learning algorithm for constraining the distribution of functions, $f(\boldsymbol{\pi})$, which are consistent with the mapping between the training set input parameters, or 'nodes' $\boldsymbol{\pi} \in \mathbb{R}^d$, and the output, $\boldsymbol{y}$, via

$$y(\boldsymbol{\pi}) = f(\boldsymbol{\pi}) + \epsilon_{\rm n}(\boldsymbol{\pi})\,, \tag{5.7}$$

where $\epsilon_{\rm n}(\boldsymbol{\pi})$ is a noise term sampled from a mean-zero Gaussian distribution with a standard deviation given by the error on the training set observable, $y(\boldsymbol{\pi})$[4]. The prediction, y^*, corresponding to an arbitrary coordinate $\boldsymbol{\pi}^*$, is then sampled from a generalisation of a Gaussian posterior probability distribution over the range of consistent functions, as expressed in equation 4.12. In other words, the GP emulator interpolates the observable from the input coordinates of the training set to the trial coordinates across a d-dimensional parameter space.

A key ingredient of the posterior is the Gaussian prior distribution of functions deemed to reasonably map between input and output. The prior is determined by a mean, conventionally taken to be zero, and a covariance function, known as the 'kernel'. The kernel can take various functional forms, each described by a vector of hyperparameters governing the kernel's behaviour. Following Rasmussen & Williams (2006); Habib et al. (2007); Heitmann et al. (2009), in this work I

[3] `https://scikit-learn.org`

[4] In this work I consider HALOFIT training predictions as a proxy for hypothetical numerical simulations with negligible error. Nevertheless, to prevent unwanted numerical instabilities I use the arbitrarily small constant default $\epsilon_{\rm n}$ in SCIKIT-LEARN for all $y(\boldsymbol{\pi})$.

adopt the squared-exponential form given in equation 4.7, the hyperparameters of which, $\boldsymbol{h} = \{A, p_1, \cdots, p_d\}$, are a kernel amplitude and correlation lengths for each dimension of the emulation. The emulator is trained by finding values for the hyperparameters such that a distribution of functions which are optimally consistent with all realisations in the training set can be defined. In this work, I again fit for these using the method built-in to SCIKIT-LEARN, which employs a gradient ascent optimisation of the marginal likelihood expressed in equation 4.14.

The accuracy of the emulator is also sensitive to the particular observable being emulated. The aim here is to predict the *pseudo* non-linear matter power spectrum discussed in Section 5.2.1, a quantity which has a dynamic range of many orders of magnitude and features a change in gradient sign between small and large scales, as shown in the upper panel of Figure 5.2 for the extensions to ΛCDM listed in Table 5.1. GP emulators on the other hand perform better at predicting smooth monotonic functions with a narrow dynamic range. Hence, following Euclid Collaboration et al. (2018), I consider the natural logarithm of the *boost factor*, $B(k, z)$, as my observable of interest, where

$$B(k, z) \equiv \frac{P_{\rm NL}^{\rm pseudo}(k, z)}{P_{\rm L}^{\rm pseudo}(k, z)}. \tag{5.8}$$

The lower panel of Figure 5.2 shows that the natural logarithm of the boost factor indeed matches the desired properties of a GP observable.

An essential step in the emulator design is data compression. For any particular cosmological model, the power spectrum is sampled both in spatial frequencies (wave numbers) and time (redshifts), for a total of $\sim 10^4$ support points. Training the emulator on each of these individual measurements requires prohibitively large computational resources. To overcome this issue, I perform a principal component analysis (PCA) following previous work (see, e.g, Habib et al., 2007; Schneider et al., 2008; Heitmann et al., 2016). For a given parameter vector, $\boldsymbol{\pi}$, and redshift, z, I apply the following linear decomposition

$$\ln B(k, z; \boldsymbol{\pi}) = \mu(k, z) + \sum_{j=1}^{n_\psi} \psi_j(k, z) w_j(\boldsymbol{\pi}) + \epsilon\,, \tag{5.9}$$

where μ and $\{\psi_1, \cdots, \psi_{n_\psi}\}$ are, respectively, the mean and basis functions of

the training set of boosts at z, and the weights w_j are the projections of the mean-subtracted observable onto ψ_i. The error term ϵ has two contributions: one component, arising from the negligible noise on the HALOFIT training predictions, given by $\epsilon_{\rm n}(\boldsymbol{\pi})$ in equation 5.7; and a second component coming from missing relevant basis functions. I find that $n_\psi = 12$ is sufficient to satisfactorily reconstruct the training set, and including more basis functions has only a marginal effect on the emulator accuracy. Ultimately, for a given $\boldsymbol{\pi}^*$ my emulator infers the boost factor by predicting the weights, $w_j(\boldsymbol{\pi}^*)$, conditioned on the training set $w_j(\boldsymbol{\pi})$.

5.3.2 Model-independent parameterisation of the pseudo cosmologies

Extensions to ΛCDM are described by an arbitrary number of additional physical parameters. Consider, for example, wCDM, $f(R)$ gravity and massive neutrino cosmologies: matter clustering in all these models depends on five standard cosmological parameters, e.g. $\boldsymbol{\pi}_\Lambda = \{\Omega_{\rm b}h^2, \Omega_{\rm m}h^2, h, n_s, \ln(10^{10}A_s)\}$, along with one or more parameters describing model-specific features, $\boldsymbol{\pi}_{\rm X} = \{f_{R0}, \sum m_\nu, \cdots\}$. To construct an emulator capable of predicting generic pseudo non-linear matter power spectra, I opt for a model-independent parameterisation mapping an arbitrary model to a certain coordinate, $\boldsymbol{\pi}^* = \{\boldsymbol{\pi}_\Lambda, \Delta\boldsymbol{\alpha}\}$, in my emulation parameter space, where $\Delta\boldsymbol{\alpha}$ is a vector of effective parameters that quantifies smooth deviations from the linear matter power spectrum of the standard cosmology. Thus, for $\Delta\boldsymbol{\alpha} = 0$ the late-time physics of the vanilla ΛCDM model is recovered. For this, I model the 'shape', namely the ratio of the *linear* pseudo matter power spectrum to that of the ΛCDM cosmology sharing the same standard parameters, as

$$\begin{aligned} S(k, z; \boldsymbol{\pi}_\Lambda, \Delta\boldsymbol{\alpha}) &\equiv \frac{P_{\rm L}^{\rm pseudo}(k, z; \boldsymbol{\pi}_\Lambda, \boldsymbol{\pi}_{\rm X})}{P_{\rm L}^{\Lambda}(k, z; \boldsymbol{\pi}_\Lambda)} \\ &\approx 1 + \sum_{i=1}^{n_\Phi} \Phi_i(k, z)\Delta\alpha_i \,, \end{aligned} \tag{5.10}$$

where $\{\Phi_1, \cdots, \Phi_{n_\Phi}\}$ are the principal components of a training set, $\mathbb{T}$, consisting

of smooth curves capturing a range of 'reasonable' deviations from ΛCDM[5].

Generation of the basis set

I assemble a set of orthogonal basis functions to reconstruct a broad class of scale-dependent linear departures from the ΛCDM power spectrum. For this purpose, I generate random curves in Fourier space over the range $10^{-4} < k\,\mathrm{Mpc}/h < 10$, with the constraint that all curves must converge to unity on large enough scales. This last requirement is motivated by the physics of well-known ΛCDM extensions, such as Generalised Brans-Dicke theories (De Felice & Tsujikawa, 2010; Park et al., 2010; Hinterbichler et al., 2011; Pogosian & Silvestri, 2016; Quiros et al., 2016; Joyce et al., 2016) and massive neutrino cosmologies (Lesgourgues & Pastor, 2006). Both these models possess a characteristic scale, $\bar{k}$, such that on scales $k \ll \bar{k}$ the linear matter clustering matches that of the vanilla ΛCDM cosmology. For $k \gg \bar{k}$ the growth of structure is either suppressed ($S < 1$) or enhanced ($S > 1$). Note that models with $\bar{k} \to 0$ exhibit in practice a scale-independent linear growth (see, e.g., Pogosian & Silvestri, 2016), and a simple rescaling of the amplitude of the vanilla ΛCDM power spectrum is sufficient to produce the corresponding pseudo power spectrum. Viable models in these two classes of theories above have $\bar{k} \gtrsim 10^{-3}\,h/\mathrm{Mpc}$ (Brax et al., 2012; Wang et al., 2012; Joyce et al., 2015; Planck Collaboration et al., 2018). I therefore adopt the value $\tilde{k} = 10^{-3}\,h/\mathrm{Mpc}$

[5]Equation 5.10 can be understood in terms of a principal component reconstruction based on the training set $\mathbb{T}$, that is

$$S(k_m) \approx \nu(k_m) + \sum_{i=1}^{n_\Phi} \Phi_i(k_m)\alpha_i\,, \tag{5.11}$$

where for simplicity I dropped the dependence on redshift and cosmology, and made explicit that I sample the shape S at $k_m \in \{k_1, \cdots, k_{n_k}\}$. The function ν is the mean across $\mathbb{T}$, and in general one expects $\nu \neq 1$. Then, for any ΛCDM cosmology we must have

$$1 - \nu(k_m) \approx \sum_{i=1}^{n_\Phi} \Phi_i(k_m)\alpha_i^\Lambda\,, \tag{5.12}$$

with α_i^Λ being the weights that exactly compensate the departure of the mean from unity, and upon using the orthonormality of the basis functions they read

$$\alpha_i^\Lambda = \sum_{m=1}^{n_k} \Phi_i(k_m)[1 - \nu(k_m)]\,. \tag{5.13}$$

Finally, the combination of Eqs. (5.11) and (5.12) gives Eq. (5.10), with $\Delta\alpha_i \equiv \alpha_i - \alpha_i^\Lambda$. Note that for a given training set, $\mathbb{T}$, the ΛCDM weights need to be computed only once, and can subsequently be hard-coded in the emulator.

for the scale at which the shapes converge to unity, and allow for generous positive smooth deviations in the range $k \in [\tilde{k}, 10]\,h/\mathrm{Mpc}$. Moreover, since the basis functions derived from my generated shapes are sufficiently general to describe the power spectrum ratios at various redshifts, I drop the z-dependence, $\Phi_i(k, z) \to \Phi_i(k)$, and use one basis set for all redshifts.

In order to reduce the sensitivity of the emulator to the particular method used to generate the set of random shapes, I employ two different and independent generating strategies. The first of these methods, which I refer to as GPCurves, draws shapes from a 1-dimensional Gaussian process with a squared-exponential kernel of the form given in equation 4.7 (where $d = 1$), conditioned on a 'training set' of densely sampled points at $k < 10^{-3}\,h/\mathrm{Mpc}$, all with values equal to unity, helping to fix the random shapes to this value on large scales. Here, for the kernel amplitude, I use $A = 5$, which corresponds to the maximum value that the random shapes can have on scales $k > 10^{-3}\,h/\mathrm{Mpc}$. I tune the only length-scale parameter, p, so that the artificial shapes exhibit at most one stationary point for $k > 10^{-3}\,h/\mathrm{Mpc}$, thus resembling features of the physical shapes in Figure 5.1. There is no intrinsic restriction in the GP preventing generated shapes from taking unphysical negative values. I therefore generate an initial set of curves from which I sample 1000 viable positive shapes.

As a second method I use Smurves (Moews et al., 2018), a novel random curve generator. The functionality of the generator is illustrated by considering a particle travelling in Fourier space from low-to-high wave numbers. Before $k = 10^{-3}\,h/\mathrm{Mpc}$, the particle experiences no forces and hence travels undisturbed at a vertical coordinate of unity. After this point, forces randomly generated from a uniform distribution perturb the particle in the vertical direction, altering its trajectory. Also randomly generated are the locations of gradient sign changes. In this manner, I produce an additional 1000 smooth curves with a maximum of one stationary point and vertical intervals in the range [0, 8]. In order to facilitate extra flexibility in the deviations from ΛCDM, the parameters specifying the curve generation in Smurves are set purposefully to create a sample of curves which are slightly broader than GPCurves.

A comparison of 10 curves randomly selected from either samples are shown in different colours in the upper panel of Figure 5.3. The upper-middle panel shows the shapes corresponding to the physical models in Table 5.1 (orange) relative to the collective sample of 2000 random curves from GPCurves and Smurves (grey). The lower-middle panel shows the orthogonal basis functions, Φ_i, obtained

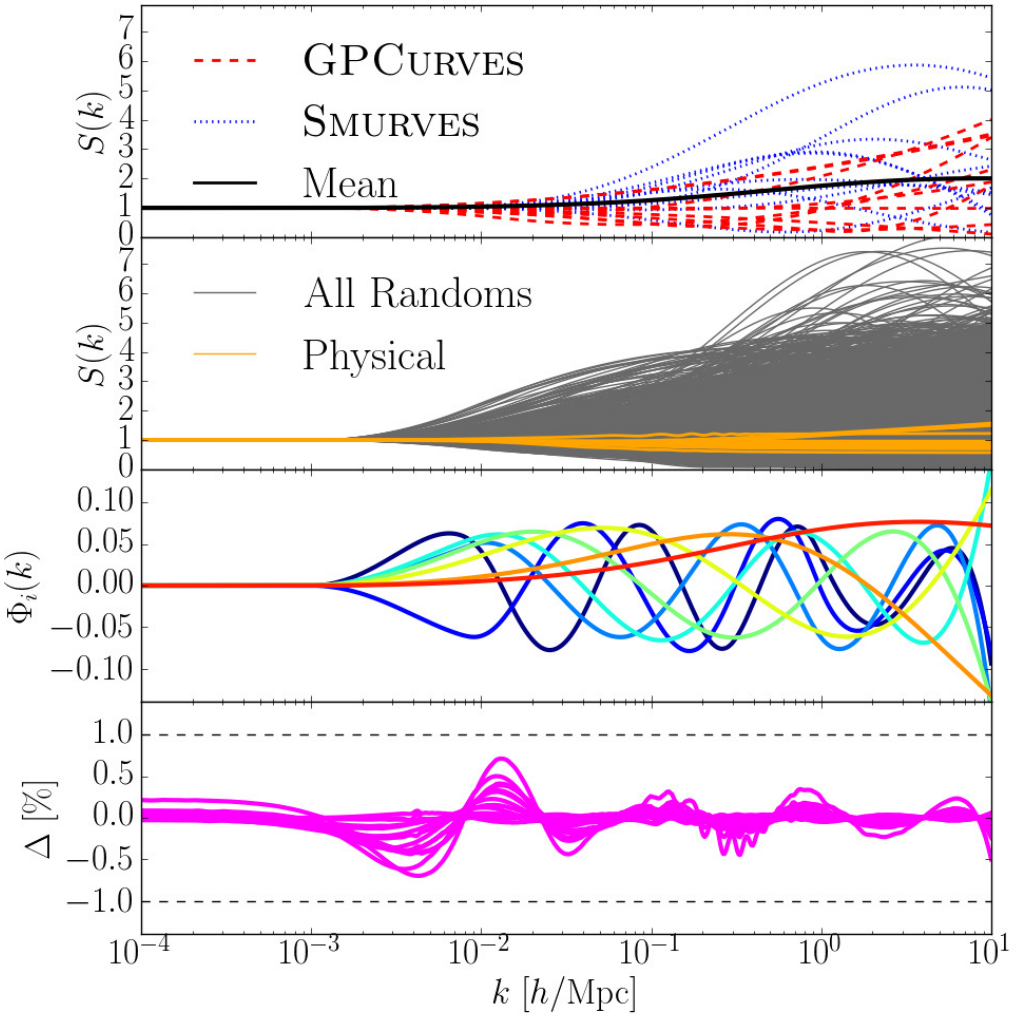

Figure 5.3 *Upper panel: A sub-sample of random shapes generated with the* GPCurves *(dashed red) and with the* Smurves *(dotted blue) methods. The mean across all 2000 random shapes is shown in solid black. Upper-middle panel: All 2000 random shapes from* GPCurves *and* Smurves *(grey) with the 14 shapes for the physical models from Figure 5.1 shown for reference (orange). Lower-middle panel: the 8 orthogonal basis functions, Φ_i, obtained via a PCA of the 2000 random shapes. The colour indicates the ranking of the basis function in the hierarchy of explaining the variance in the shapes, with red indicating the most important, and dark blue indicating the least. Lower panel: the accuracy, Δ, in reconstructing the physical model shapes, from Figure 5.1, using the basis functions, Φ_i.*

via a PCA of the 2000 curves. Note that the resulting basis set allows for a broader range of departures from ΛCDM than my test physical models.

Each basis function, Φ_i, appearing in the shape reconstruction (equation 5.10) contributes one extra dimension to the parameter space of the emulator, with the additional parameters corresponding to the PCA weights, $\Delta\alpha_i$. Therefore, in order for my emulator to be computationally efficient whilst still achieving accuracies of $\lesssim$1%, I keep the smallest number of basis functions that guarantees sub-per-cent reconstruction errors on all physical test shapes (see Figure 5.1). I find that $n_\Phi = 8$ basis functions are sufficient to reconstruct $f(R)$ gravity shapes with negligible errors. For cosmologies with massive neutrinos, however, changes to the BAO dynamics at early times produce rapid oscillations in their late-time power spectrum shapes. These features require one additional step in the reconstruction, but no extra parameters need be included in the emulation. Further details on this can be found in Appendix 5.A. The resultant shape reconstruction accuracy for all physical models is shown in the lower panel of Figure 5.3. Overall my emulation volume is therefore carved out of a 13-dimensional space: 5 ΛCDM parameters, $\boldsymbol{\pi}_\Lambda$, and 8 shape parameters, $\Delta\boldsymbol{\alpha}$.

