Cosmic tomography with weak gravitational lensing

Li, S.S.

CHAPTER 1

Introduction

Approximately 13.8 billion years ago (Planck Collaboration et al. 2020), an event known as the 'Big Bang' marked the inception of the physical Universe. As the Universe continues to expand, its energy density decreases, leading to the separation of the four fundamental forces and the cooling of particles. Eventually, the temperature fell to such a level that the initial density fluctuations could grow by gravitationally attracting surrounding matter, resulting in the formation of gas clouds, stars, galaxies, and the large-scale structures we observe today.

Unveiling this big picture of the history of our Universe, as depicted in Figure 1.1, stands as one of the most significant accomplishments of twentieth-century cosmological research, if not the most significant. The theoretical and observational advancements in the field are truly remarkable. This is especially evident with the establishment of the spatially flat Λ Cold Dark Matter (ΛCDM) model, which hinges on only six parameters and yet successfully describes several key observations with astounding accuracy, including the expansion of the Universe, the measurements of the Cosmic Microwave Background (CMB), and the matter distribution of cosmic large-scale structures.

As we move into the 21st century, the field of cosmology continues to thrive, intertwined with advancements in observational techniques and the subsequent challenges in accurately interpreting increasingly precise measurements. On the one hand, the development of various cosmological probes, facilitated by modern technology, yields unprecedented precision in our measurements. This progress provides a massive opportunity to unravel the mysteries surrounding some critical elements of our current cosmological model, such as the origins of initial density fluctuations and the enigmatic nature of dark matter and dark energy. On the other hand, the technical advancements come with the challenging task of controlling various systematic effects that stem from both observational conditions and astrophysical contamination. Therefore, devising effective strategies to control these systematics, in order to harness the full statistical power of advanced cosmological probes, remains a pivotal topic in 21st-century cosmological studies.

Among the techniques that show promise as cosmological probes but present practical challenges is weak gravitational lensing. According to general relativity, light

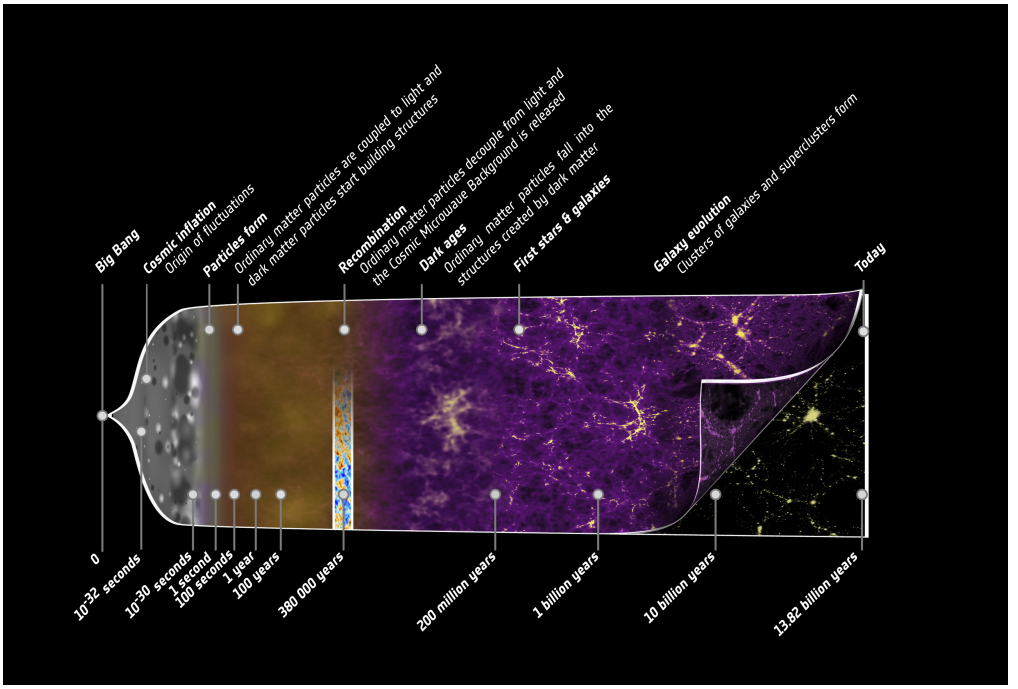

Figure 1.1: An overview of the history of our Universe, from its initial state—characterised by near uniformity, punctuated by small fluctuations—to the complex cosmic structure we observe today. [Credit: ESA – C. Carreau]

from distant galaxies is distorted by the gravitational fields associated with foreground matter. Consequently, measuring these small yet coherent distortions offers a direct probe into the matter distribution in large-scale structures, or those associated with individual massive objects, such as galaxy groups or clusters. However, in practice, robustly measuring these small lensing-induced distortions is difficult due to distortions introduced by observational conditions and instrumental effects. The task is further complicated by the challenges involved in accurately determining the distances to the galaxies. Even with robust weak lensing measurements in hand, extracting cosmological information from these measured signals is non-trivial due to astrophysical effects like the intrinsic alignment of galaxies and baryonic effects.