5.3.3 Building the training and trial sets

Having established a model-independent description for arbitrary cosmological models, I now discuss how I generate training and trial sets for my pseudo matter power spectrum emulator. In this Section, I give details on the Latin hypercube (LH) space-filling strategies I adopt to efficiently sample my 13-dimensional space. I then define suitable ranges for the cosmological and shape parameters, $\{\boldsymbol{\pi}_\Lambda, \Delta\boldsymbol{\alpha}\}$, such that the emulated boost factors are representative of viable extensions to ΛCDM. The nodes of the LHs take values in the range $[0, 1]$, which I then properly rescale to obtain the "physical" coordinates $\boldsymbol{\pi}$ used to generate the training and trial sets.

Latin hypercube sampling

The use of LH-based distributions of points to evenly sample a parameter space was first introduced by McKay et al. (1979). The LH is a generalisation of the concept of a Latin square in which each row and column of a two-dimensional grid features exactly one sample, taking the form of a chess board with a number

of rooks that do not threaten each other. Owing to the desirable space-filling properties, LH sampling of the input parameters for simulations has become a standard practice (Sacks et al., 1989; Currin et al., 1991; Heitmann et al., 2006; Schneider et al., 2008; Agarwal et al., 2012; Liu et al., 2015; Garrison et al., 2018; Harnois-Deraps et al., 2019).

Assembling 13-dimensional LH designs therefore presents an ideal way to construct predictions to both train the emulator and test its accuracy for different regions of the emulation space. The spatial distribution of the trial cosmologies does not require optimisation. Therefore I sample 300 trial coordinates from a 13-dimensional LH generated with the pyDOE Python package[6]. The training nodes, however, impact significantly on the emulator accuracy, and sampling methods able to cover the high-dimensional volume as uniformly as possible are thus preferable. For this purpose, space-filling criteria based on the distance between samples are commonly used. One of the most prominent approaches is the Maximin LH design, which maximises the minimum distance between samples (Morris & Mitchell, 1995). While this approach leads to the sought-after space-filling properties in the bulk of the parameter space, most points cluster in the corners or at the boundaries of the lower-dimensional projection spaces. Since the matter power spectrum is most sensitive to subsets of cosmological and shape parameters, space-filling strategies also optimising projection space sampling would give an obvious advantage.

To this end, Joseph et al. (2015) introduced the maximum projection design (MaxPro), in which a weighted Euclidean distance, E, is minimised, resulting in the new criterion

$$\min_{\boldsymbol{D}} E(\boldsymbol{D}) = \left(\frac{1}{\binom{N}{2}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{1}{\prod_{l=1}^{d} (x_{il} - x_{jl})^2} \right)^{\frac{1}{d}}, \tag{5.14}$$

where $\boldsymbol{D} = \{\boldsymbol{x}_1, \ldots \boldsymbol{x}_N\}$ is an $N \times d$ design matrix containing N samples occupying the d-dimensional parameter space. x_{il} denotes the value of the i-th sample's coordinate in the l-th dimension. The MaxPro criterion maximises the bulk and projection space-filling properties, since for any dimension l, if $x_{il} = x_{jl}$ and $i \neq j$, $E(\boldsymbol{D}) = \infty$. This ensures that in the design minimising $E(\boldsymbol{D})$ no two points can be close to each other in any of the projections.

[6] `https://pythonhosted.org/pyDOE/`

The MaxPro design optimisation implemented in this work is based on the simulated annealing algorithm described in Morris & Mitchell (1995)[7]. This consists of selecting an initial configuration before searching for progressively better design choices by randomly perturbing the current design, keeping the new matrix should it reduce the cost function, $E(\boldsymbol{D})$, or else do so with a given probability. Finally this implementation arrives at a locally optimal MaxPro design by ascending the gradient given by

$$\frac{\partial E^d(\boldsymbol{D})}{\partial x_{rs}} = \frac{2}{\binom{N}{2}} \sum_{i \neq r} \frac{1}{(x_{is} - x_{rs})} \frac{1}{\prod_{l=1}^{d} (x_{il} - x_{rl})^2}. \tag{5.15}$$

To investigate the performance of my emulator I generate sets of training nodes with sizes $N = 100, 300, 500$ and 600, with the additional property of having an optimal space-filling pattern in each projection. Furthermore, to investigate the sensitivity of my results to the specific configuration of nodes in each training set, I produce 10 optimised designs of each size and measure the range of achieved accuracies across them.

Setting parameter ranges

I shall now map my dimensionless training and trial nodes belonging to the unit LHs onto my parameters, $\boldsymbol{\pi} \in \{\boldsymbol{\pi}_\Lambda, \Delta\boldsymbol{\alpha}\}$. To do this, I must identify appropriate ranges for these parameters. I begin with the ΛCDM parameters, and note that my emulator is intended for cosmological analyses employing high-quality data from the next-generation galaxy surveys. Hence, I set the boundaries on $\boldsymbol{\pi}_\Lambda$ to be roughly consistent with the 95% marginalised constraints from the combined TT, TE, EE + lowE analysis in Planck Collaboration et al. (2018), namely,

$$\begin{aligned} 0.0215 < \Omega_{\rm b} h^2 < 0.0235, \\ 0.12 < \Omega_{\rm m} h^2 < 0.155, \\ 0.6 < h < 0.8, \\ 0.9 < n_s < 1.05, \\ 2.92 < \log(10^{10} A_s) < 3.16. \end{aligned} \tag{5.16}$$

These ranges are illustrated in Figure 5.4, which shows the 2D projections of the

[7] `https://CRAN.R-project.org/package=MaxPro`

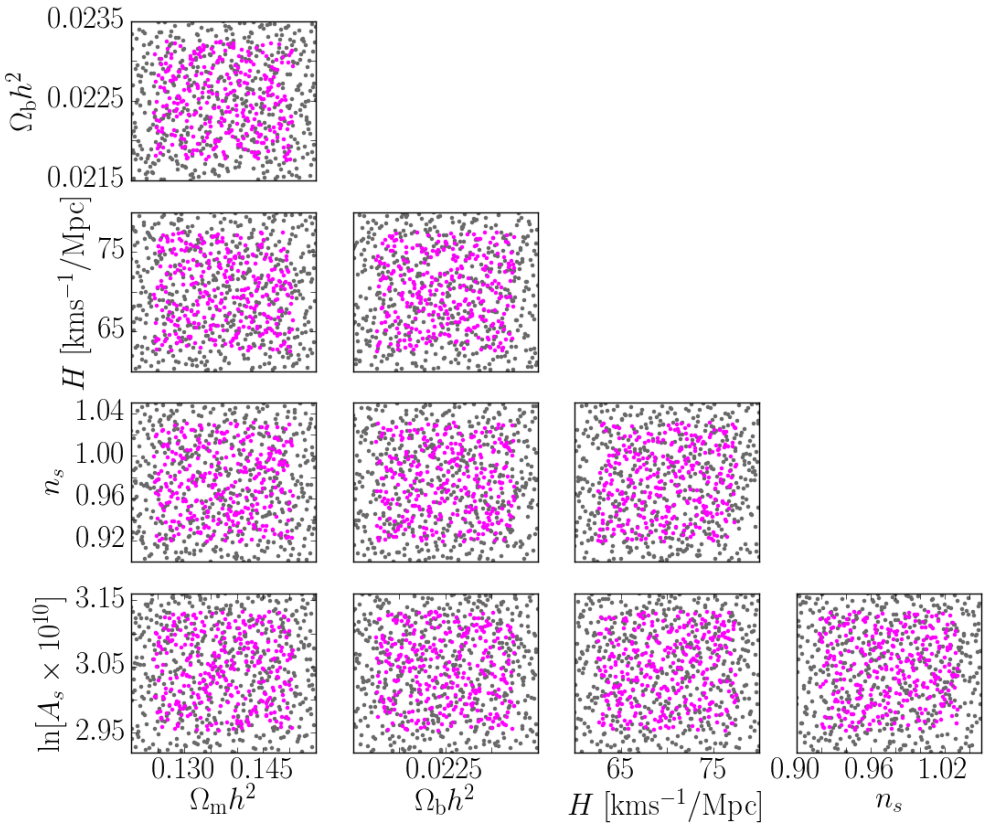

Figure 5.4 *The 2D projections of the 13-dimensional emulation space for the five ΛCDM parameters only. The grey points correspond to a 500-node training set obtained by maximising the distance between nodes in each projection, whereas the magenta points correspond to the trial coordinates, which are simply sampled from a standard Latin hypercube (as described earlier in this Section). The axes show the full allowed range of the training points. The trial coordinates are confined to a smaller hypercube with sides measuring 75% of the training range, which reduces the impact of boundary effects on the emulator accuracy.*

ΛCDM sub-space, with the grey points corresponding to one of the $N = 500$ MaxPro distributed training designs.

Turning to the shape parameters, $\Delta\boldsymbol{\alpha}$, I take inspiration from the physical models listed in Table 5.1, and simply set the range of values in my training set to be slightly broader than those defined by the $\Delta\boldsymbol{\alpha}$'s associated with $f(R)$ gravity, massive neutrino and combined cosmologies. The upper panel of Figure 5.5 shows the $\Delta\alpha_i$ for the physical models at $z = 0$ in orange, and those related to one of the designs with 500 optimally spaced training nodes in grey. The convergence of the $\Delta\alpha_i$ values towards zero as the PCA rank, i, increases, is an indication that I am using a sufficient number of basis functions to reconstruct my test models.

The lower panel of Figure 5.5 shows the shapes corresponding to the $\Delta\boldsymbol{\alpha}$ vectors shown in the upper panel, both for the 500-node training set (in grey) and for the physical models (in orange), reproduced from Figure 5.1. One can see that the training shapes cover a complex range of behaviours, whilst still encompassing the shapes of the physical models.

For the trial set, I opt for narrower parameter ranges, such that in each dimension the trial coordinates have values in the central 75% of the training set intervals. In doing this I test the emulator accuracy only in the bulk of the parameter space, so that my conclusions are less influenced by edge effects, where emulator accuracy degrades due to a shortage of nodes near the boundaries. This is visualised in Figure 5.4 by the distribution of trial nodes (in magenta). I also test the performance of my emulator for strictly ΛCDM cosmologies, where all shape parameters are set to zero. This ΛCDM trial ensemble is used to test whether the emulator is able to make accurate predictions for the standard model in the non-linear regime, even when ΛCDM cosmologies are completely absent from the training set.

Generating boost factors

With the training and trial nodes now assembled, the final task is to produce the corresponding linear as well as non-linear matter power spectra, and hence the logarithm of the boost factor, $B(k, z)$ (see equation 5.8), which the emulator is trained on and predicts. Conventionally, the non-linear statistics are measured from cosmological simulations, but for the purposes of this proof-of-concept study halofit can serve as a proxy for a suite of N-body simulations. Note that this

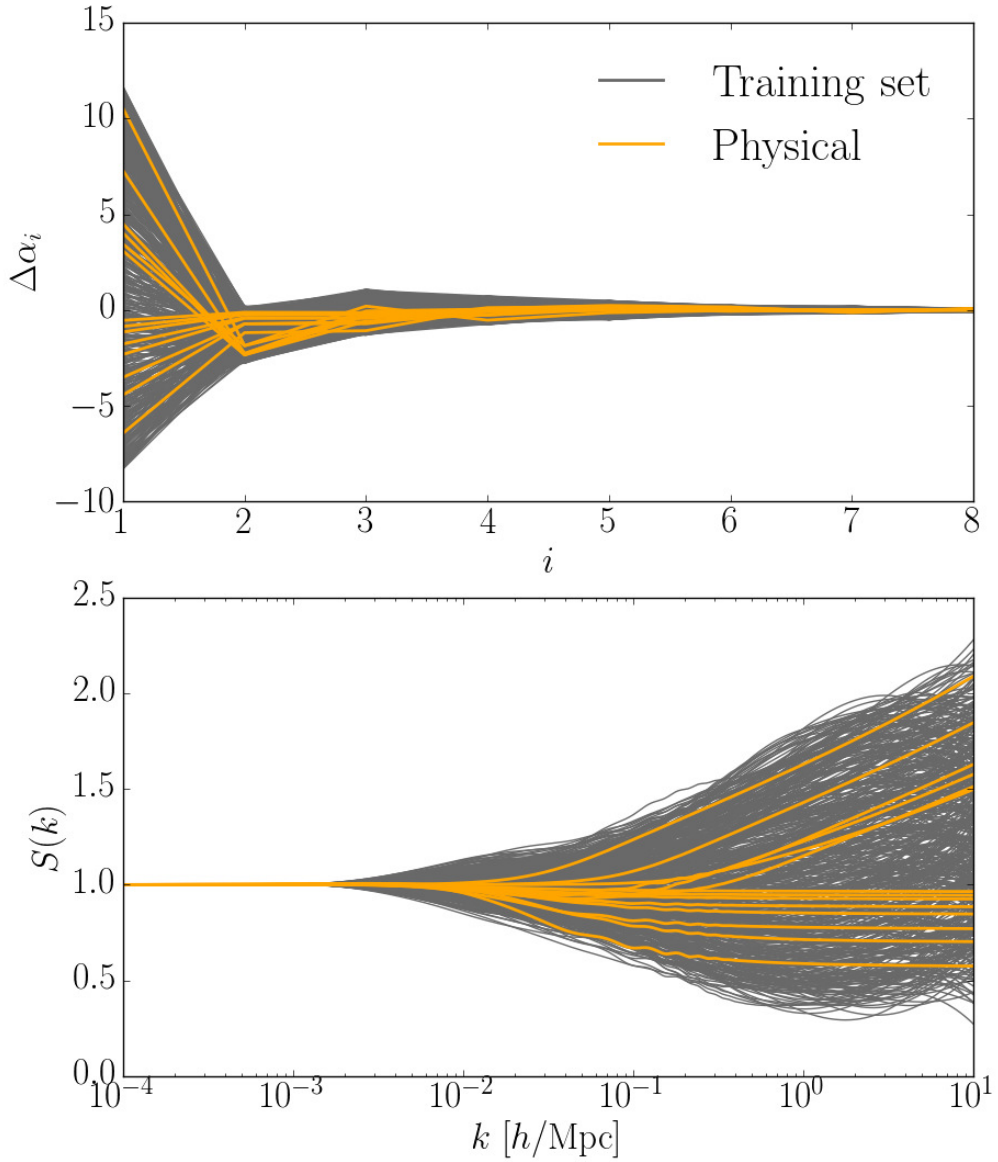

Figure 5.5 *Upper panel: the $\Delta\alpha_i$ values, as a function of PCA rank, i, which parameterise deviations from ΛCDM in the $z = 0$ pseudo linear matter power spectra, for a 500 node training set (grey) and physical models (orange). Lower panel: synthetic and physical shapes corresponding to the $\Delta\alpha_i$ weights above.*

approach is self-consistent as long as the growth of structure and background evolution of the 'simulated' cosmologies matches those of the ΛCDM model, which is indeed the case for the pseudo cosmologies. To estimate the surrogate $P_{\rm NL}^{\rm pseudo}$ I repeat the process described in Section 5.2.1 for the physical models: for a given pseudo cosmology, $\boldsymbol{\pi} = \{\boldsymbol{\pi}_\Lambda, \Delta\boldsymbol{\alpha}\}$, my modified version of CAMB accepts as input the random shape described by $\Delta\boldsymbol{\alpha}$, and internally multiplies this by the ΛCDM linear power spectrum associated with $\boldsymbol{\pi}_\Lambda$. This generates the pseudo linear power spectrum which together with $\boldsymbol{\pi}_\Lambda$ serves as input for HALOFIT non-linear predictions. The boost factor is then produced by simply taking the ratio of the non-linear to the linear power spectrum. I generate boost factors for the training and trial sets of my emulator at redshifts 0 and 1. The $z = 0$ pseudo linear and non-linear matter power spectra, along with the corresponding boost factors, are shown in Figure 5.6 for a 500-node training set (in grey), as well as for the physical models (in orange) also illustrated in Figure 5.2.

5.4 Results

The non-linear matter power spectrum is key to the derivation of a number of cosmological observables, therefore I use this rather than the emulated statistic, $\ln B(k, z)$, to evaluate the accuracy of my emulator. I consider wave numbers in the range $k \in [0.01, 10]\, h/\mathrm{Mpc}$, excluding larger scales where the boost factor goes to unity in all models, presenting no challenge to the emulator. For each trial cosmology and redshift separately, I quantify the emulator accuracy as

$$\epsilon^{\rm emu} \equiv \max_{0.01<k<10} \left| \frac{P_{\rm NL}^{\rm emu}(k) - P_{\rm NL}^{\rm True}(k)}{P_{\rm NL}^{\rm True}(k)} \right| \tag{5.17}$$

where $P_{\rm NL}^{\rm emu}$ and $P_{\rm NL}^{\rm True}$ refer to the emulated and true pseudo non-linear power spectra respectively, and for the purposes of this study I take HALOFIT non-linear predictions as the truth. Figure 5.7 presents the fraction of the trial cosmologies with emulation errors smaller than a threshold value, $\bar{\epsilon}$, for various sizes of the training sets at both $z = 0$ (solid lines) and $z = 1$ (dashed). The fraction shown is the average across the 10 realisations per training set size. The left panel shows the results for the pseudo cosmology trials, whilst the right panel corresponds to the pure ΛCDM trial ensemble.

At $z = 0$, for the pseudo cosmology trials and training sets of size $N = 500$ or

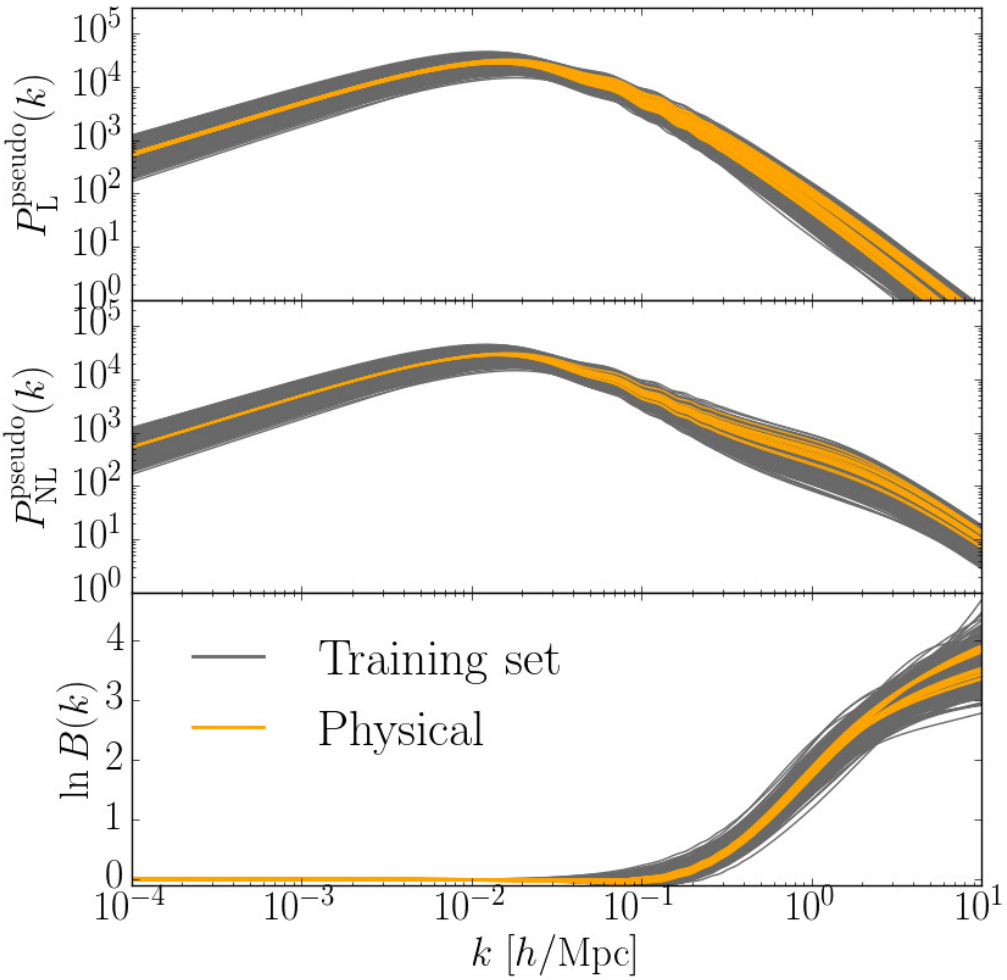

Figure 5.6 *The $z = 0$ linear (upper) and non-linear (middle) matter power spectra for the 500 pseudo cosmologies of a training set realisation (grey) and for the physical models (orange). The lower panel shows the logarithm of the boost factor, defined as the natural logarithm of the ratio of the middle to the upper panels (see equation 5.8). By design my training set comfortably encompasses the physical model predictions.*

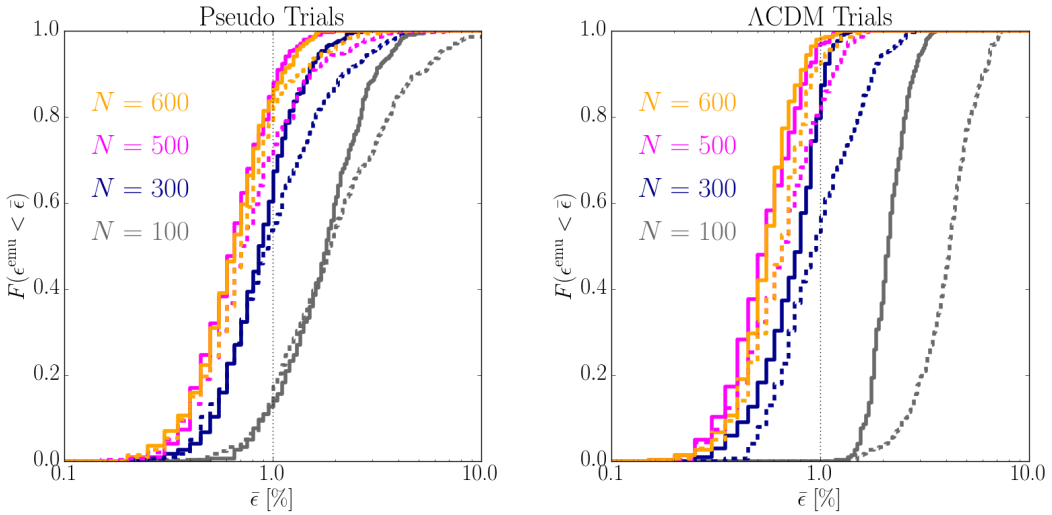

Figure 5.7 *The cumulative distribution, F, of trial predictions for which $\epsilon^{\rm emu}$, the maximum absolute fractional deviation between the emulated and* HALOFIT *pseudo non-linear matter power spectra over the range $k \in [0.01, 10]\,h/\mathrm{Mpc}$ (equation 5.17), is within a threshold, $\bar{\epsilon}$, for $z = 0$ (solid lines) and $z = 1$ (dashed lines). F is averaged across the 10 realisations per training set size, N, increasing from 100 nodes (grey) to 600 (orange). The left panel corresponds to the 300 pseudo cosmology trials. The right panel illustrates the performance for the 300 pure ΛCDM trials (see Section 5.3.3).*

600, the emulation is accurate to better than 1% for about 85% of the trials, and the full ensemble is recovered to better than 2% accuracy. As expected, the emulator performs noticeably worse when trained on a smaller number of nodes, with only 13% (60%) of the pseudo non-linear spectra emulated to better than 1% when $N = 100$ (300). This is because the emulator has less information on the complex relationship between input parameters, $\boldsymbol{\pi}$, and the boost factors. However, I find that the largest inaccuracy across all trial models is still $< 6\%$ ($< 3\%$) for $N = 100$ (300).

At $z = 1$, the fraction of pseudo trials with $\epsilon^{\rm emu} < 1\%$ are similar to those at $z = 0$. However, the reduced steepness of the cumulative distributions at $z = 1$, compared to those at $z = 0$, suggests an increasing fraction of outliers with redshift. This is caused by the training and trial boost factors at $z = 1$ having a broader dynamic range compared with those at $z = 0$, meaning the higher redshift poses greater challenge to emulation. The relative broadening is a consequence of my chosen range for the $\Delta\boldsymbol{\alpha}$ parameters, designed so as to tightly bracket the physical models at $z = 0$ (see Figures 5.5–5.6). If instead I had designed the parameters to follow the physical models at $z = 1$, the trend seen in the emulation accuracies with redshift would reverse. For future application of this work, I advise that the $\Delta\boldsymbol{\alpha}$ ranges be set so as to anchor the training set boost factors about a redshift most strongly probed by the data. As this is a proof of concept study however, choosing $z = 0$ to be the epoch of optimal emulation accuracy is sufficient for illustrative purposes.

The right panel of Figure 5.7 shows similar trends for the ΛCDM trial ensemble. For the two largest training sets I find that over 90% of the trial cosmologies are reproduced with better than per cent accuracy at both redshifts analysed. The larger emulation space dimensionality compared to similar emulator schemes designed for wCDM cosmologies only (see, e.g. Heitmann et al., 2014) demands a three-fold increase in the number of nodes to achieve equivalent accuracies for ΛCDM cosmologies. In fact, for my smallest training set ($N = 100$) the emulator predictions can have errors as large as 4% (7%) at $z = 0$ ($z = 1$), comparable to those obtained by Heitmann et al. (2014). Although my emulation strategy requires a few hundred training nodes to reach per cent accuracy, it is important to realise that it provides non-linear power spectrum predictions for a much broader class of cosmologies, including the ΛCDM model. Moreover, with the aid of *pairing and fixing* techniques (Angulo & Pontzen, 2016; Euclid Collaboration et al., 2018) the total number of training simulations required is analogous to

that of other GP emulators (Lawrence et al., 2010, 2017).

Finally I assess the accuracy of the emulator for the physical models in Table 5.1. Motivated by my results for the synthetic trial cosmologies, for this second test I use $N = 500$ training nodes, since a larger training set yields no significant improvement in the emulator performance.

The full range of the fractional errors of the emulated $P_{\rm NL}^{\rm pseudo}$ relative to the HALOFIT predictions, across the 10 training set realisations and all physical models, are shown in Figure 5.8 for $z = 0$ (upper panel) and $z = 1$ (lower panel). The outer (pale) and inner (dark) bands show the ranges including or excluding the xMNU_0.4 cosmology, respectively. Comparison of these reveals that much of the observed error is associated with this extreme model, which resides outside of the central 75%-per-side hypervolume used in evaluating the accuracies with the synthetic pseudo cosmologies. I find that the remaining models are recovered to better than the target per cent accuracy for $k < 10\,h/\mathrm{Mpc}$ at $z = 0$, with the majority of the training set realisations. I note that the range plotted, showing the total variation obtained across the training set realisations, presents a conservative estimate of the emulator performance; the 2σ range of *absolute* emulation accuracies, for example, is considerably tighter. I observe a slightly higher outlier rate at $z = 1$, consistent with my observations with the synthetic pseudo cosmologies. Of the models included in the dark band, I find, once again, that the largest inaccuracies occur for cosmologies near the boundary of the sub-volume used to test the synthetic shapes, MNU_0.4 and xMNU_0.2, both lying many standard deviations away from the *Planck* best fit cosmology (Planck Collaboration et al., 2018).

5.5 Conclusions

In this Chapter, I have designed and tested a methodology essential for predicting accurate non-linear matter power spectra in arbitrary extensions to the ΛCDM cosmology in the context of the *halo model reaction* framework (Cataneo et al., 2018, 'C18'). The effectiveness of the C18 strategy rests on the availability of accurate baseline ΛCDM non-linear power spectra evolved from non-standard initial conditions, the *pseudo* power spectra (see Section 5.2.1). I showed that such quantities can be readily predicted with a Gaussian process emulator whilst simultaneously satisfying the stringent accuracy requirements. I demonstrated

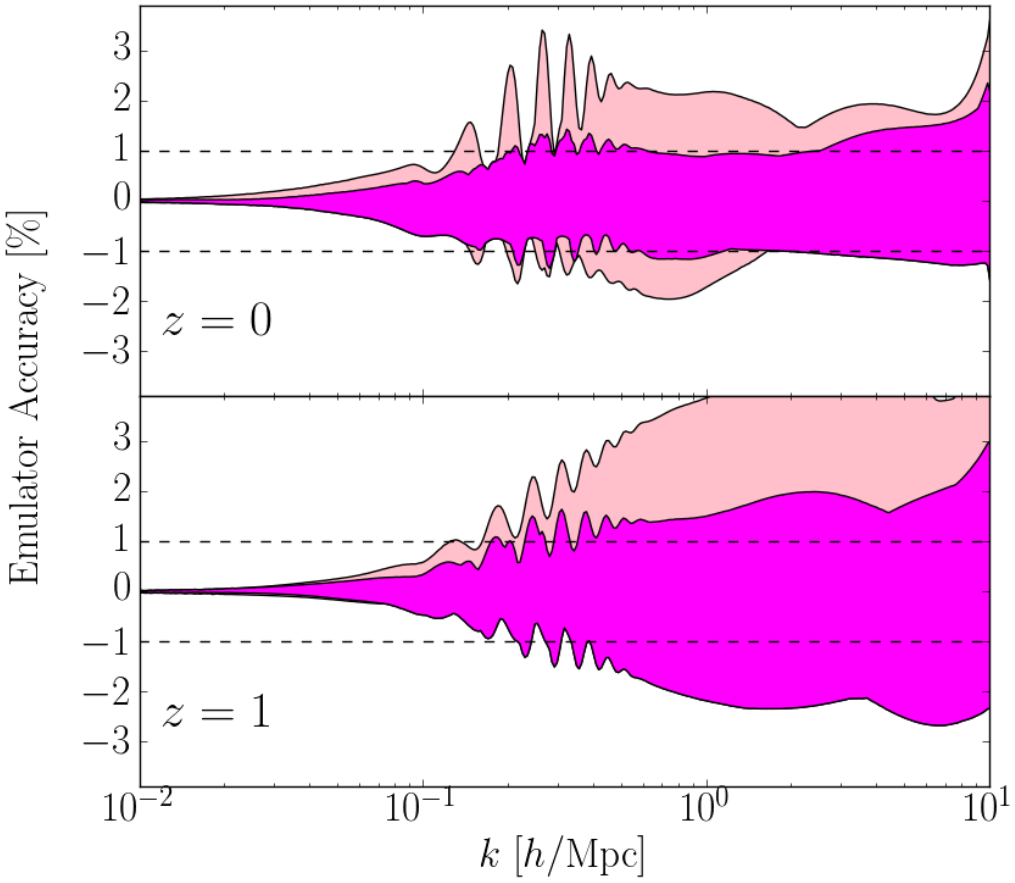

Figure 5.8 *The full range of accuracies of the emulated pseudo matter power spectra for the physical models at $z = 0$ (upper panel) and $z = 1$ (lower panel), obtained with my 10 training sets of 500 nodes. The difference in the outer (pale) and inner (dark) band is whether the xMNU_0.4 cosmology is included in calculating the range of accuracies. Hence, much of the observed error bar can be attributed to this extreme model, which lies close to the edge of my training sets.*

the power of this technique for theoretically motivated models, such as $f(R)$ gravity and massive neutrino cosmologies, as well as for phenomenological models, where in both cases structures on different scales grow at different rates.

Based on the results of this work, I advocate the following implementation of this methodology in the future:

1. A 13-dimensional optimised Latin hypercube is defined in terms of: $\boldsymbol{\pi}_\Lambda\,(\mathrm{dim}=5)$, the cosmological parameters that define the linear ΛCDM power spectrum, and $\Delta\boldsymbol{\alpha}\,(\mathrm{dim}=8)$, the shape parameters which modify the linear ΛCDM power spectrum to create the linear power spectrum for an arbitrary beyond-ΛCDM model.

2. Two standard ΛCDM N-body simulations, with initial conditions *paired* and *fixed* following Angulo & Pontzen (2016), are run for each node in the LH, i.e. with the input cosmology $\boldsymbol{\pi}_\Lambda$, and rescaling of the primordial power spectrum with the shape defined by $\Delta\boldsymbol{\alpha}$.

3. From the output of each N-body simulation, the pseudo non-linear matter power spectrum, $P_{\mathrm{NL}}^{\mathrm{pseudo}}(k,z)$, and hence the natural logarithm of the boost factor, $\ln B(k,z)$ (see equation 5.8), are measured. An emulator is subsequently trained to predict this quantity for any $\{\boldsymbol{\pi}_\Lambda, \Delta\boldsymbol{\alpha}\}$ combination.

4. A user will provide the analytically computed linear power spectrum in their chosen beyond-ΛCDM model, $P_{\mathrm{L}}^{\mathrm{real}}(k,z;\boldsymbol{\pi}_\Lambda,\boldsymbol{\pi}_X)$, where $\boldsymbol{\pi}_X$ is a vector of parameters describing the particular extension to the standard cosmology, and the corresponding vector of base cosmological parameters, $\boldsymbol{\pi}_\Lambda$. A PCA decomposition of the ratio, $P_{\mathrm{L}}^{\mathrm{real}}(k,z;\boldsymbol{\pi}_\Lambda,\boldsymbol{\pi}_X)/P_{\mathrm{L}}^{\Lambda}(k,z;\boldsymbol{\pi}_\Lambda)$, is then performed, where the denominator is the standard ΛCDM linear spectrum, to map $\boldsymbol{\pi}_X$ to the eight PCA weights (see equation 5.10).

5. Given the set of values of the base cosmological parameters and PCA weights for the queried model, the emulator predicts $\ln B(k,z)$.

6. The pseudo non-linear matter power spectrum is then simply obtained as $P_{\mathrm{NL}}^{\mathrm{pseudo}}(k,z) = B(k,z) \times P_{\mathrm{L}}^{\mathrm{real}}(k,z)$.

7. Finally, a rescaling of the pseudo non-linear power spectrum by the C18 halo model reaction, in equation 5.4, produces the *real* non-linear power

spectrum including the effects of physics for the chosen beyond-ΛCDM model.

The advantage of my emulator scheme over existing approaches (Heitmann et al., 2014; Lawrence et al., 2017; Euclid Collaboration et al., 2018; Winther et al., 2019) is twofold: the model-independent parameterisation enables predictions for beyond-one-parameter extensions to ΛCDM, effectively expanding the domain of applicability to a much broader class of cosmologically interesting models; and the simulations required for the training phase are all based on the standard ΛCDM cosmology, which makes them easier and faster to run.