This thesis focuses on this very topic, with a specific emphasis on the role of weak gravitational lensing in deriving cosmological parameters and investigating dark matter halos. Our primary objective is to enhance the accuracy of weak lensing analyses by improving the calibration of signal measurements and scrutinising the inference pipeline through consistency and sensitivity tests. Additionally, we aim to deepen our understanding of the relationship between dark matter halos and galaxies. This relationship is critical not only for achieving robust cosmological inference but also for enhancing our understanding of galaxy formation and evolution.

1.1 Foundations of the modern cosmology

Before delving into the cosmological applications of weak lensing, it is instructive to review the foundational elements of contemporary cosmological models. For the sake of brevity and specificity, we focus on the current standard cosmological model, namely the ΛCDM model. We begin with an examination of the isotropic and homogeneous average Universe in Sect. 1.1.1, followed by an analysis of the evolution of inhomogeneities at small scales in Sect. 1.1.2. Finally, we conclude with a discussion of key observational tests to the standard cosmological model in Sect. 1.1.3.

1.1.1 The homogeneous and isotropic Universe

The standard cosmological model is built upon two main foundations: the cosmological principle and the validity of general relativity at cosmological scales. The cosmological principle states that, on sufficiently large scales, the average matter distribution is both homogeneous and isotropic. This assumption leads to the Friedmann-Lemaître-Robertson-Walker (FLRW) metric, which defines the geometry of space-time as follows

$$\mathrm{d}s^2 \equiv -g_{\mu\nu}\mathrm{d}x^\mu x^\nu = c^2\mathrm{d}t^2 - a^2(t)\mathrm{d}l^2 \,, \tag{1.1}$$

where c stands for the speed of light. The metric tensor, $g_{\mu\nu}$, establishes the relationship between the coordinate values and the interval $\mathrm{d}s$ in the four-dimensional space-time. The scale factor, $a(t)$, captures the expansion of space. The comoving spatial element, $\mathrm{d}l$, is then defined as

$$\mathrm{d}l^2 \equiv \left[\frac{\mathrm{d}r^2}{1-Kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\ \mathrm{d}\phi^2)\right] \,, \tag{1.2}$$

where r, θ, and ϕ represent comoving spatial polar coordinates, and K is the curvature parameter. A K value of zero corresponds to a spatially flat Euclidean space, $K > 0$ to a positively curved, spherical space, and $K < 0$ to a negatively curved, hyperbolic space.

The relationship between the space-time geometry and the energy content in the Universe is described by the Einstein field equation:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} - \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \,. \tag{1.3}$$

Here, $R_{\mu\nu}$ and R denote the Ricci tensor and scalar, respectively, which describe the local curvature of space-time. The cosmological constant, Λ, was initially introduced by Einstein to achieve a static Universe, but it is now used to explain the observed accelerating expansion of the Universe. G is the Newtonian gravitational constant, and $T_{\mu\nu}$ represents the energy-momentum tensor.

In the case of a perfect fluid—an apt approximation for the average matter distribution at large scales under the cosmological principle—the energy-momentum tensor is formed by a simple diagonal matrix $T^\mu{}_\nu = \mathrm{diag}(\rho c^2, -P, -P, -P)$, where ρc^2 represents the

energy density, and P denotes pressure. By inserting this energy-momentum tensor and the FLRW metric into the Einstein field equation, we obtain the Friedmann equations, which describe the dynamics of the cosmological expansion:

$$\left[\frac{\dot{a}(t)}{a(t)}\right]^2 = \frac{8\pi G}{3}\rho(t) - \frac{c^2}{a^2(t)}K + \frac{c^2}{3}\Lambda\,, \tag{1.4}$$

and

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}\left(\rho + \frac{3}{c^2}P(t)\right) + \frac{c^2}{3}\Lambda\,. \tag{1.5}$$

The connection between the cosmological constant and the energy content is apparent from these equations. The influence of the cosmological constant is equivalent to an energy component with a constant density and pressure that satisfy $\rho_\Lambda = -P_\Lambda/c^2 = \Lambda c^2/(8\pi G)$. As such, the cosmological constant is also regarded as the simplest form of the mysterious dark energy, which is considered to be driving the observed accelerating expansion of the Universe.

Broadly speaking, the matter components in the Universe can be categorised into two groups: non-relativistic matter, which includes cold dark matter and ordinary baryons, and relativistic matter, consisting of photons and relativistic neutrinos. These two types of matter show different pressure-density relationships. For a perfect fluid, this relationship is characterised by the equation of state parameter w, defined as $w \equiv P/(\rho c^2)$. For non-relativistic matter, the pressure is negligible, corresponding to a $w_\mathrm{m} = 0$. On the other hand, relativistic matter has a $w_\mathrm{rad} = 1/3$. In similar terms, the cosmological constant can be interpreted as a dark energy component with a $w_\Lambda = -1$.