This work is a first step towards a proper and accurate non-linear matter power spectrum emulator, one with a training set built from the output of N-body simulations. Here, instead, my goal was to perform a feasibility study, in which the approximate semi-analytical predictions of HALOFIT act as surrogate simulations (replacing step (ii) above). Taking inspiration from the linear power spectrum shape of well-studied extensions to the standard cosmology, I computed the training pseudo non-linear power spectra by feeding smooth, random modifications of the ΛCDM linear power spectrum to HALOFIT. Despite its intrinsic 5-10% inaccuracies, previous studies showed that HALOFIT provides a quick and robust way to design, construct and test the emulator (Heitmann et al., 2009; Euclid Collaboration et al., 2018).

With the aid of *pairing and fixing* techniques for the initial conditions (Angulo & Pontzen, 2016), I estimate that a total of 1000 ΛCDM-*like* simulations (two realisations for each of the 500 training nodes) is enough to build a per-cent-level accurate pseudo non-linear matter power spectrum emulator. This setup also allows for predictions of the ΛCDM non-linear matter power spectrum with $\lesssim$1% accuracy, even with the standard cosmology being completely absent from the training set. Despite the remarkable amount of computational resources required, the number of simulations is not too dissimilar from that used for the training of previous, more limited emulators (Lawrence et al., 2010, 2017).

Although not shown here, this methodology can also potentially be extended to include the effects on baryonic physics on non-linear scales (see, e.g., Schneider et al., 2019). This correction, which could be incorporated into the halo model reaction, is left for future investigation.

The results presented here should be regarded as conservative, as other emulation

schemes, such as *sparse polynomial chaos expansion* (Blatman & Sudret, 2011; Euclid Collaboration et al., 2018), and further design optimisation (Rogers et al., 2019; Caron et al., 2019) could help overcome the shortcomings of Gaussian process regression. Furthermore, I have only considered departures from ΛCDM on scales of $k > 10^{-3}\,h/\mathrm{Mpc}$. This requirement could be relaxed by incorporating only one additional emulation parameter, $\Delta\bar{k}$, describing translations of the scale at which departures occur about this point. Since the scope of this preliminary study was to show the feasibility of pseudo cosmology emulation, these topics are also left to future work.

My emulation method, coupled to the halo model reactions, enables fast, accurate, and flexible predictions of the matter power spectrum deep in the non-linear regime, where we are more likely to find imprints of new physics beyond the concordance cosmology, if any (see, e.g., Heymans & Zhao, 2018). Moreover, the approach developed in this work can be extended to predict other pseudo cosmology statistics, for instance galaxy clustering (Zhai et al., 2019), as well as mean halo properties, such as their abundance and concentration (Kwan et al., 2013; McClintock et al., 2019).

5.A Modelling BAO residuals

As shown in Figure 5.1, models with massive neutrinos have shapes characterised by a rapidly oscillating component superimposed over a smooth function. These wiggles are associated with changes in the baryon acoustic oscillations (BAO) produced by differences in the sound horizon at the end of the drag epoch and shallower gravitational potentials before recombination (Hu & Sugiyama, 1996; Hu & White, 1996; Eisenstein & Hu, 1998; Lattanzi & Gerbino, 2017) compared to a massless neutrino ΛCDM cosmology with the same standard parameters, i.e. $\boldsymbol{\pi}_{\Lambda}^{m_\nu=0} = \boldsymbol{\pi}_{\Lambda}^{m_\nu\neq0}$, and $\Omega_{\mathrm{c}}^{m_\nu=0} = \Omega_{\mathrm{c}}^{m_\nu\neq0} + \Omega_{\nu}^{m_\nu\neq0}$. Depending on the value of the standard cosmological parameters and neutrino mass, the amplitude of the BAO residual in the shape can be up to ∼1% of the slowly varying component. The PCA reconstruction outlined in Section 5.3.1 requires a large number of components to simultaneously capture long- and short-scale variations. Hence I find errors $\gtrsim$1% in reconstructing the $\sum m_\nu$-induced oscillations with only $n_\Phi = 8$ basis functions, as shown by the grey curves in the upper panel of Figure 5.9.

In order to improve the reconstruction accuracy of the PCA, I isolate and remove

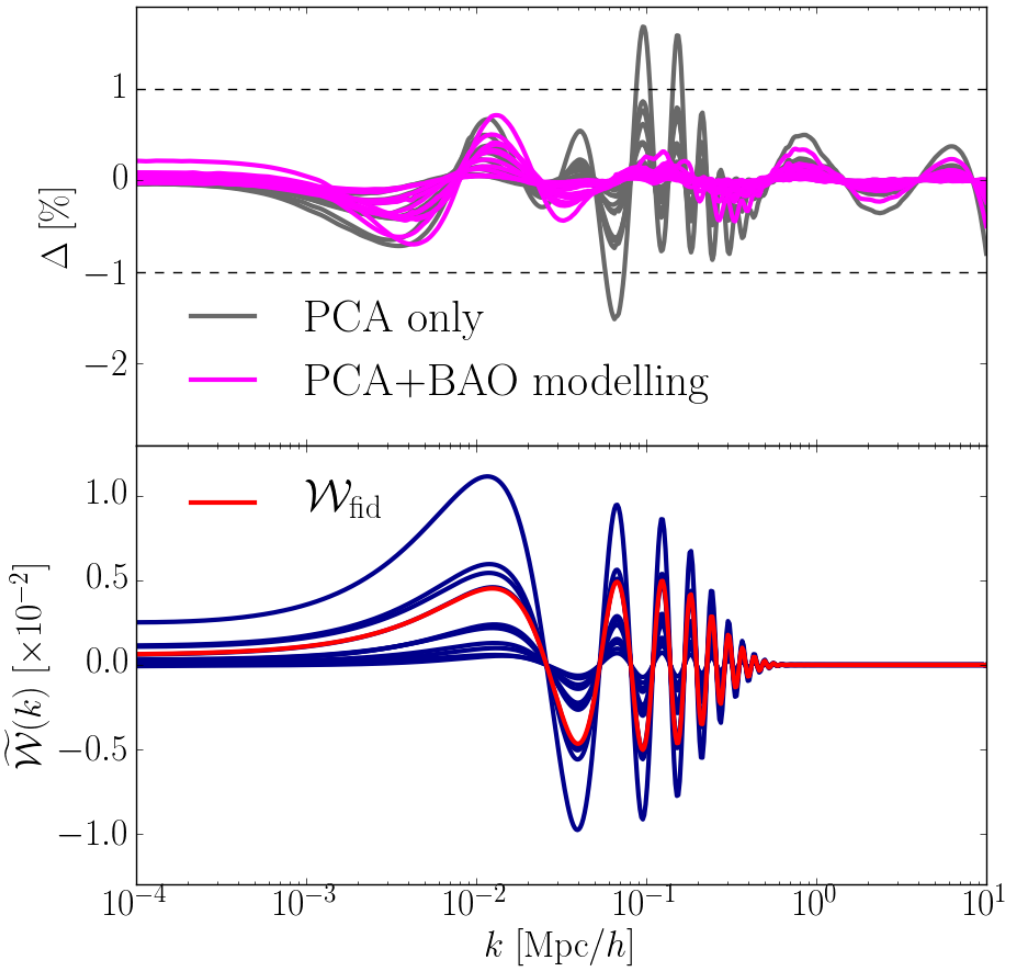

Figure 5.9 *Upper panel: the accuracy of the shape reconstruction, Δ, for the physical models featuring massive neutrinos (see Table 5.1). The grey curves show the results for direct PCA reconstruction using the 8 orthogonal basis functions derived in Section 5.3.2. The magenta curves, reproduced from Figure 5.3, illustrate the benefit of including the BAO residuals reconstruction via the two-step process in equation 5.21: the BAO residuals are first extracted with the de-wiggling algorithm presented in Appendix 5.B and then modelled using the template (see equation 5.20), whilst the PCA reconstruction is performed on the smooth component only. Lower panel: the red curve shows the BAO residual associated with the MNU_0.4 cosmology, which I use as the fiducial template. The blue curves show the best fit functions, $\widetilde{\mathcal{W}}(k;\bar{a},\bar{b})$, to the oscillatory component of the shapes for the remaining physical models with massive neutrinos. All curves here correspond to $z = 0$; results are very similar at $z = 1$.*

the BAO residuals by applying the de-wiggling algorithm presented in Appendix 5.B. In short, the BAO in Fourier space maps to a localised bump in the discrete sine transform of the power spectrum. By removing this bump and performing the inverse transform, I obtain a smooth, "de-wiggled" version of the power spectrum. This is performed on both the model featuring massive neutrinos and on the ΛCDM baseline, before calculating the de-wiggled shape given by

$$S^{\rm dw}(k,z) = \frac{P_{\rm L}^{\rm pseudo,dw}(k,z)}{P_{\rm L}^{\Lambda,{\rm dw}}(k,z)}\,, \tag{5.18}$$

where the superscript 'dw' on the quantities on the right-hand side denote power spectra obtained from applying the de-wiggling algorithm. This smoothed component of the shape is captured very accurately by the basis functions. Reconstruction of the *full* shape then rests on modelling the remaining oscillatory component, given by

$$\mathcal{W}(k,z) = S(k,z) - S^{\rm dw}(k,z). \tag{5.19}$$

Notice that all quantities in equations 5.18 and 5.19 depend implicitly on the cosmological parameters. I experimented with fitting these oscillations with the function

$$\widetilde{\mathcal{W}}(k;a,b) = [a + b\log_{10}(k)]\mathcal{W}_{\rm fid}(k)\,, \tag{5.20}$$

where a and b are free parameters controlling the amplitude and modulation of a fiducial wiggle template, $\mathcal{W}_{\rm fid}$, which I identify with the wiggle contribution of the MNU_0.4 cosmology in Table 5.1, represented by the red curve in the lower panel of Figure 5.9. I checked that my fiducial template choice in equation 5.20 is robust against changes in redshift, standard cosmological parameters, sum of neutrino masses and neutrino mass eigenstates. This ensures that the parameters $\{a,b\}$ are sufficient to capture BAO variations in the family of pseudo cosmologies treated in this work. The blue curves in the lower panel of Figure 5.9 show the behaviour of my fitting formula, from equation 5.20, for all of the test cosmologies in Table 5.1 containing massive neutrinos.

In summary, for each redshift and cosmology separately, I reconstructed the total

shapes of the physical models as

$$S(k;\boldsymbol{\pi}_\Lambda,\Delta\boldsymbol{\alpha},a,b)=S^{\rm dw}(k;\boldsymbol{\pi}_\Lambda,\Delta\boldsymbol{\alpha})+\widetilde{\mathcal{W}}(k;a,b)\,, \tag{5.21}$$

where only the smooth, de-wiggled term is modelled via the PCA reconstruction given in equation 5.10. The magenta curves in the upper panel of Figure 5.9 show the overall accuracy of the two-step reconstruction, originally presented in the lower panel of Figure 5.3. I find that the inaccuracies in the BAO regime are greatly suppressed compared to performing the PCA directly on the full shape (grey curves).

The fact that the cosmologies with massive neutrinos require two extra parameters to summarise their shapes to sub-per-cent accuracy, might imply that the emulator need be informed of $\{a,b\}$ to achieve this target accuracy in its predictions, thereby increasing the dimensionality from 13 to 15. To test this, I produced MaxPro training sets optimised in 15 dimensions, with a range of wiggle parameters, a and b, slightly broader than those obtained for the physical models with massive neutrinos. I then compared the emulator's performance with the physical models, when trained on these 15-dimensional sets, and when trained on the same sets but in 13-dimensions, omitting a and b. I found, in fact, that the emulator is mostly insensitive to these two parameters, which therefore contribute only marginally to improving its accuracy.

Figure 5.10 holds the key to this somewhat surprising result, where I adopt the xMNU_0.4 cosmology as a worst case scenario among my physical models. Plotted in solid orange is the $S/S^{\rm dw}$ ratio, which is equivalent to the ratio of pseudo linear power spectra with or without changes to the BAO scale compared to ΛCDM (see equation 5.18). In dashed dark blue is the ratio of their non-linear counterparts obtained with halofit. Solid magenta shows the ratio between the boost factors derived from using the full or de-wiggled shape, which is also equal to the ratio of the dark blue and orange curves. We see that although the BAO residual amounts to $\sim$1% of the non-linear power spectrum, its contribution to the boost factor is much smaller. In turn, this means that the emulator can reconstruct the boost factor with the required accuracy irrespective of whether the information on the BAO residual is provided or not. In other words, the boost factor effectively damps down the correction associated with changes to the BAO caused by massive neutrinos, eliminating the need for the additional

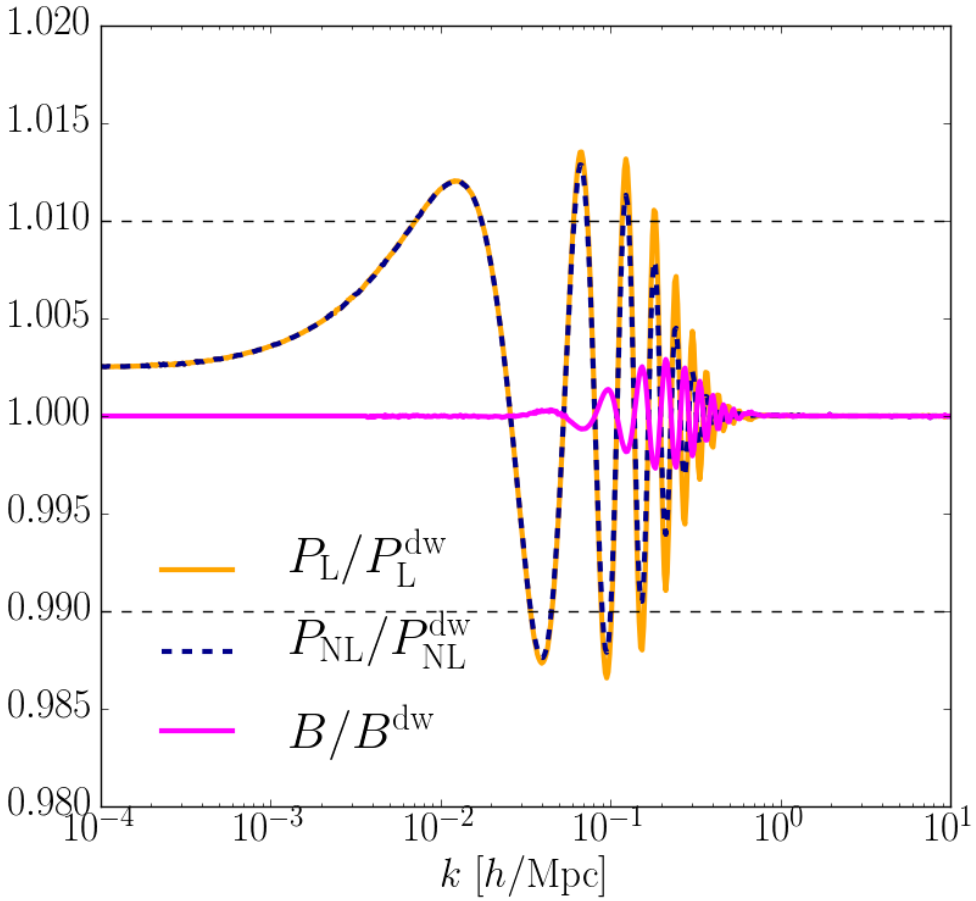

Figure 5.10 *Full to de-wiggled ratios for the boost factor, linear and non-linear matter power spectra of xMNU_0.4 in Table 5.1. Quantities in the legend are defined as follows:* $P_{\rm L} = S \times P_{\rm L}^{\Lambda}$; $P_{\rm L}^{\rm dw} = S^{\rm dw} \times P_{\rm L}^{\Lambda}$; $P_{\rm NL} = {\rm HF}[P_{\rm L}]; P_{\rm NL}^{\rm dw} = {\rm HF}[P_{\rm L}^{\rm dw}]$. *The de-wiggled shape,* $S^{\rm dw}$, *is obtained as in equation 5.18, while HF is shorthand for* HALOFIT. *The solid orange and dashed blue lines represent the linear and non-linear BAO residuals, respectively, whose ratio is equivalent to the ratio of the boost factors,* $B = P_{\rm NL}/P_{\rm L}$ *and* $B^{\rm dw} = P_{\rm NL}^{\rm dw}/P_{\rm L}^{\rm dw}$ *(solid magenta). The contribution of the BAO residual to the non-linear power spectrum exceeds 1%, but is at most* $\simeq 0.3\%$ *in the boost, which implies I can safely ignore the wiggle parameters,* $\{a, b\}$, *in emulation.*

wiggle modelling (equation 5.20)[8].

The fact that early-time modifications to the BAO physics have such a weak effect on the boost factor is a strength of my emulation scheme; an emulator designed to directly predict the pseudo non-linear power spectra is likely to incur $\gtrsim$1% errors unless detailed modelling of the BAO residual is employed. The unimportance of the wiggle parameters means that the 13-dimensional emulator is sufficient to achieve my target accuracy. In Appendix 5.C, I investigate the performance of the emulator with cosmologies featuring extra relativistic degrees of freedom, for which changes in the acoustic oscillations prior to the epoch of last scattering leave especially prominent wiggles in the late-time shape ratios. I find the emulator achieves sub-per-cent accuracy at $z = 0$, but less consistent results at higher redshift for this model.

Although modelling of the BAO residuals proves to be unnecessary, their extraction through the de-wiggling algorithm remains important. By acting as a low-pass filter, the PCA reconstruction can in principle remove rapid oscillations and capture the broadband shape. However, the PC weights thus derived are somewhat different from those obtained after de-wiggling, generating an error that ultimately propagates to the boost factor. I compare the emulator performance with and without de-wiggling in Appendix 5.D.

5.B De-wiggling methodology

Here I outline in more detail the method, discussed briefly in Appendix 5.A, for de-wiggling linear matter power spectra following Baumann et al. (2018). This procedure is performed on the pseudo and ΛCDM spectra separately, and therefore I denote both with $P_{\rm L}(k)$ in the following:

1. I firstly interpolate the linear matter power spectra featuring BAO wiggles, $P_{\rm L}(k)$, from the logarithmically-spaced k values onto a linear-spaced array, k_n, sampled at 2^n points, where n is an integer. This is so as to improve the computational efficiency of the Fast Discrete Sine Transform (DST), which is then performed on $\log_{10}[k_n P_{\rm L}(k_n)]$ using the orthonormalised type-II sine transform. The resulting array, which has indices denoted by $i_{\rm dst}$ and length

[8]Although I verified this statement with HALOFIT I expect it to remain valid in simulations as well, assuming that HALOFIT provides accurate predictions for the ratio $P_{\rm NL}/P_{\rm NL}^{\rm dw}$.

2^n, is split into two separate arrays, one with the even $i_{\rm dst}$ indices and the other with odd $i_{\rm dst}$ indices.

2. The DST of the even and odd arrays each feature a bump, which is a localised manifestation of the BAO in the k-space of the matter power spectrum. This is shown by the solid lines for one of the physical models in the upper panel of Figure 5.11. In order to de-wiggle the power spectrum, I must remove these bumps. To this end, I take the second derivative of the DST, converting the observed bump in each array into a prominent single-period oscillation, as shown by the lower panel of Figure 5.11. I then identify $i_{\rm dst}^{\rm min}$ and $i_{\rm dst}^{\rm max}$ values bracketing the oscillation, which are close to the local minimum on the left-hand side of the approximate centre of this feature and the local maximum on the right-hand side respectively.

3. I remove the corresponding $i_{\rm dst}$-range from the even and odd DST arrays, multiply each by $(i_{\rm dst}+1)^2$ and interpolate across the gap between $i_{\rm dst}^{\rm min}$ and $i_{\rm dst}^{\rm max}$ with cubic splines. The $(i_{\rm dst}+1)^2$ factors are then divided out, generating smooth "de-bumped" even and odd DST arrays, where the information corresponding to the BAO has been removed. This is shown by the dashed lines in the upper panel of Figure 5.11.

4. I recombine the even and odd DST arrays into a single array. Performing an inverse type-II DST produces the logarithmic de-wiggled linear matter power spectra weighted by the linear-spaced k_n array, $\log_{10}[k_n P_{\rm L}^{\rm dw}(k_n)]$. From this I obtain the de-wiggled linear spectra itself, $P_{\rm L}^{\rm dw}(k_n)$.

5. After performing this on both the pseudo and ΛCDM power spectra, I take the ratio of the respective $P_{\rm L}^{\rm dw}(k_n)$ to obtain $S^{\rm dw}(k_n)$, the de-wiggled shape sampled at the linearly-spaced wave numbers. Finally I interpolate the de-wiggled shape back to the logarithmically-spaced k sampling, obtaining $S^{\rm dw}(k)$. I find that there is less numerical noise at the upper and lower wave number bounds of the de-wiggled shape when I take the ratio of the linearly-sampled de-wiggled power spectra before interpolating, rather than the other way around.

Whereas Baumann et al. (2018) chose fixed values of the $i_{\rm dst}^{\rm min}$ and $i_{\rm dst}^{\rm max}$ relative to the oscillation in the second derivative of the DST arrays, I optimise for these parameters in my analysis with the physical models. Specifically, I cycle through a grid of $(i_{\rm dst}^{\rm min}, i_{\rm dst}^{\rm max})$ combinations which are centred on the local minimum and

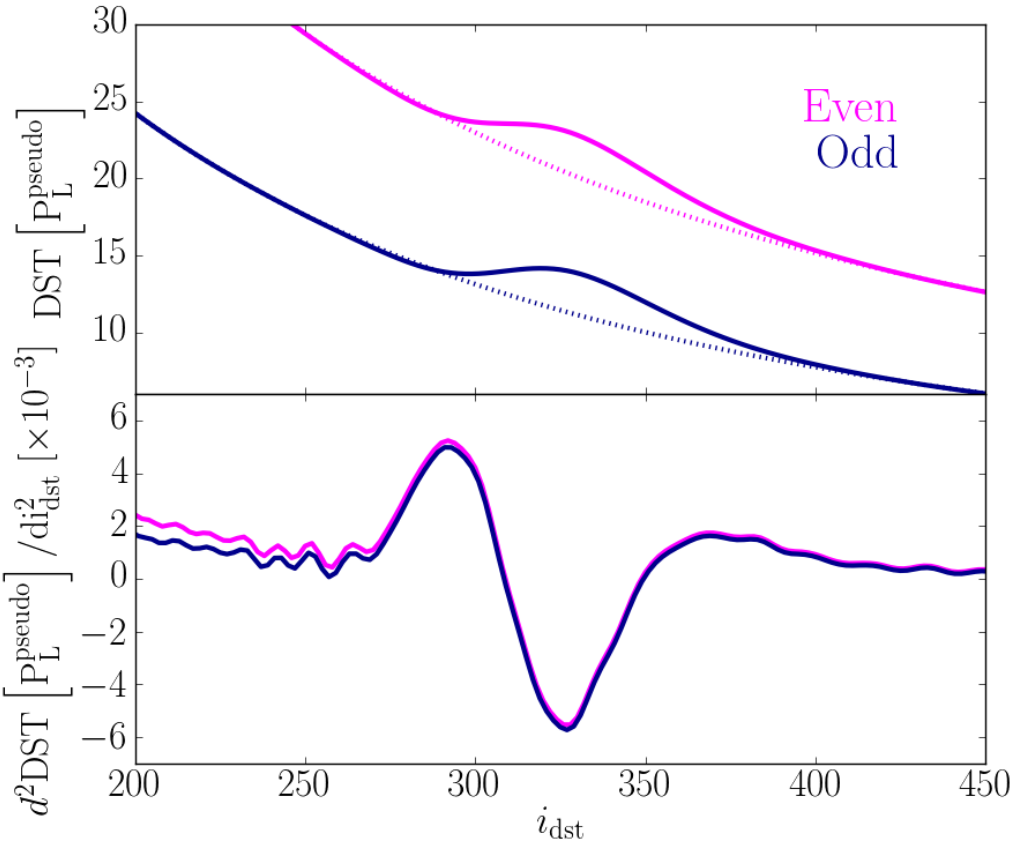

Figure 5.11 *Upper panel: The discrete sine transform (DST) of the pseudo linear matter power spectrum for the F4-MNU_0.4 model, split into two arrays with even (magenta) and odd (dark blue) indices,* i_{dst}*, before (solid lines) and after (dashed lines) the bump is removed. The bump is a localised manifestation of the BAO wiggles present in the power spectrum. Lower panel: The second derivative of the DST used to identify the* $i^{\mathrm{min}}_{\mathrm{dst}}$ *and* $i^{\mathrm{max}}_{\mathrm{dst}}$ *values which bracket the bump in the DST. I adjust the values of these for each model, optimising the smoothness of the returned de-wiggled shape. The final* $i^{\mathrm{min}}_{\mathrm{dst}}$ *and* $i^{\mathrm{max}}_{\mathrm{dst}}$ *values are close to, respectively, the local minimum on the left-hand side of the centre of the prominent oscillation in the second-derivative, and the local maximum on the right-hand side.*

maximum on either side of the oscillation, and identify those which return the smoothest de-wiggled shape. The smoothness is measured by integrating the second derivative of $S^{\rm dw}(k)$, convolved with a Gaussian filter, in the range of the BAO wiggles. The optimal $(i_{\rm dst}^{\rm min}, i_{\rm dst}^{\rm max})$ combination is that which minimises this integral. I find the de-wiggled shapes are insensitive to the width of the Gaussian filter, which is employed so as to ensure numerical noise does not bias identification of the optimal $(i_{\rm dst}^{\rm min}, i_{\rm dst}^{\rm max})$ range.

5.C Extension to extra relativistic degrees of freedom

Appendix 5.A demonstrated the unimportance of modelling the BAO residuals in the shapes of the physical cosmologies, justifying the 13-dimensional emulation scheme. In order to further test the validity of this, I try my emulator scheme against a theoretically motivated cosmology with an especially prominent BAO residuals in the shape ratio.

The existence of a relic sea of neutrinos with effective relativistic degrees of freedom, $N_{\rm eff} = 3.046$, is a general prediction of the standard model (Mangano et al., 2005). However, physics beyond this paradigm includes scenarios with additional relativistic particles (or 'dark radiation') at the epoch of decoupling, typically quantified by the variation $\Delta N_{\rm eff}$, such that

$$N_{\rm eff} = 3.046 + \Delta N_{\rm eff}. \tag{5.22}$$

This has the effect of reducing the expansion rate and increasing the acoustic scale (Archidiacono et al., 2013), thereby acting to ease the tension between the early- and late-time measurements of the Hubble constant (Riess et al., 2019, see Figure 1.10). I consider a model with $\Delta N_{\rm eff} = 1$ extra relativistic degrees of freedom, which lies several standard deviations away from the best-fit result of the Planck Collaboration et al. (2018) analysis. The BAO residuals in the shape of this cosmology at $z = 0$ are shown in the upper panel of Figure 5.12.

Following the same procedure employed for the physical models in Table 5.1, I obtain a smoothed version of this shape using the de-wiggling algorithm (see Appendix 5.B), perform a PCA decomposition to obtain the PCA weights, $\Delta\boldsymbol{\alpha}$, but refrain from modelling the BAO residual component as described in Appendix

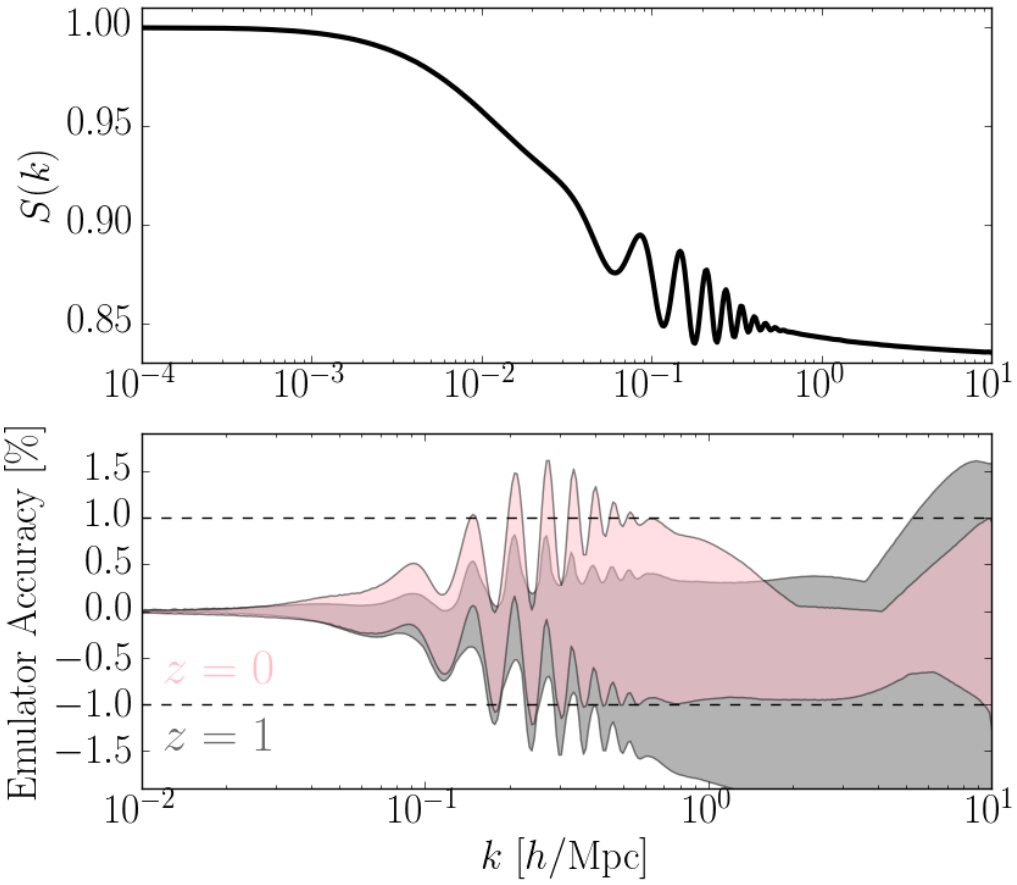

Figure 5.12 *Upper: the $z = 0$ shape ratio for the model with $\Delta N_{\rm eff} = 1$ extra relativistic degrees of freedom. The BAO residual is at the level of $\simeq 2\%$, about twice as large as the residuals for physical models with massive neutrinos, shown in Figure 5.1. Lower: the full range of emulation accuracies achieved with the 10 realisations of the $N = 500$ training set at $z = 0$ (pink) and $z = 1$ (grey). Despite the larger prominence of the BAO residual to the shape, I find its contribution to the boost factor remains at the sub-per-cent level, facilitating reconstruction of the $z = 0$ pseudo non-linear power spectrum to better than the target accuracy with almost all of the training set realisations. The larger inaccuracies observed at $z = 1$ suggest that this model cannot reliably be constructed at higher redshifts with the current emulation scheme.*

5.A. I then infer the $P_{\rm NL}^{\rm pseudo}$ for this model using the emulator in the fiducial 13-dimensional setup and compare to the prediction from HALOFIT.

The results using the 500-node training sets are shown in the lower panel of Figure 5.12. I find that despite the BAO residual in the shape being as large as 2% for this cosmology[9], the emulator reproduces the $P_{\rm NL}^{\rm pseudo}$ to better than 1% accuracy at $z = 0$ (pink) in almost all of the training set realisations. The reason that these accuracies are achievable, as with the physical models, is due to the boost factor suppressing the early-time modifications of the BAO to sub-per-cent levels. I also note that changes in the linear matter power spectrum for models with $\Delta N_{\rm eff} \neq 0$ (see Figure 5.12) are caused by processes happening well before the matter dominated epoch. At redshifts $\gtrsim$100 the background and growth evolution in these extensions are indistinguishable from those in ΛCDM. Therefore, the *full* non-linear matter power is entirely captured by the pseudo cosmology (i.e. the reaction is unity for all scales, see Section 5.2.1). At $z = 1$, however, I find that this model cannot be reliably predicted to within the target accuracy across the training set realisations. This is perhaps an indication that models with such marked BAO residuals are beyond the capabilities of my emulator as specified in this work. This may be improved upon given an alternative emulation method, such as *sparse polynomial chaos expansion* (Blatman & Sudret, 2011; Euclid Collaboration et al., 2018), or with careful optimisation of the training set configuration (Rogers et al., 2019; Caron et al., 2019).

5.D Emulation without de-wiggling

Although modelling of the BAO residuals in the shapes to obtain the wiggle parameters, a and b, has negligible effect on the results (see Appendix 5.A), I find that shape de-wiggling remains important for improving emulation accuracy. This is because the PCA is a low-pass filter, effective at capturing only the low-frequency component of the shapes. This property is imparted by the basis set, Φ_i, derived from smooth curves (see Section 5.3.2), such that the corresponding weights, $\Delta\alpha_i$, are designed to contain smooth component information. Consequently, performing the PCA directly on the full shape leads to misestimation of the $\Delta\alpha_i$ values, and bias in the emulation.

Figure 5.13 presents the effect of this for the xMNU_0.4 cosmology (see Table

[9]For comparison, this feature never exceeds 1% in my other test cosmologies.