These equations of state enable us to re-formulate the Friedmann equations into the following form:

$$H^2(t) = H_0^2\left[\Omega_{\Lambda,0} + \Omega_{\mathrm{m},0}\,a^{-3}(t) + \Omega_{\mathrm{rad},0}\,a^{-4}(t)\right] - \frac{c^2}{a^2(t)}K\,, \tag{1.6}$$

where the Hubble parameter, $H(t) \equiv \dot{a}(t)/a(t)$, is introduced. Ω_i is defined as the ratio of the density ρ_i to the critical density $\rho_\mathrm{crit}(t) \equiv 3H^2(t)/(8\pi G)$. The subscript 0 indicates values measured at the current epoch ($a(t_0) = 1$).

The Hubble parameter at the current epoch, H_0, is also known as the Hubble constant, and it measures the current rate of cosmological expansion. Historically, the exact value of the Hubble constant has been uncertain, with estimates ranging from $50 - 100$ km s^{-1} Mpc^{-1}. Consequently, it is common to introduce a little h factor, defined as $H_0 = 100h$ km s^{-1} Mpc^{-1}, to absorb the uncertainties in the exact value of H_0. This convention will be used throughout this chapter.

The cosmological parameters in Eq. (1.6) were precisely determined through the *Planck* CMB observations, although some discrepancies persist among results from different cosmological probes (as discussed in Sect. 1.1.3). By combining information from temperature and polarisation maps, as well as lensing reconstruction, and adopting a spatially flat ΛCDM model ($K = 0$), Planck Collaboration et al. (2020) constrained

$H_0 = 67.36 \pm 0.54$ km s^{-1} Mpc^{-1}, $\Omega_{\Lambda,0} = 0.6847 \pm 0.0073$, $\Omega_{\mathrm{m},0} = 0.3153 \pm 0.0073$, and $\Omega_{\mathrm{rad},0} < 10^{-4}$.

1.1.2 Structure evolution

An entirely homogeneous and isotropic Universe would not spontaneously form the complex cosmological structures we observe today. The existence of these large-scale structures, such as groups and clusters of galaxies, implies the presence of initial density fluctuations. These early cosmological inhomogeneities, serving as the seeds of our current observed structures, grew due to gravitational instability. While the origin of these primordial density fluctuations remains a subject of debate — with the current popular theories suggesting quantum fluctuations associated with inflation or topological defects from early-time Universe phase transitions — the evolution of these fluctuations is well-described by the current standard cosmological model. In fact, when these primordial cosmological perturbations are considered as initial conditions, the standard ΛCDM model accurately predicts both the observed anisotropy in the CMB and the matter distribution in the late-time Universe. We briefly review some key aspects of structure formation within the framework of the standard ΛCDM model in this section.

Density fluctuations and two-point statistical measures

Typically, the cosmological density fluctuations are described by a perturbation field, δ, defined as the contrast between the local density and the mean background density as follows:

$$\delta(\boldsymbol{x}) \equiv \frac{\rho(\boldsymbol{x})}{\langle\rho\rangle} - 1 \,. \tag{1.7}$$

Here, $\langle\cdots\rangle$ denotes the averaging operator, and $\boldsymbol{x}$ represents the comoving spatial coordinate. It is also convenient to consider the perturbation field as a superposition of many signal modes, which in a flat comoving geometry, is quantified by the Fourier transform:

$$\delta(\boldsymbol{x}) = \sum_{\boldsymbol{k}} \delta_{\boldsymbol{k}} \exp(i\boldsymbol{k}\cdot\boldsymbol{x}) \,; \qquad \delta_{\boldsymbol{k}} = \frac{1}{V}\int \mathrm{d}^3\boldsymbol{x}\, \exp(-i\boldsymbol{k}\cdot\boldsymbol{x})\, \delta(\boldsymbol{x}) \,, \tag{1.8}$$

where V denotes the volume of a sufficiently large box, inside which the perturbations are assumed to be periodic, and $\boldsymbol{k}$ represents the wavevector. The cosmological principle implies that while the perturbation field characterises inhomogeneities, its statistical properties should exhibit homogeneity and isotropy like the background density field (see Chapter 16 of Peacock 1999 for a detailed discussion). This enables us to simplify calculations by reducing the three-dimensional coordinates of $\boldsymbol{x}$ and the wavevector $\boldsymbol{k}$ to single dimensions, x and k respectively.

Given the stochastic nature of the initial fluctuations, the meaningful predictions about the cosmological structure concern its statistical properties. Thus, it is necessary to introduce some statistical measures of the perturbation field. One of the basic measures

is the two-point correlation function, defined as

$$\xi(r) \equiv \langle \delta(x)\delta(x+r) \rangle \,, \tag{1.9}$$

where r denotes the separation between two spatial points in comoving units. This correlation function quantifies the clumpiness of the density field, and thus can serve as a means to characterise the cosmological structure. An equivalent measure in the Fourier space is the power spectrum, which is the Fourier transform of the two-point correlation function:

$$P(k) \equiv V\langle |\delta_k|^2 \rangle = 4\pi \int \mathrm{d}r\, \xi(r) \frac{\sin(kr)}{kr} r^2 \,, \tag{1.10}$$

In this equation, we integrated the angular coordinates in the 3D Fourier transform (as defined in Eq. 1.8), based on the assumption that the perturbation field is homogeneous and isotropic.