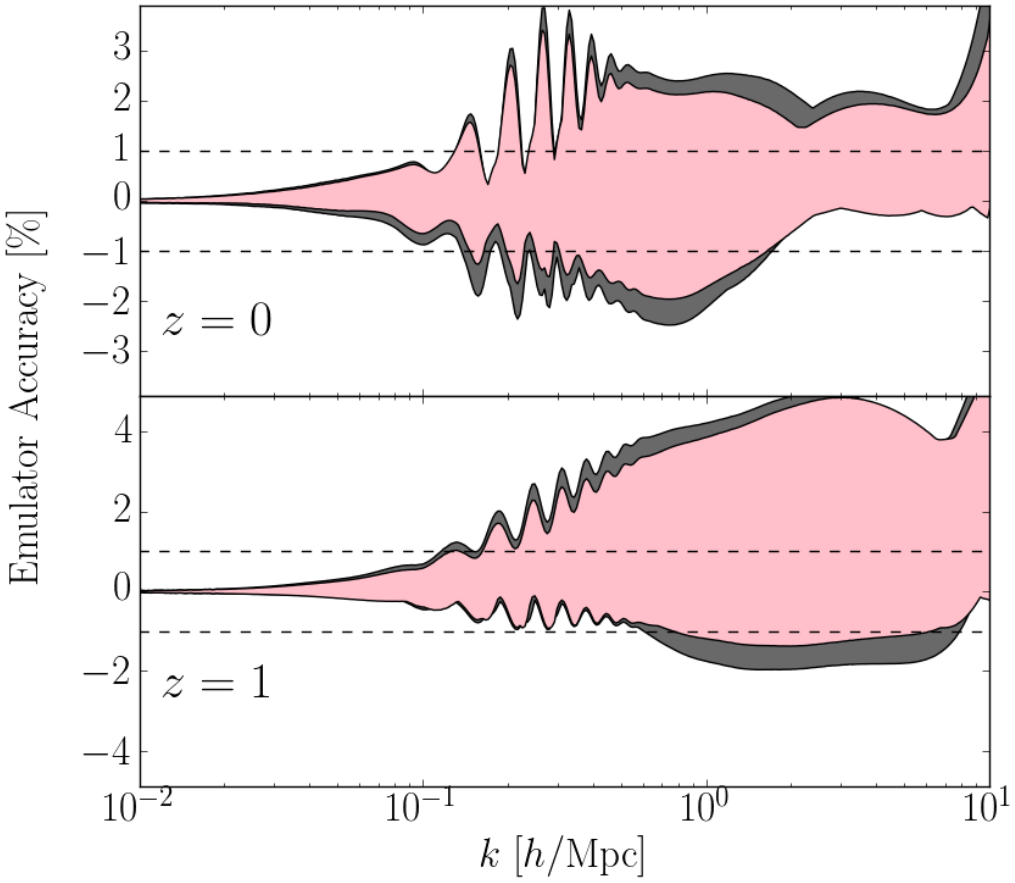

Figure 5.13 *The range of accuracies of the emulated pseudo matter power spectra for the xMNU_0.4 cosmology (see Table 5.1) at $z = 0$ (upper panel) and $z = 1$ (lower panel), obtained with my 10 training sets of 500 nodes. The inner (pink) band corresponds to my fiducial emulation method, reproduced from Figure 5.8. The outer (grey) band corresponds to the case with no de-wiggling and PCA decomposition performed directly on the full shape.*

5.1) at $z = 0$ (upper panel) and $z = 1$ (lower panel). The pink band shows the full range of emulation accuracies achieved with the 10 realisations of the 500-node training set for my fiducial emulation method incorporating the de-wiggling, reproduced from Figure 5.8. The grey band shows this quantity for the case with no de-wiggling, where the PCA decomposition is performed directly on the full shape. Indeed, excess inaccuracies of $\sim$0.5% are observed in the absence of de-wiggling, for the offset PCA weights map to a boost factor related to a slightly different shape compared to the HALOFIT predictions. Although currently subdominant, these errors will become more relevant for future emulators designed to have target performance of $\sim$0.5%, a requirement imposed by irreducible inaccuracies in the halo model reactions.

The road and the tale have both been long, would you not say so? The trip has been long and the cost has been high... but no great thing was ever attained easily. A long tale, like a tall Tower, must be built a stone at a time.

Stephen King, The Dark Tower

6

Conclusions

We live in an era of *precision cosmology*, where model parameters are inferred with accuracies approaching the per cent level. However, these are also testing times for this field. The comparison of independent observational probes of the low and high redshift Universe has highlighted a *cosmological discordance*, providing significant evidence for either unaccounted systematics or an incompleteness in the standard model. Now, more than ever, I believe novel methodologies are called for, to shed further light on this puzzle, and to strengthen our capabilities to reach precision parameter constraints with the next generation of cosmological data.

In this pursuit, I have championed the method of weak gravitational lensing, wherein the large scale mass distributions of the cosmic web are probed via the coherent statistical correlations in the images of background galaxies. Much information on the overdense, virilised structures in particular is contained in the lensing signal, and yet it lies beyond the reach of the more simple, linear-prescription diagnostics. Specifically, this thesis has focused on how enhanced cosmological constraints can be extracted from the non-linear side of the Universe, by combining: well-established and newly developed theoretical

modelling techniques in weak lensing; state-of-the-art numerical simulations, supporting a broad range of cosmologies; and recent advances in machine learning algorithms. In this final Chapter, I will summarise my main findings, and forecast how this work may benefit the analyses of data from future lensing surveys.

6.1 A summary of findings

Weak lensing presents an effective but challenging route to access cosmological information. In Chapters 1 and 2, I presented an in-depth discussion of the theoretical grounds to why constraining the cosmological model is possible via this method, with the latter including the myriad of systematics which serve to mire our inference. The three main weak lensing surveys, DES, KiDS and HSC, however, argue that the main sources of bias, presented in Section 2.6, have been accounted for in their recent cosmic shear analyses (Troxel et al., 2017; Hildebrandt et al., 2018; Hamana et al., 2019, respectively). One would conclude, therefore, that these cannot be the cause of the discrepancies in the measurements of the clustering parameter, $S_8 = \sigma_8/\sqrt{\Omega_\mathrm{m}/0.3}$, from weak lensing and the cosmic microwave background (Planck Collaboration et al., 2018).

In Chapter 2 (Section 2.6.1), I also presented my early work in investigating the reliability of shear estimation with the latest KiDS data, featuring 1000 deg^2 of deep multi-band imaging. This included executing, and inspecting the quality of, dedicated image processing. Furthermore, I presented results quantifying the impact of imperfections in the PSF modelling, finding it to be at the suitably low, sub per cent level over the range of angular scales relevant to cosmic shear measurement.

In Chapter 3, I developed the use of a novel field transformation, known as "clipping", and for the first time applied this technique in a cosmic shear analysis with real data, the first 450 deg^2 of KiDS imaging (KiDS-450; Hildebrandt et al., 2017). Subsequently, I measured clipped analogues of the conventional, "unclipped" shear correlation functions, showing that the combination of these two types of statistic, yield improvements in the precision of the inferred S_8, by 17%, over the standard analysis alone. At least part of the information gain in clipping comes from the non-linear Universe, with the selective truncation of the signal from overdensities informing constraints, in a similar manner to the method of peak abundance statistics (see Section 2.5.4 for further discussion).

Measurement of any cosmological statistic presents challenges, and clipping is no exception. Most notable of these, is the requirement of numerical simulations, varying broadly in their input cosmologies, to determine the expectation value of the clipped signal. In Chapter 3 I used the Dietrich & Hartlap (2010, DH10) suite for this purpose, although irreducible noise present in these mocks inevitably manifested in the inferred likelihood contours. I concluded that credible inference of the power of clipping could be inferred for the KiDS-450 data, but more advanced simulations would be needed for applying this methodology to larger surveys with greater statistical power.

In Chapter 4, I addressed this concern, introducing the *cosmo*-SLICS suite (Harnois-Deraps, Giblin & Joachimi, 2019). These simulations feature 26 different 4-dimensional input cosmologies, with larger box sizes and improved mass resolution over DH10. Additionally, the initial conditions of the pair of simulations executed at each node, are fixed to achieve variance cancellation (see, for example, Angulo & Pontzen, 2016), facilitating a low-noise estimate of the expectation value of inferred cosmological statistics. In this Chapter, I tailored an emulator, employing Gaussian process regression, to the *cosmo*-SLICS. This allowed for robust interpolation of lensing statistics measured in the simulations, to arbitrary cosmologies, across the full parameter space consistent with the current cosmic shear constraints. I demonstrated that the emulator achieves $\simeq$1% interpolation accuracy, and that including the intrinsic noise of the simulations, degraded this to the $\lesssim$5% mark - still competitive for concurrent parameter inference with cosmic shear.

The true benefit of the *cosmo*-SLICS suite and emulator, lies in the freedom they provide in pursuing new and powerful lensing probes, those without an analytical recipe. To demonstrate this, I applied clipping to these simulations, and used the emulator, trained on the resultant clipped correlation functions, to explore and constrain cosmology with these techniques. The low levels of noise in the *cosmo*-SLICS emulator's predictions allowed for a more robust determination of the improvement in precision yielded by combining clipped and unclipped probes: a 25% and 21% reduction in the uncertainties in $\Omega_{\rm m}$ and S_8 respectively, over the unclipped alone. I further showed that clipping is in fact also sensitive to the Hubble and present dark energy equation of state parameters, h and w_0, with the combination of probes yielding 26% and 18% improvements in precision, respectively. This finding was inaccessible to the Chapter 3 analysis, featuring the DH10 mocks with fixed values of these parameters, but is realised in Chapter

4 thanks to the specifications of the *cosmo*-SLICS design, in which h and w_0 vary.

Finally, in Chapter 5, I exercised the emulator method alongside new developments in theoretical modelling, that of halo model reaction functions (Cataneo et al., 2018), to create a general, computationally efficient framework, in which accurate non-linear matter power spectrum predictions can be generated for arbitrary cosmological models. This requires only standard ΛCDM gravity-only N-body simulations, with initial power spectra modulated by smooth random functions, generated to encompass the range of small-scale departures from the standard model permitted by observational constraints. Once trained on the output of these simulations - the 'pseudo' non-linear matter power spectra - the emulator would permit accurate interpolation of this quantity for practically any input cosmology not currently ruled out by data. The emulator's predictions need only be scaled by the appropriate reaction function, inexpensively computed from the halo model (Cataneo et al., 2018), to achieve a ∼1% accurate power spectrum out to the highly non-linear scale, $k = 10\,h/\mathrm{Mpc}$.

This Chapter served as a proof-of-concept for the production and emulation of the pseudo cosmologies, with the non-linear component of the pseudo power spectra approximated using the fitting function formalism from Takahashi et al. (2012), enacted in HALOFIT, rather than computed accurately and expensively with simulations. This allowed me to forecast the number of simulations one would need for this work to be realised in reality, arriving at an estimate of 1000 standard model simulations (500 input cosmologies with two matched-pair simulations at each node). I demonstrated that a training set of this size, boosts the performance of the emulator in predicting the pseudo power spectra, to the required accuracy for a broad class of cosmological extensions to ΛCDM. These included $f(R)$ gravity, massive neutrino cosmologies, combinations, and indeed synthetic cosmologies which deviate from the standard model on small scales. The training set requirements, although undoubtedly costly to initially produce, are similar in size to previous simulation suites (Lawrence et al., 2010, 2017; Harnois-Déraps et al., 2018), and most importantly, permit non-linear predictions in *arbitrary* extensions to ΛCDM. This capability cannot be matched by any other emulator proposed to date.

Common to all of the research in Chapters 3-5, is the extraction of cosmological content from the murky, non-linear, underbelly of the Universe. In this, I employed a powerful trinity of resources: theory, simulations and machine learning. Clipping, made possible by advances in simulations, is but one member

of a family of non-Gaussian statistics, including peak statistics (Jain & Van Waerbeke, 2000; Kacprzak et al., 2016; Shan et al., 2018; Martinet et al., 2018), higher-order statistics (Fu et al., 2014), Minkowski functionals (Petri et al., 2015), and density-split lensing (Gruen et al., 2018; Brouwer et al., 2018), whose sole *raison d'être* is this purpose. The state-of-the-art simulations, *cosmo*-SLICS, twinned with an emulator reliant on machine learning methodology, exist primarily to provide practical means to apply such methods to lensing data sets. Finally, we have the pseudo cosmology emulator, coupled with the novel theoretical framework of halo model reactions. These too, are designed to unlock the cosmological constraining power buried deep in the regime of non-linear gravitational collapse.

6.2 The future of cosmological lensing

The past and present status of cosmology and weak lensing, have so far been the primary focus of discussion in this thesis. It is now high time we turned our attention to the future of these fields. In this final Section of my thesis, I provide the answers to the following: How will we continue to investigate the nature of dark matter, dark energy, and cosmological discordance? What observational data will be available to satisfy our curiosity? What role will the work presented in this thesis play, in future cosmological inquiry?

As for the question of observations, the hopes and aspirations of next-generation cosmological lensing, rests primarily on three instruments: the Large Synoptic Survey Telescope (LSST Dark Energy Science Collaboration, 2012), Euclid (Laureijs et al., 2011), and the Wide Field Infrared Survey Telescope (WFIRST; Spergel et al., 2015). Table 6.1 presents a comparison of the end-of-mission forecasts for the associated observation campaigns, alongside those of the combined KiDS-VIKING survey.

LSST is a ground-based telescope with an 8.4 m diameter, funded by the US National Science Foundation and Department of Energy. With a 3.2 Gigapixel camera, this instrument will survey 18,000 deg^2 of sky, of which $\sim$14,000 deg^2 will be useful for measuring cosmic shear, imaging about 10 billion weakly lensed galaxies in multiple optical and near-infrared bands, to unprecedented depths. This galaxy ensemble will exceed the size of the KiDS-VIKING survey at the end of its observation campaign, by a factor of $\sim$100. In addition to

Table 6.1 *A comparison of forecasts for the combined KiDS-VIKING survey, and the next-generation weak lensing surveys, LSST, Euclid and WFIRST, at the end of their observation campaigns. Values for the KiDS-VIKING galaxy density, seeing and median $z_{\rm source}$ correspond to those reported for the KV450 cosmic shear analysis (see Hildebrandt et al., 2018, and references therein). The galaxy density quoted for Euclid is the science requirement (Laureijs et al., 2011), and those of LSST and WFIRST are reported in Chang et al. (2013) and Akeson et al. (2019), respectively. The band(s) used for shape measurement is(are) highlighted in bold on the sixth row. The seeing quoted for the space-based Euclid and WFIRST are the full-width-half-maxima of the PSFs, whereas those for KiDS-VIKING and LSST are approximate atmospheric values. LSST/Euclid/WFIRST median source galaxy redshifts are reported in, respectively, Chang et al. (2013); Laureijs et al. (2011); Spergel et al. (2015).*

	KiDS(-VIKING)	LSST	Euclid	WFIRST
Ground-/space-based	Ground	Ground	Space	Space
Telescope diameter [m]	2.6	8.4	1.2	2.4
Footprint [$\times 10^3$ deg^2]	1.35	14	15	2.4
Galaxy number [billion]	0.1	$\sim$10	$\sim$1	$\sim$0.4
Galaxy density [arcmin^{-2}]	$\simeq$7	$\simeq$26	$\simeq$30	$\simeq$45
Bands	$ugri(ZYJHK_{\rm s})$	$ug\boldsymbol{ri}ZY$	$\boldsymbol{riZ}YJH$	$\boldsymbol{YJHF}$**184**
Seeing [arcsec]	$\sim$0.7	$\sim$0.7	0.16	0.12
Median $z_{\rm source}$	$\sim$0.7	$\sim$1.2	$\sim$0.8	$\sim$1.0

weak lensing, LSST will probe the nature of dark matter and dark energy with an assortment of other probes: baryon acoustic oscillations (BAOs), Type Ia supernovae, galaxy cluster abundance and strong lensing (see Sections 1.3.2-1.3.3 for further discussion).

Whereas LSST will still suffer the irreducible systematic distortion of galaxy shapes caused by Earth's atmosphere ($\sim$0.7 arcsec in magnitude), the space-based Euclid and WFIRST are limited only in this respect by the size of their instrumental PSFs, (0.16 and 0.12 arcsec respectively). The larger of the two surveys, in terms of weak lensing science goals, is Euclid, developed by the European Space Agency. This mission will acquire comparable sky coverage to LSST, using a 1.2 m telescope and 0.6 Gigapixel camera to image about 1 billion source galaxies in optical and near-infrared bands. The NASA instrument, WFIRST, on the other hand, will feature a 2.4 m telescope and 0.3 Gigapixel camera, and will survey a smaller area of sky than Euclid, 2400 deg^2, but to greater depth, collecting images of around 400 million galaxies solely in the near-infrared. Euclid and WFIRST will also probe cosmology with alternative probes, including BAOs, Type Ia supernovae and galaxy clusters.

The costs of the aforementioned surveys, in terms of both finance and working-hours invested, are extraordinary. It is therefore, paramount, that every avenue to optimally exploiting the cosmological content of the acquired data sets is explored. Of the most significant obstacles to this endeavour, however, is the inaccuracy of analytical prescriptions, for the cosmological dependence of the non-linearities in large-scale structure. Therein, lies key motivation for the research presented in this thesis: preparing methodologies for application to these future surveys, with aims to better understand this illusive side of the Universe.

With the feasibility of clipping for enhancing cosmic shear, having been demonstrated conclusively for the first time with observations in Chapter 3, clipping is thus collated with the other non-Gaussian statistics mentioned in Section 6.1 (peak abundance, high-order statistics, Minkowski functionals and density splitting), as prominent methodologies for tangibly improving our grasp of cosmological information beyond standard two-point functions, such as the power spectrum and correlation functions. Applying these techniques to the data from LSST, Euclid or WFIRST, however, requires they are tried and tested with the currently available observations.

Herein the *cosmo*-SLICS suite and emulator come to the fore, providing practical

and straight-forward means to measure the signal of these various probes in the simulations, and subsequently interpolate to arbitrary cosmologies for comparison with the data (as shown in Chapter 4). The $\lesssim 5\%$ accuracy of these resources does not permit *direct* application to the next-generation lensing data, rather, their utility is in aiding the development of analysis pipelines and identifying potential unforeseen systematics in any given method, well in advance of the first-light of each instrument. In addition, they readily facilitate an investigation of the *combined* constraining power of these probes, which may highlight an optimal route to take forward for next-generation cosmic shear analyses. Furthermore, Figures 4.3-4.4 in this Chapter demonstrated significant gains in interpolation accuracy should the *cosmo*-SLICS be augmented in size by a factor of two. It is plausible that such an extension to the suite, may yield model predictions which are sufficiently precise for these resources to be used alongside LSST, Euclid or WFIRST data.