Assuming that the initial fluctuations follow a Gaussian random field, a condition which aligns with current observational data and the predictions of the popular inflation theory, the power spectrum or two-point correlation function entirely captures the statistical properties of the field. However, for non-Gaussian random fields, which could develop at later times due to non-linear dynamics, higher-order correlation functions become necessary. For the sake of brevity, we shall not extend our discussion to these higher-order statistics in this short introduction.

Linear evolution in the early stage

During the early stage of structure evolution, fluctuations are typically small ($\delta_k \ll 1$). In these circumstances, we can employ a perturbative approach to analytically solve the evolution problem, subject to certain approximations. The choice between Newtonian mechanics or a relativistic treatment is influenced by the properties of the matter content and the size of the perturbation. The simplest case involves sub-horizon perturbations—fluctuations on scales much smaller than the horizon size—in pressureless matter, which comprises cold dark matter and non-relativistic baryons.

In this scenario, the Newtonian perturbation theory is applicable. Moreover, if we disregard the entropy perturbation, the evolution of each Fourier mode of these perturbations can be described by the equation:

$$\ddot{\delta}_k + 2H(t)\dot{\delta}_k = 4\pi G\langle \rho \rangle \delta_k \,. \tag{1.11}$$

In this equation, a damping term $2H(t)\dot{\delta}_k$ arises due to the cosmological expansion, resulting in the growth of the fluctuations following a power law in time, rather than an exponential growth. This slow growth rate enables probing the primordial fluctuations using observations from the present epoch, as the evolved matter distribution still retains significant information about the initial state.

Non-linear evolution and gravitational collapse

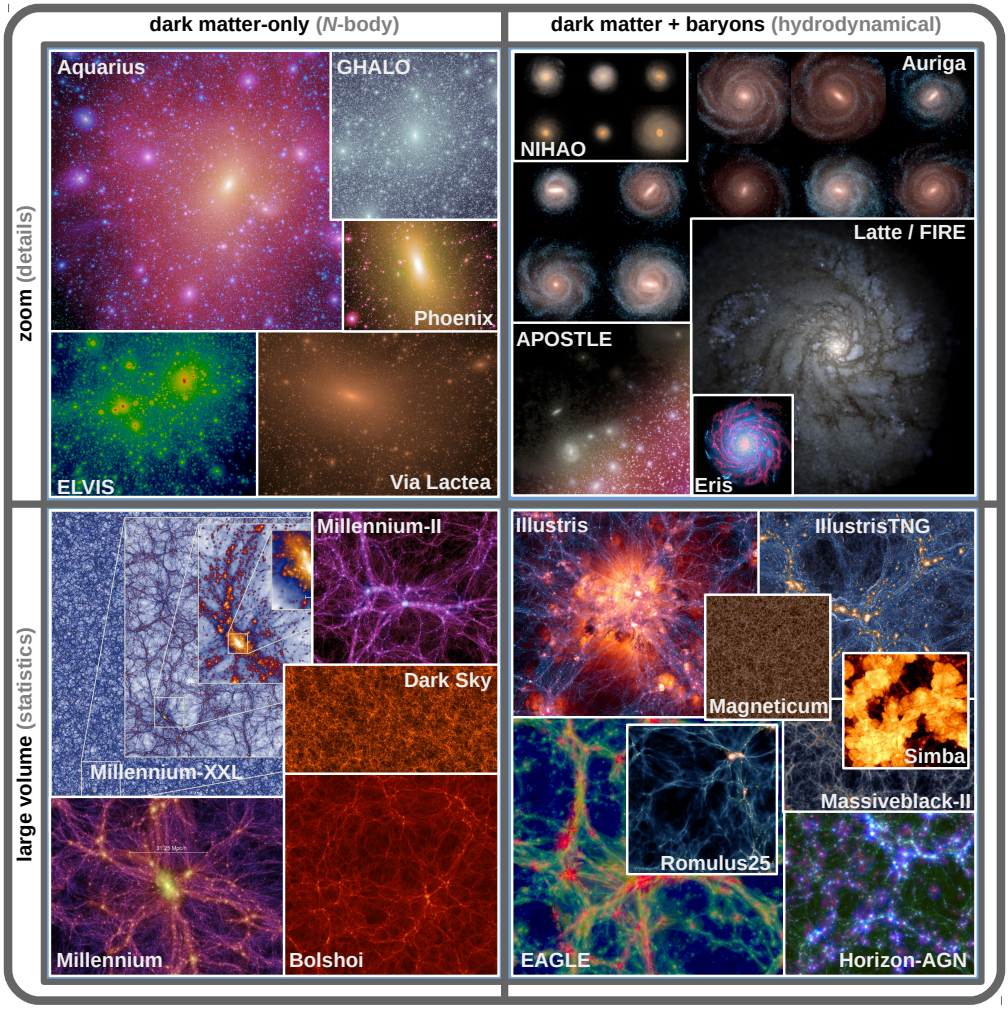

Figure 1.2: Visual representations of some numerical simulations. The simulations are categorised into large-volume simulations, which statistically represent large-scale structures, and zoom-in simulations, focusing on resolving smaller scales. Each of these types is further classified into two classes: dark matter-only N-body simulations, which consider only gravitational effects, and hydrodynamical simulations, which also include non-gravitational processes such as gas cooling and feedback mechanisms. [Credit: Vogelsberger et al. (2020)]