Whilst striving to enable analysis with non-Gaussian probes, I have also sought to improve the standard procedure to constraining cosmological constraints: employing two-point probes such as the matter power spectrum. This is, after all, the road most-travelled, through which the most precise inference of model parameters and progress in understanding systematics, has been made. In forecasting the required simulated resources for emulating pseudo cosmology matter power spectra (Chapter 5), the range of input cosmologies was specifically designed to facilitate constraining the data collected by the next-generation surveys. Though the requirements are great, I argue that the pay-off is greater: finally, a robust "one-stop-shop" to hold not only the standard model, but practically *any* cosmological model, up to the harsh light of observations.

If there is but one conclusion I would have the reader take from my doctoral research, it would be that there is no single road I would advocate to future cosmological parameter constraints. Theoretical modelling, numerical simulations and machine learning, are all promising methods to employ, but in my experience, the best results are obtained when their powers are combined, as demonstrated in Chapters 3-5 of this thesis. In regimes where theoretical arguments generate poor predictions, simulations offer accurate ones. Where simulations may be inflexible, in terms of sampling cosmological models and their associated input parameters, emulators facilitate generalisation to arbitrary input. Where machine learning may seem opaque, theoretical modelling provides interpretability. In this fashion, this trinity of methods effectively compensate for their individual

shortfalls, thereby yielding enhanced prospects for both non-Gaussian (Chapters 3 and 4) *and* the standard two-point (Chapter 5) statistical probes. To my mind, therefore, the most accurate and precise future cosmological constraints, will be reached with the combined efforts of theorists, simulators and computer scientists.

On the subject of extraterrestrial life, Sir Arthur C. Clarke, once said *"Two possibilities exist: either we are alone in the Universe or we are not. Both are equally terrifying."* It is with similar mindset, one might consider the future of cosmology. Either the next generation of surveys will reveal an incompleteness in the standard model, and the century of theoretical and observational work underpinning it, or they will not. Perhaps Einstein and his disciples were right all along, and indeed we must face facts: some 96% of the energy density truly *is* the elusive, headline-grabbing dark stuff; the early Universe actually *was* inflated by an as-of-yet unknown energy source. *Or*, new physics, maybe even an entirely new description of cosmological gravity, is out there, waiting to be discovered. I believe it will be with a mixture of terror and exhilaration, that the cosmological community will embrace the outcome.

Bibliography

Abbott B. P., et al., 2016, Phys. Rev. Lett., 116, 221101

Abbott B. P., et al., 2017, The Astrophysical Journal, 848, L12

Agarwal S., Abdalla F. B., Feldman H. A., Lahav O., Thomas S. A., 2012, MNRAS, 424, 1409

Akeson R., et al., 2019, preprint, (arXiv:1902.05569)

Alam S., Ho S., Silvestri A., 2016, MNRAS, 456, 3743

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409

Almosallam I. A., Jarvis M. J., Roberts S. J., 2016, MNRAS, 462, 726

Alsing J., Heavens A., Jaffe A. H., Kiessling A., Wandelt B., Hoffmann T., 2016, MNRAS, 455, 4452

Angulo R. E., Pontzen A., 2016, MNRAS, 462, L1

Archidiacono M., Giusarma E., Hannestad S., Mena O., 2013, preprint, (arXiv:1307.0637)

Auld T., Bridges M., Hobson M. P., Gull S. F., 2007, Monthly Notices of the Royal Astronomical Society: Letters, 376, L11

Aviles A., Cervantes-Cota J. L., 2017, Phys. Rev. D, 96, 123526

Bahcall N. A., Cen R., 1992, ApJ, 398, L81

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15

Bartelmann M., Maturi M., 2017, Scholarpedia, 12, 32440

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

Battinelli P., Demers S., 2006, A&A, 447, 473

Baumann D., Green D., Wallisch B., 2018, J. Cosmology Astropart. Phys., 8, 029

Ben-Dayan I., Durrer R., Marozzi G., Schwarz D. J., 2014, Phys. Rev. Lett., 112, 221301

Benítez N., 2000, ApJ, 536, 571

Benjamin J., van Waerbeke L., Ménard B., Kilbinger M., 2010, MNRAS, 408, 1168

Benjamin J., et al., 2013, MNRAS, 431, 1547

Bergé J., Price S., Amara A., Rhodes J., 2012, MNRAS, 419, 2356

Bernardeau F., 2005, A&A, 441, 873

Bernstein G. M., Armstrong R., 2014, MNRAS, 438, 1880

Bernstein G. M., Jarvis M., 2002, AJ, 123, 583

Bernstein G. M., Armstrong R., Krawiec C., March M. C., 2016, MNRAS, 459, 4467

Bertin E., Arnouts S., 1996, A&AS, 117, 393

Betoule M., et al., 2014, A&A, 568, A22

Bird S., Viel M., Haehnelt M. G., 2012, MNRAS, 420, 2551

Blas D., Garny M., Konstandin T., Lesgourgues J., 2014, Journal of Cosmology and Astro-Particle Physics, 2014, 039

Blatman G., Sudret B., 2011, Journal of Computational Physics, 230, 2345

Bonamente M., Joy M. K., LaRoque S. J., Carlstrom J. E., Reese E. D., Dawson K. S., 2006, ApJ, 647, 25

Bond J. R., Efstathiou G., 1984, ApJ, 285, L45

Bonnet H., Mellier Y., 1995, A&A, 303, 331

Bonnett C., 2015, MNRAS, 449, 1043

Bose B., Koyama K., Lewandowski M., Vernizzi F., Winther H. A., 2018, Journal of Cosmology and Astro-Particle Physics, 2018, 063

Brainerd T. G., Blandford R. D., Smail I., 1996, in American Astronomical Society Meeting Abstracts #188. p. 13.02

Brandt T. D., 2016, The Astrophysical Journal, 824, L31

Brax P., Valageas P., 2013, Phys. Rev. D, 88, 023527

Brax P., Davis A.-C., Li B., Winther H. A., 2012, Phys. Rev. D, 86, 044015

Bridle S., King L., 2007, New Journal of Physics, 9, 444

Bridle S., et al., 2009.

Brouwer M. M., et al., 2018, MNRAS, 481, 5189

Carleo G., Cirac I., Cranmer K., Daudet L., Schuld M., Tishby N., Vogt-Maranto L., Zdeborová L., 2019, preprint, (arXiv:1903.10563)

Caron S., Heskes T., Otten S., Stienen B., 2019, preprint, (arXiv:1905.08628)

Carroll S. M., Press W. H., Turner E. L., 1992, ARA&A, 30, 499

Cataneo M., et al., 2015, Phys. Rev. D, 92, 044009

Cataneo M., Lombriser L., Heymans C., Mead A., Barreira A., Bose S., Li B., 2018, preprint, (arXiv:1812.05594)

Catelan P., Kamionkowski M., Blandford R. D., 2001, Monthly Notices of the Royal Astronomical Society, 320, L7

Chang C., et al., 2013, Monthly Notices of the Royal Astronomical Society, 434, 2121

Chang C., et al., 2017, preprint, (arXiv:1708.01535)

Chisari N. E., et al., 2017, MNRAS, 472, 1163

Chisari N. E., et al., 2018, MNRAS, 480, 3962

Clowe D., Bradač M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, ApJ, 648, L109

Collister A., Lahav O., 2003, Publications of the Astronomical Society of the Pacific, 116

Copeland D., Taylor A., Hall A., 2018, MNRAS, 480, 2247

Croft R. A. C., Metzler C. A., 2000, ApJ, 545, 561

Currin C., Mitchell T., Morris M., Ylvisaker D., 1991, Journal of the American Statistical Association, 86, 953

Cyr-Racine F.-Y., Sigurdson K., Zavala J., Bringmann T., Vogelsberger M., Pfrommer C., 2016, Phys. Rev. D, 93, 123527

D'Aloisio A., Natarajan P., 2011, MNRAS, 411, 1628

DES Collaboration et al., 2017, preprint, (arXiv:1708.01530)

DES Collaboration et al., 2018, preprint, (arXiv:1810.02499)

Dakin J., Hannestad S., Tram T., 2019, preprint, (arXiv:1904.11773)

De Felice A., Tsujikawa S., 2010, Journal of Cosmology and Astro-Particle Physics, 2010, 024

Dietrich J. P., Hartlap J., 2010, MNRAS, 402, 1049

Doroshkevich A. G., Zel'dovich Y. B., Syunyaev R. A., 1978, Soviet Ast., 22, 523

Dubois Y., et al., 2014, MNRAS, 444, 1453

Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, Monthly Notices of the Royal Astronomical Society, 463, 3948

Efstathiou G., 2014, Monthly Notices of the Royal Astronomical Society, 440, 1138

Eifler T., Schneider P., Hartlap J., 2009, A&A, 502, 721

Einstein A., 1915, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 844-847.,

Eisenstein D. J., Hu W., 1998, ApJ, 496, 605

Erben T., et al., 2005, Astronomische Nachrichten, 326, 432

Euclid Collaboration et al., 2018, preprint, (arXiv:1809.04695)

Fairbairn M., Marsh D. J. E., Quevillon J., 2017, Phys. Rev. Lett., 119, 021101

Fenech Conti I., Herbonnet R., Hoekstra H., Merten J., Miller L., Viola M., 2017, MNRAS, 467, 1627

Fixsen D. J., 2009, ApJ, 707, 916

Fluri J., Kacprzak T., Lucchi A., Refregier A., Amara A., Hofmann T., Schneider A., 2019, preprint, (arXiv:1906.03156)

Folatelli G., et al., 2010, AJ, 139, 120

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Forero D. V., Tórtola M., Valle J. W. F., 2014, Phys. Rev. D, 90, 093006

Frieman J., Dark Energy Survey Collaboration 2013, in American Astronomical Society Meeting Abstracts #221. p. 335.01

Fu L., et al., 2014, MNRAS, 441, 2725

Gaia Collaboration et al., 2018, A&A, 616, A1

Gao F., et al., 2016, ApJ, 817, 128

Garrison L. H., Eisenstein D. J., Ferrer D., Tinker J. L., Pinto P. A., Weinberg D. H., 2018, ApJS, 236, 43

Gentile M., Courbin F., Meylan G., 2013, A&A, 549, A1

Giblin B., et al., 2018, MNRAS, 480, 5529

Giblin B., Cataneo M., Moews B., Heymans C., 2019, Monthly Notices of the Royal Astronomical Society

Gil-Marín H., et al., 2016, MNRAS, 460, 4188

Gil-Marín H., et al., 2018, MNRAS, 477, 1604

Green A. M., 2016, Phys. Rev. D, 94, 063530

Gruen D., et al., 2018, Phys. Rev. D, 98, 023507

Gupta A., Matilla J. M. Z., Hsu D., Haiman Z., 2018, Phys. Rev. D, 97, 103515

Habib S., Heitmann K., Higdon D., Nakhleh C., Williams B., 2007, Phys. Rev. D, 76, 083503

Hamana T., et al., 2019, preprint, (arXiv:1906.06041)

Hamuy M., Phillips M. M., Suntzeff N. B., Schommer R. A., Maza J., Aviles R., 1996, AJ, 112, 2391

Handley W., Lemos P., 2019, preprint, (arXiv:1902.04029)

Harnois-Déraps J., van Waerbeke L., 2015, MNRAS, 450, 2857

Harnois-Déraps J., Pen U.-L., Iliev I. T., Merz H., Emberson J. D., Desjacques V., 2013, MNRAS, 436, 540

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

Harnois-Deraps J., Giblin B., Joachimi B., 2019, preprint, (arXiv:1905.06454)

Harrison E. R., 1970, Phys. Rev. D, 1, 2726

Hartlap J., Simon P., Schneider P., 2007, A&A, 464, 399

Hatt D., et al., 2017, ApJ, 845, 146

Heavens A., Refregier A., Heymans C., 2000, MNRAS, 319, 649

Heisenberg L., Bartelmann M., 2019, preprint, (arXiv:1901.01041)

Heitmann K., Higdon D., Nakhleh C., Habib S., 2006, ApJ, 646, L1

Heitmann K., Higdon D., White M., Habib S., Williams B. J., Lawrence E., Wagner C., 2009, ApJ, 705, 156

Heitmann K., White M., Wagner C., Habib S., Higdon D., 2010, ApJ, 715, 104

Heitmann K., Lawrence E., Kwan J., Habib S., Higdon D., 2014, ApJ, 780, 111

Heitmann K., et al., 2016, ApJ, 820, 108

Heymans C., Heavens A., 2003, preprint, (arXiv:astro-ph/0310495)

Heymans C., Zhao G.-B., 2018, International Journal of Modern Physics D, 27, 1848005

Heymans C., et al., 2013, MNRAS, 432, 2433

Hikage C., et al., 2018, preprint, (arXiv:1809.09148)

Hilbert S., Xu D., Schneider P., Springel V., Vogelsberger M., Hernquist L., 2017, MNRAS, 468, 790

Hildebrandt H., et al., 2010, A&A, 523, A31

Hildebrandt H., et al., 2012, MNRAS, 421, 2355

Hildebrandt H., et al., 2017, MNRAS, 465, 1454

Hildebrandt H., et al., 2018, preprint, (arXiv:1812.06076)

Hinshaw G., et al., 2013, ApJS, 208, 19

Hinterbichler K., Khoury J., Levy A., Matas A., 2011, Phys. Rev. D, 84, 103521

Hirata C. M., Seljak U., 2004, Phys. Rev. D, 70, 063526

Hirata C. M., et al., 2004, MNRAS, 353, 529

Hirata C. M., Mandelbaum R., Ishak M., Seljak U., Nichol R., Pimbblet K. A., Ross N. P., Wake D., 2007, Monthly Notices of the Royal Astronomical Society, 381, 1197

Hložek R., Marsh D. J. E., Grin D., Allison R., Dunkley J., Calabrese E., 2017, Phys. Rev. D, 95, 123511

Hoekstra H., Franx M., Kuijken K., Squires G., 1998, New Astron. Rev., 42, 137

Hoessel J. G., Gunn J. E., Thuan T. X., 1980, ApJ, 241, 486

Hogg D. W., 1999, preprint, (arXiv:astro-ph/9905116)

Hojjati A., Pogosian L., Zhao G.-B., 2011, Journal of Cosmology and Astro-Particle Physics, 2011, 005

Hoyle B., et al., 2018, MNRAS, 478, 592

Hu W., Dodelson S., 2002, ARA&A, 40, 171

Hu W., Sawicki I., 2007, Phys. Rev. D, 76, 064004

Hu W., Sugiyama N., 1996, ApJ, 471, 542

Hu W., White M., 1996, ApJ, 471, 30

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Huff E., Mandelbaum R., 2017, preprint, (arXiv:1702.02600)

Huterer D., Shafer D. L., 2018, Reports on Progress in Physics, 81, 016901

Huterer D., et al., 2015, Astroparticle Physics, 63, 23

Ilbert O., et al., 2006, A&A, 457, 841

Jain B., Van Waerbeke L., 2000, ApJ, 530, L1

Jarvis M., 2015, TreeCorr: Two-point correlation functions, Astrophysics Source Code Library (ascl:1508.007)

Jarvis M., Bernstein G., Jain B., 2004, MNRAS, 352, 338

Jarvis M., et al., 2016, MNRAS, 460, 2245

Joachimi B., Mandelbaum R., Abdalla F. B., Bridle S. L., 2011, A&A, 527, A26

Johnson A., et al., 2017, MNRAS, 465, 4118

Jones B. J. T., 2017, Precision Cosmology: The First Half Million Years. Cambridge University Press, doi:10.1017/CBO9781139027809

Joseph V. R., Gul E., Ba S., 2015, Biometrika, 102, 371

Joudaki S., et al., 2017, MNRAS, 471, 1259

Joudaki S., et al., 2018, MNRAS, 474, 4894

Joudaki S., et al., 2019, preprint, (arXiv:1906.09262)

Joyce A., Jain B., Khoury J., Trodden M., 2015, Phys. Rep., 568, 1

Joyce A., Lombriser L., Schmidt F., 2016, Annual Review of Nuclear and Particle Science, 66, 95

Jullo E., Natarajan P., Kneib J. P., D'Aloisio A., Limousin M., Richard J., Schimd C., 2010, Science, 329, 924

Jullo E., Pires S., Jauzac M., Kneib J.-P., 2014, MNRAS, 437, 3969

Kacprzak T., Zuntz J., Rowe B., Bridle S., Refregier A., Amara A., Voigt L., Hirsch M., 2012, MNRAS, 427, 2711

Kacprzak T., Bridle S., Rowe B., Voigt L., Zuntz J., Hirsch M., MacCrann N., 2014, MNRAS, 441, 2528

Kacprzak T., et al., 2016, MNRAS, 463, 3653

Kaiser N., 1987, MNRAS, 227, 1

Kaiser N., 1992, ApJ, 388, 272

Kaiser N., Squires G., 1993, ApJ, 404, 441

Kaiser N., Squires G., Broadhurst T., 1995, ApJ, 449, 460

Kamionkowski M., Spergel D. N., Sugiyama N., 1994, ApJ, 426, L57

Kannawadi A., et al., 2019, A&A, 624, A92

Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, MNRAS, 450, 1349

Kilbinger M., 2015, Reports on Progress in Physics, 78, 086901

Kilbinger M., Schneider P., 2005, A&A, 442, 69

Kilbinger M., et al., 2009, A&A, 497, 677

Kilbinger M., et al., 2017a, MNRAS, 472, 2126

Kilbinger M., et al., 2017b, MNRAS, 472, 2126

Kim A. G., et al., 2013, ApJ, 766, 84

King L., Schneider P., 2002, A&A, 396, 411

Kitching T. D., et al., 2013, ApJS, 205, 12

Kitching T. D., et al., 2014, MNRAS, 442, 1326

Kochanek C. S., 2003, ApJ, 583, 49

Köhlinger F., et al., 2017, preprint, (arXiv:1706.02892)