As the fluctuations continue to grow, the linear approximation becomes less and less accurate, and eventually a full non-linear treatment of gravitational collapse is required to effectively study galaxy formation. While the underlying physical principles remain the same, finding analytical solutions becomes increasingly challenging or even impossible

due to the complexity of these non-linear dynamical systems. This is where numerical simulations prove indispensable. Figure 1.2 showcases visual representations from some recent simulations. Such cosmological simulations are essential for understanding the intricate process of structure and galaxy formation, and they serve as key tools in interpreting observations from modern cosmological surveys.

While non-linear dynamics are complex, valuable insights can still be drawn from simplified analytical models. Comparing non-linear results with their linear counterparts can provide a qualitative understanding of the evolution of cosmic structures. A useful example involves the evolution of spherically symmetric, sub-horizon fluctuations in pressureless matter. For these small-scale fluctuations, the Newtonian equation $\ddot{r} = -GM/r^2$ is applicable, where r is the radius of a given mass shell, and M is the mass enclosed within that shell. Assuming that M remains constant over time and that the mass shell initially expands but ultimately collapses due to a negative net energy, we can solve the Newtonian equation to find the times of maximum expansion ($\dot{r} = 0$) and collapse ($r = 0$). Substituting these times into the formula derived from the linear treatment gives corresponding linear perturbation values of approximately 1.06 and 1.69, respectively. While these values do not match the actual amplitude of the fluctuations, they allow us to use linear perturbation theory to predict the timing of the turn-around and collapse.

The final density of the collapsed objects cannot be directly estimated from the aforementioned simplified model because the Newtonian equation diverges when $r \to 0$. This divergence arises because the assumption of a constant M breaks down when r is small and particles move across different mass shells. Ultimately, the kinetic energy from the random motion of particles balances the gravitational potential energy, and the collapsed system reaches virial equilibrium. Consequently, the virial theorem can be used to estimate the final overdensity of the collapsed system. Under the assumption of a matter-dominated Universe, the overdensity of the collapsed system in the spherical collapse model is found to be $\Delta_c = 18\pi^2 \approx 178$. Although this value is dependent on the idealised assumptions of the spherical collapse model, it implies that virialised regions can be effectively described by a sphere with an average density roughly 200 times the mean density of the Universe. This definition of virialised objects continues to be widely used in contemporary studies.

Smoothed perturbation filed and Press-Schechter formalism

A notable application of the aforementioned simple spherical collapse model is the Press-Schechter formalism proposed by Press & Schechter (1974). This method predicts the mass function of collapsed objects from the initial perturbation field, bypassing the need for detailed non-linear dynamic calculations. The basic idea is to evolve the perturbation field using linear theory and connect it to the mass function of collapsed objects. This connection is established by using a smoothed version of the perturbation

field, defined as

$$\Delta(\boldsymbol{x}; R) \equiv \int \mathrm{d}^3\boldsymbol{x}'\ \delta(\boldsymbol{x}')W(\boldsymbol{x}+\boldsymbol{x}'; R)\ ; \qquad \int \mathrm{d}^3\boldsymbol{x}\ W(\boldsymbol{x}; R)\ = 1, \tag{1.12}$$

where $W(\boldsymbol{x}; R)$ is a spherical window function with a characteristic radius R. One of the most commonly used window functions is the top-hat window function defined as

$$W(\boldsymbol{x}; R) = \begin{cases} \left(\frac{4\pi}{3}R^3\right)^{-1} & \text{if } |\boldsymbol{x}| \leq R \\ 0 & \text{otherwise} \end{cases}. \tag{1.13}$$

With this window function, it is straightforward to establish a connection between the characteristic radius and the average mass contained within the window volume: $M = 4\pi R^3\langle\rho\rangle/3$. A similar correlation between the characteristic radius and mass can be established for any chosen window function.

In practice, a top-hat window function with a radius of $8h^{-1}$ Mpc is often used because the variance of the cosmic structure at present day, when smoothed with this function, approximates one. This window function, when applied to smooth the present linear perturbation field—computed by evolving the initial perturbation field to the present day using the linear treatment—yields a root-mean-square measure known as σ_8, which is commonly used to quantify the strength of cosmic density fluctuations.

The Press-Schechter formalism assumes that the probability $\Delta(R) > \delta_\mathrm{c}$, denoted as $\mathcal{P}(>\delta_\mathrm{c})$, is equivalent to the fraction of collapsed objects with a mass greater than M, represented as $F(>M)$. Here, δ_c is the overdensity of collapsed objects inferred from the linear treatment. In the spherical collapse model discussed earlier, $\delta_\mathrm{c} \approx 1.69$. However, this assumption contains an inherent flaw due to the limits of linear treatment, which suggest only the regions initially overdense can result in collapsed objects. Consequently, this leads to a loss of half the total mass when transitioning from $\mathcal{P}(>\delta_\mathrm{c})$ to $F(>M)$.

To remedy this issue, Press & Schechter (1974) assumed without proof that the remaining mass in the underdense regions will be accreted into the collapsed objects, resulting in doubling all collapsed object masses but keeping the shape of the mass function unchanged, thus, $F(>M) = 2\mathcal{P}(>\delta_\mathrm{c})$. In the case of $\Delta(\boldsymbol{x}; R)$ following a Gaussian random field, the resulting Press-Schechter mass function is as follows:

$$n(M)\mathrm{d}M = \sqrt{\frac{2}{\pi}}\frac{\langle\rho\rangle}{M^2}\frac{\delta_\mathrm{c}}{\sigma_M}\exp\left(-\frac{\delta_\mathrm{c}^2}{2\sigma_M^2}\right)\left|\frac{\mathrm{d}\ln\sigma_M}{\mathrm{d}\ln M}\right|\mathrm{d}M\ , \tag{1.14}$$

where σ_M represents the dispersion of the smoothed overdensity field, as filtered by a window function corresponding to mass M.

Despite its reliance on several idealised assumptions and some unjustified ansatz, the Press-Schechter mass function effectively captures the general behaviour of the true mass function. Specifically, it correctly predicts the exponential decrease at the high mass end and the power-law increase at the low mass end. Further refinements of the

Press-Schechter formalism, including the extended Press-Schechter formalism by Bond et al. (1991) and the merger-tree method (e.g. Lacey & Cole 1993), continue to play significant roles in current hierarchical models of galaxy formation and serve as the backbone of modern semi-analytical models of galaxy formation.

1.1.3 Testing the standard cosmological model

Apart from theoretical developments, which primarily focus on parametrising the Universe with cosmological models and understanding structure formation, another crucial aspect of cosmological studies is the use of astronomical observations to test and refine these models. These observational endeavours play a significant role in falsifying established cosmological theories. With the advent of increasingly powerful cosmological surveys and the new observational window opened by the detection of gravitational waves, we can anticipate significant advancements in the coming decades. In this section, we provide a brief overview of some key cosmological probes that have played — and will continue to play — a central role in shaping our current standard cosmological model.

Determining the Hubble constant

The start of observational cosmology can be traced back to the pioneering work of Hubble (1929), who observationally confirmed the expansion of the Universe by measuring the recession of galaxies, as had been theoretically predicted by Lemaître (1927). The slope of the relationship between the distance to a galaxy and its recession velocity is now termed the Hubble constant H_0, which quantifies the current expansion rate of the Universe.

Accurately determining the value of H_0 remains an active area of research in observational cosmology. Figure 1.3 provides a summary of the current constraints on H_0 derived from various indirect and direct methods. The indirect method typically involves using CMB measurements to infer H_0 values based on assumed cosmological models. The direct method currently involves three approaches: the 'standard candle' method, which is based on bright objects with absolute luminosity that can be inferred from non-distance-related measurements; the 'standard siren' method based on gravitational waves, which enables a direct measurement of H_0 without forming any 'cosmic distance ladder'; and time-delay measurements from strongly lensed quasars.

Beyond the impressive precision of the current constraints from these varied methods, the most striking finding is the 4σ to 6σ discrepancy between the results from the indirect method and those from the direct method. This statistically significant H_0 tension triggered extensive interest in exploring potential systematic effects inherent in either method or in refining the current cosmological model (see Di Valentino et al. 2021 for a recent review).