Kratochvil J. M., Haiman Z., May M., 2010, Phys. Rev. D, 81, 043519

Kristian J., Sachs R. K., 1966, ApJ, 143, 379

Kuijken K., et al., 2015, MNRAS, 454, 3500

Kuijken K., et al., 2019, A&A, 625, A2

Kwan J., Bhattacharya S., Heitmann K., Habib S., 2013, ApJ, 768, 123

LSST Dark Energy Science Collaboration 2012, preprint, (arXiv:1211.0310)

Landy S. D., Szalay A. S., 1993, ApJ, 412, 64

Lapuschkin S., Wäldchen S., Binder A. e., Montavon G., Samek W., Müller K.-R., 2019, Nature Communications, 10, 1096

Lattanzi M., Gerbino M., 2017, preprint, (arXiv:1712.07109)

Laureijs R., et al., 2011, preprint, (arXiv:1110.3193)

Lawrence E., Heitmann K., White M., Higdon D., Wagner C., Habib S., Williams B., 2010, ApJ, 713, 1322

Lawrence E., et al., 2017, preprint, (arXiv:1705.03388)

Lee J., Pen U.-L., 2000, ApJ, 532, L5

Leistedt B., Mortlock D. J., Peiris H. V., 2016, MNRAS, 460, 4258

Lesgourgues J., Pastor S., 2006, Phys. Rep., 429, 307

Lesgourgues J., Matarrese S., Pietroni M., Riotto A., 2009, Journal of Cosmology and Astro-Particle Physics, 2009, 017

Levi M., Vlah Z., 2016, preprint, (arXiv:1605.09417)

Levi M., et al., 2013, preprint, (arXiv:1308.0847)

Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473

Li C., Jing Y. P., Faltenbacher A., Wang J., 2013, ApJ, 770, L12

Liddle A. R., 1998, An introduction to modern cosmology

Lima M., Cunha C. E., Oyaizu H., Frieman J., Lin H., Sheldon E. S., 2008, MNRAS, 390, 118

Limber D. N., 1953, ApJ, 117, 134

Liu X., Wang Q., Pan C., Fan Z., 2014, ApJ, 784, 31

Liu J., Petri A., Haiman Z., Hui L., Kratochvil J. M., May M., 2015, Phys. Rev. D, 91, 063507

Liu X., et al., 2016, Physical Review Letters, 117, 051101

Liu J., Bird S., Zorrilla Matilla J. M., Hill J. C., Haiman Z., Madhavacheril M. S., Petri A., Spergel D. N., 2018, J. Cosmology Astropart. Phys., 3, 049

Lombriser L., 2014, Annalen der Physik, 264, 259

Lombriser L., Simpson F., Mead A., 2015, Physical Review Letters, 114, 251101

Lu T., Zhang J., Dong F., Li Y., Liu D., Fu L., Li G., Fan Z., 2017, AJ, 153, 197

Luppino G. A., Kaiser N., 1997, The Astrophysical Journal, 475, 20

Mandelbaum R., 2018, Annual Review of Astronomy and Astrophysics, 56, 393

Mandelbaum R., Hyper Suprime-Cam (HSC) Collaboration 2017, in American Astronomical Society Meeting Abstracts #229. p. 226.02

Mandelbaum R., Hirata C. M., Ishak M., Seljak U., Brinkmann J., 2006, Monthly Notices of the Royal Astronomical Society, 367, 611

Mangano G., Miele G., Pastor S., Pinto T., Pisanti O., Serpico P. D., 2005, Nuclear Physics B, 729, 221

Marian L., Smith R. E., Hilbert S., Schneider P., 2012, Monthly Notices of the Royal Astronomical Society, 423, 1711

Marian L., Smith R. E., Hilbert S., Schneider P., 2013, Monthly Notices of the Royal Astronomical Society, 432, 1338

Marsh D. J. E., 2016, preprint, (arXiv:1605.05973)

Marsh D. J. E., Silk J., 2014, MNRAS, 437, 2652

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

Martinet N., et al., 2018, MNRAS, 474, 712

Massara E., Villaescusa-Navarro F., Viel M., 2014, Journal of Cosmology and Astro-Particle Physics, 2014, 053

Massey R., et al., 2007, MNRAS, 376, 13

McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936

McClintock T., et al., 2019, ApJ, 872, 53

McKay M. D., Beckman R. J., Conover W. J., 1979, Technometrics, 21, 239

McQuinn M., White M., 2013, MNRAS, 433, 2857

Mead A. J., 2017, MNRAS, 464, 1282

Mead A. J., Peacock J. A., Heymans C., Joudaki S., Heavens A. F., 2015, MNRAS, 454, 1958

Mehta K. T., Seo H.-J., Eckel J., Eisenstein D. J., Metchnik M., Pinto P., Xu X., 2011, ApJ, 734, 94

Melchior P., Viola M., 2012, MNRAS, 424, 2757

Ménard B., Scranton R., Schmidt S., Morrison C., Jeong D., Budavari T., Rahman M., 2013, preprint, (arXiv:1303.4722)

Miller L., Kitching T. D., Heymans C., Heavens A. F., van Waerbeke L., 2007, MNRAS, 382, 315

Miller L., et al., 2013, MNRAS, 429, 2858

Moews B., de Souza R. S., Ishida E. E. O., Malz A. I., Heneka C., Vilalta R., Zuntz J., 2018, preprint (arXiv:1812.09786)

Morris M.-D., Mitchell T. J., 1995, J. Stat. Plan. Inference, 43, 381

Mörtsell E., Dhawan S., 2018, J. Cosmology Astropart. Phys., 9, 025

Neal M. R., 1994, Bayesian Learning for Neural Networks PhD thesis. Lect. Notes Statist. Vol. 118, doi:10.1007/978-1-4612-0745-0_2,

Newman J. A., 2008, ApJ, 684, 88

Newman J. A., et al., 2015, Astroparticle Physics, 63, 81

Neyrinck M. C., Szapudi I., Szalay A. S., 2009, The Astrophysical Journal Letters, 698, L90

Niikura H., et al., 2019, Nature Astronomy, p. 238

Nori M., Baldi M., 2018, MNRAS, 478, 3935

Oesch P. A., et al., 2016, ApJ, 819, 129

Padmanabhan N., White M., 2009, Phys. Rev. D, 80, 063508

Panagia N., 1999, in Chu Y.-H., Suntzeff N., Hesser J., Bohlender D., eds, IAU Symposium Vol. 190, New Views of the Magellanic Clouds. p. 549

Park M., Zurek K. M., Watson S., 2010, Phys. Rev. D, 81, 124008

Parkinson D., et al., 2012, Phys. Rev. D, 86, 103518

Paulin-Henriksson S., Amara A., Voigt L., Refregier A., Bridle S. L., 2008, A&A, 484, 67

Peacock J. A., 1999, Cosmological Physics

Peacock J. A., Dodds S. J., 1994, MNRAS, 267, 1020

Pedregosa F., et al., 2011, Journal of Machine Learning Research, 12, 2825

Peebles P. J. E., 1980, The large-scale structure of the universe

Peebles P. J. E., Yu J. T., 1970, ApJ, 162, 815

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Perlmutter S., et al., 1999, ApJ, 517, 565

Petri A., Liu J., Haiman Z., May M., Hui L., Kratochvil J. M., 2015, Phys. Rev. D, 91, 103511

Petrillo C. E., et al., 2019, MNRAS, 484, 3879

Phillips M. M., 1993, ApJ, 413, L105

Planck Collaboration et al., 2016a, A&A, 594, A13

Planck Collaboration et al., 2016b, A&A, 594, A27

Planck Collaboration et al., 2018, preprint, (`arXiv:1807.06209`)

Pogosian L., Silvestri A., 2016, Phys. Rev. D, 94, 104014

Poulin V., Serpico P. D., Lesgourgues J., 2016, Journal of Cosmology and Astro-Particle Physics, 2016, 036

Prat J., et al., 2019, MNRAS, 487, 1363

Press W. H., Schechter P., 1974, ApJ, 187, 425

Quiros I., García-Salcedo R., Gonzalez T., Horta-Rangel F. A., Saavedra J., 2016, European Journal of Physics, 37, 055605

Rasmussen C., Williams C., 2006, Gaussian Processes for Machine Learning. Adaptive Computation and Machine Learning, MIT Press, Cambridge, MA, USA

Rau M. M., Wilson S., Mandelbaum R., 2019, preprint, (arXiv:1904.09988)

Refregier A., Kacprzak T., Amara A., Bridle S., Rowe B., 2012, MNRAS, 425, 1951

Ribli D., Dobos L., Csabai I., 2019, preprint, (arXiv:1902.08161)

Riess A. G., et al., 1998, AJ, 116, 1009

Riess A. G., et al., 2011, ApJ, 730, 119

Riess A. G., et al., 2016, ApJ, 826, 56

Riess A. G., et al., 2018a, ApJ, 855, 136

Riess A. G., et al., 2018b, ApJ, 861, 126

Riess A. G., Casertano S., Yuan W., Macri L. M., Scolnic D., 2019, ApJ, 876, 85

Rogers K. K., Peiris H. V., Pontzen A., Bird S., Verde L., Font-Ribera A., 2019, Journal of Cosmology and Astro-Particle Physics, 2019, 031

Ross A. J., et al., 2017, MNRAS, 464, 1168

Rowe B., 2010, MNRAS, 404, 350

Rozo E., et al., 2010, ApJ, 708, 645

Sacks J., Welch W. J., Mitchell T. J., Wynn H. P., 1989, Statist. Sci., 4, 409

Sadeh I., Abdalla F. B., Lahav O., 2016, Publications of the Astronomical Society of the Pacific, 128, 104502

Salvador A. I., et al., 2019, MNRAS, 482, 1435

Schaye J., et al., 2015, MNRAS, 446, 521

Schirmer M., 2013, ApJS, 209, 21

Schmidt S. J., Ménard B., Scranton R., Morrison C., McBride C. K., 2013, MNRAS, 431, 3307

Schneider P., 1996, MNRAS, 283, 837

Schneider P., 2005, preprint, (arXiv:astro-ph/0509252)

Schneider A., 2015, MNRAS, 451, 3117

Schneider P., van Waerbeke L., Jain B., Kruse G., 1998, MNRAS, 296, 873

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

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

Schneider P., van Waerbeke L., Kilbinger M., Mellier Y., 2002c, A&A, 396, 1

Schneider M. D., Knox L., Habib S., Heitmann K., Higdon D., Nakhleh C., 2008, Phys. Rev. D, 78, 063529

Schneider M., Frenk C. S., Cole S., 2012, in American Astronomical Society Meeting Abstracts #219. p. 248.09

Schneider M. D., Hogg D. W., Marshall P. J., Dawson W. A., Meyers J., Bard D. J., Lang D., 2015, ApJ, 807, 87

Schneider A., Teyssier R., Stadel J., Chisari N. E., Le Brun A. M. C., Amara A., Refregier A., 2019, Journal of Cosmology and Astro-Particle Physics, 2019, 020

Schrabback T., et al., 2018, MNRAS, 474, 2635

Schwarzschild K., 1916, Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.), 1916, 189

Scoccimarro R., 2004, Phys. Rev. D, 70, 083007

Sehgal N., et al., 2011, ApJ, 732, 44

Seitz S., Schneider P., 1996, A&A, 305, 383

Seitz C., Schneider P., 1997, A&A, 318, 687

Sellentin E., Heavens A. F., 2016, MNRAS, 456, L132

Sellentin E., Heymans C., Harnois-Déraps J., 2017, preprint, (arXiv:1712.04923)

Semboloni E., Schrabback T., van Waerbeke L., Vafaei S., Hartlap J., Hilbert S., 2011a, MNRAS, 410, 143

Semboloni E., Hoekstra H., Schaye J., van Daalen M. P., McCarthy I. G., 2011b, MNRAS, 417, 2020

Senatore L., Zaldarriaga M., 2017, preprint, (arXiv:1707.04698)

Seo H.-J., Sato M., Dodelson S., Jain B., Takada M., 2011, ApJ, 729, L11

Shan H., et al., 2018, MNRAS, 474, 1116

Sheldon E., 2015, NGMIX: Gaussian mixture models for 2D images (ascl:1508.008)

Sheldon E. S., Huff E. M., 2017, ApJ, 841, 24

Sheth R. K., Tormen G., 1999, MNRAS, 308, 119

Sheth R. K., Tormen G., 2002, MNRAS, 329, 61

Simpson F., James J. B., Heavens A. F., Heymans C., 2011, Physical Review Letters, 107, 271301

Simpson F., Heavens A. F., Heymans C., 2013, Phys. Rev. D, 88, 083510

Simpson F., et al., 2016a, Phys. Rev. D, 93, 023525

Simpson F., Harnois-Déraps J., Heymans C., Jimenez R., Joachimi B., Verde L., 2016b, MNRAS, 456, 278

Simpson F., Jimenez R., Pena-Garay C., Verde L., 2017, J. Cosmology Astropart. Phys., 2017, 029

Singh S., Mandelbaum R., More S., 2015, Monthly Notices of the Royal Astronomical Society, 450, 2195

Slipher V. M., 1917, Proceedings of the American Philosophical Society, 56, 403

Smith R. E., et al., 2003, MNRAS, 341, 1311

Sorce J. G., Tully R. B., Courtois H. M., 2012, The Astrophysical Journal, 758, L12

Spergel D., et al., 2015, preprint, (arXiv:1503.03757)

Springel V., 2005, MNRAS, 364, 1105

Starobinskii A. A., 1985, Soviet Astronomy Letters, 11, 133

Strubell E., Ganesh A., McCallum A., 2019, preprint, (arXiv:1906.02243)

Sugiyama N., 1995, ApJS, 100, 281

Sunyaev R. A., Zeldovich Y. B., 1970, Ap&SS, 7, 3

Suyu S. H., et al., 2013, ApJ, 766, 70

Takada M., Jain B., 2002, MNRAS, 337, 875

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

Taruya A., Nishimichi T., Saito S., 2010, Phys. Rev. D, 82, 063522

Taylor P. L., Kitching T. D., McEwen J. D., 2018, Phys. Rev. D, 98, 043532

Tenneti A., Mandelbaum R., Di Matteo T., 2016, MNRAS, 462, 2668

Terukina A., Lombriser L., Yamamoto K., Bacon D., Koyama K., Nichol R. C., 2014, Journal of Cosmology and Astroparticle Physics, 2014, 013

Thomas D. B., Kopp M., Markovič K., 2019, preprint, (arXiv:1905.02739)

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709

Tröster T., Ferguson C., Harnois-Déraps J., McCarthy I. G., 2019, MNRAS, 487, L24

Troxel M. A., et al., 2017, preprint, (arXiv:1708.01538)

Troxel M. A., et al., 2018, preprint, (arXiv:1804.10663)

Tyson J. A., Valdes F., Jarvis J. F., Mills A. P. J., 1984, ApJ, 281, L59

Valdes F., Tyson J. A., Jarvis J. F., 1983, ApJ, 271, 431

Van Waerbeke L., et al., 2013, MNRAS, 433, 3373

VanderPlas J. T., Connolly A. J., Jain B., Jarvis M., 2012, ApJ, 744, 180

Vanzella E., et al., 2004, A&A, 423, 761

Velliscig M., et al., 2015, MNRAS, 454, 3328

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

Vogelsberger M., et al., 2014, MNRAS, 444, 1518

Voigt L. M., Bridle S. L., 2010, MNRAS, 404, 458

Wang J., Hui L., Khoury J., 2012, Phys. Rev. Lett., 109, 241301

Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, Phys. Rep., 530, 87

White S. D. M., Davis M., Frenk C. S., 1984, MNRAS, 209, 27P

White S., Navarro J., Evrard A., Frenk C., 1993, Nature, 366

Wilson M. J., 2016, preprint, (arXiv:1610.08362)

Winther H., Casas S., Baldi M., Koyama K., Li B., Lombriser L., Zhao G.-B., 2019, preprint, (arXiv:1903.08798)

Yoon M., Jee M. J., Tyson J. A., Schmidt S., Wittman D., Choi A., 2018, preprint, (arXiv:1807.09195)

Zel'dovich Y. B., 1970, A&A, 5, 84

Zentner A. R., Semboloni E., Dodelson S., Eifler T., Krause E., Hearin A. P., 2013, Phys. Rev. D, 87, 043509

Zhai Z., et al., 2019, ApJ, 874, 95

Zhao G.-B., 2014, ApJS, 211, 23

Zhao G.-B., Pogosian L., Silvestri A., Zylberberg J., 2009, Phys. Rev. D, 79, 083513

Zumalacárregui M., Bellini E., Sawicki I., Lesgourgues J., Ferreira P. G., 2017, Journal of Cosmology and Astro-Particle Physics, 2017, 019

Zuntz J., et al., 2018, MNRAS, 481, 1149

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

de Jong J. T. A., Verdoes Kleijn G. A., Kuijken K. H., Valentijn E. A., 2013, Experimental Astronomy, 35, 25

de Jong J. T. A., et al., 2017, A&A, 604, A134

van Uitert E., et al., 2018, MNRAS, 476, 4662

van Waerbeke L., 1998, A&A, 334, 1

von der Linden A., et al., 2014, MNRAS, 443, 1973