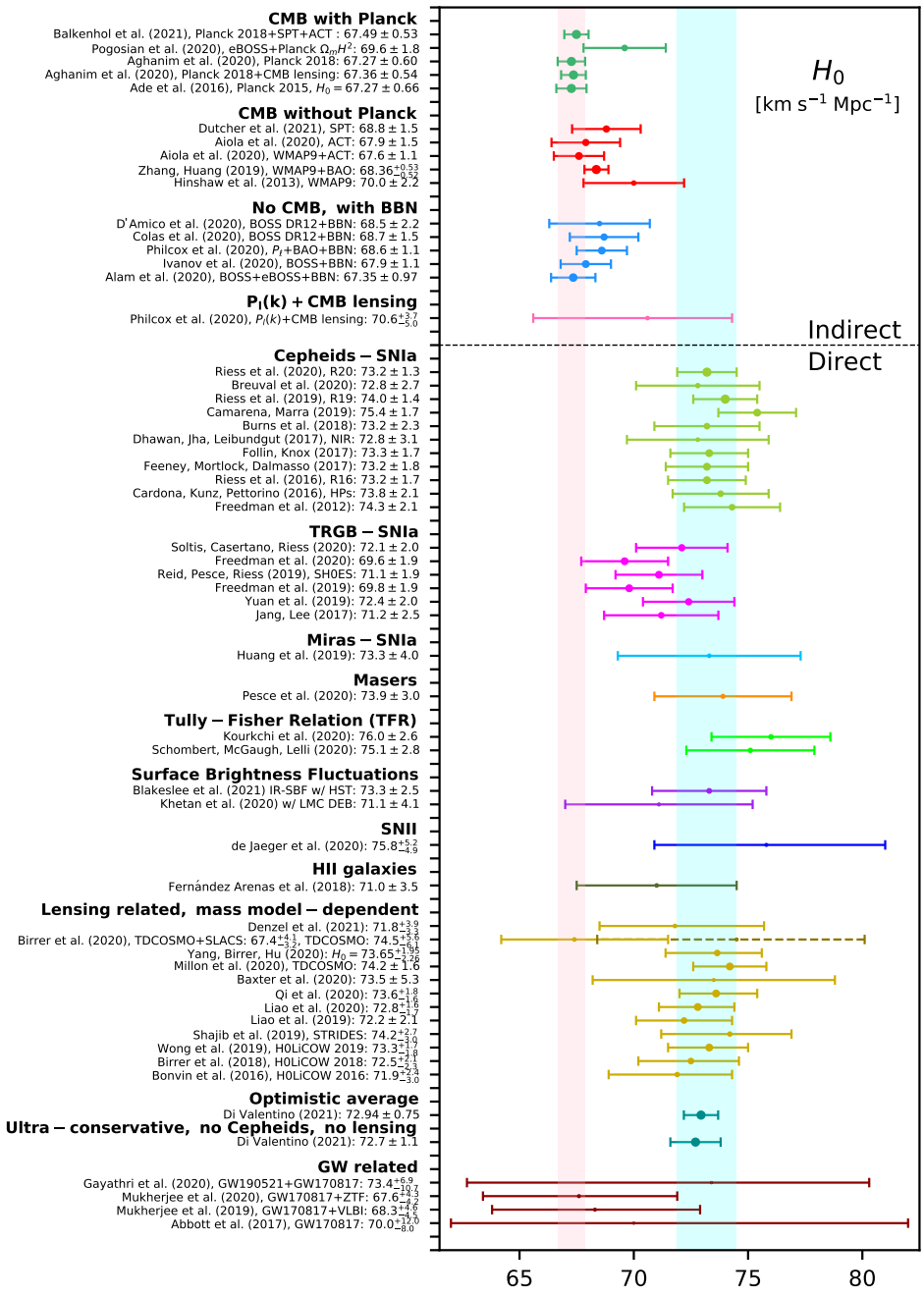

Figure 1.3: A summary plot showing the constraints on H_0 derived from various indirect and direct cosmological probes. The error bars denote the 68% credible intervals. The y-axis labels show the sources of the measurements, which are grouped based on their respective methods. [Credit: Di Valentino et al. (2021)]

Cosmic microwave background

The discovery of the CMB by Penzias & Wilson (1965) served as a crucial piece of evidence in support of the 'Big Bang' theory. The properties of the CMB observed today are consistent with predictions from this theory. According to the 'Big Bang' theory, the early Universe was extremely hot and dense, resulting in ionisation of all atoms and frequent photon scattering. This intense scattering resulted in a blackbody spectrum. As the Universe expanded, its temperature decreased, eventually reaching a point where protons and electrons could combine to form neutral atoms. This 'recombination epoch' rendered the Universe transparent, as neutral atoms could no longer scatter photons via Thomson scattering. This allowed photons to travel freely through space. The photons emitted from this 'last scattering surface' are what we now observe as the CMB, providing invaluable insights into the early stages of the Universe.

The pattern of CMB anisotropies is influenced by all cosmological parameters, making it an extremely powerful tool for constraining these parameters. Currently, the most precise constraints on the majority of cosmological parameters are provided by the *Planck* CMB measurements (Planck Collaboration et al. 2020). The Atacama Cosmology Telescope (Fowler et al. 2007) and South Pole Telescope (Carlstrom et al. 2011) experiments continue these efforts, with a focus on achieving higher angular resolution in CMB measurements and increasing sensitivity in detecting CMB polarisations. Such advances will not only further refine the constraints on cosmological parameters, but also aid in studying distant galaxy clusters and their environments by leveraging the interactions between CMB photons and the hot ionised gas within these clusters, a phenomenon known as the Sunyaev-Zeldovich effect (Sunyaev & Zeldovich 1972).

Measuring large-scale matter distribution

Another major class of observational methods involves the direct measurement of matter distribution in the late-time Universe, using techniques such as galaxy clustering and weak gravitational lensing, with the latter being the primary focus of this thesis. As detailed in Sect. 1.1.2, the evolution of cosmic structure within an expanding Universe encodes valuable information about initial density fluctuations. According to the standard ΛCDM model, the current matter distribution is mainly determined by the initial perturbation field and the horizon scale at the time of matter-radiation equality ($\rho_\mathrm{m} = \rho_\mathrm{rad}$). This horizon scale, as shown by Bardeen et al. (1986) and Sugiyama (1995), is characterised by the parameter $\Omega_\mathrm{m}h$ and to a lesser extent, the baryonic mass fraction. Consequently, observations of the large-scale structure offer direct constraints on the cosmological parameters Ω_m and σ_8. A standard approach to simplify the comparison of cosmological results from different probes is the introduction of a structure growth parameter, $S_8 = \sigma_8\sqrt{\Omega_\mathrm{m}/0.3}$.

Galaxy clustering operates on the assumption that galaxy locations trace the cosmic matter distribution. However, galaxies, being the outcome of complex, non-linear processes, can deviate from this assumption, resulting in a statistical difference between

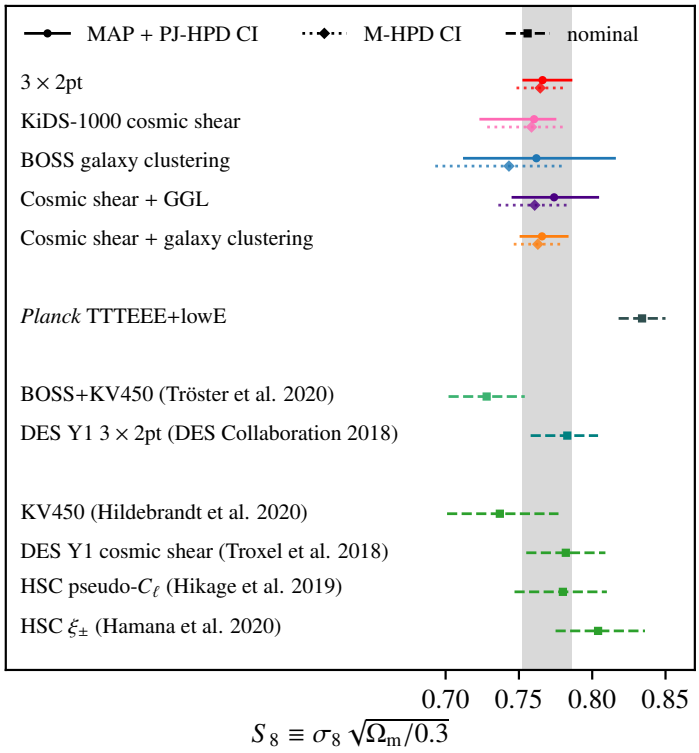

Figure 1.4: Constraints on S_8 from different cosmological probes. The label 3×2pt refers to results from a joint analysis combining three sets of two-point statistics: galaxy position-position correlations (galaxy clustering), galaxy shear-position correlations (galaxy-galaxy lensing), and galaxy shear-shear correlations (cosmic shear). Other labels represent the respective survey teams and the probes used. The error bars correspond to the 68% credible intervals. [Credit: Heymans et al. (2021)]

the galaxy and underlying matter distributions (see Desjacques et al. 2018 for a review). This galaxy bias complicates the interpretation of measured galaxy clustering signals. Furthermore, constructing a 3D galaxy distribution depends on galaxy redshift estimates, which encapsulate both the cosmological expansion and the line-of-sight peculiar velocity, thus introducing contamination from intricate dynamical effects.

On the other hand, the weak lensing effect, which directly correlates with the gravitational field of foreground matter, provides an unbiased trace of the large-scale structure. However, the minute distortions induced by the weak lensing effect are challenging to measure and susceptible to various systematic errors associated with the measurement process. Despite advances in high-quality imaging surveys and innovative shear measurement algorithms, careful control of systematic effects remains essential as survey statistical power increases.

A promising approach entails a joint analysis of these two methods, as they

probe the same underlying large-scale structure of the Universe. This method can break the degeneracy between cosmological parameters found in individual observables, significantly increasing the precision of cosmological parameter constraints. Furthermore, because different observables are affected by different systematic effects, a joint analysis permits some degree of self-calibration, leading to more accurate cosmological results.

Figure 1.4 presents the S_8 constraints from a recent joint analysis by the Kilo-Degree Survey (KiDS, Heymans et al. 2021), comparing results with the *Planck* CMB measurements and other contemporary galaxy surveys such as the Baryon Oscillation Spectroscopic Survey (BOSS, Alam et al. 2015), the Dark Energy Survey (DES, Dark Energy Survey Collaboration et al. 2016), and the Hyper Suprime-Cam (HSC) survey (Aihara et al. 2018). It shows that the precision of the current joint analysis of galaxy clustering and weak lensing is already comparable to the *Planck* CMB measurements. Interestingly, a mild 2σ to 3σ discrepancy exists between the S_8 values measured from large-scale matter distribution and those inferred from CMB measurements. Explaining this mild S_8 tension is more subtle than resolving the more significant H_0 tension. Given the high precision of modern surveys, any slight systematic residuals in any probe could induce such level of difference, underscoring the importance of identifying and controlling measurement and astrophysical systematic errors. More detailed discussions on this aspect will be presented in the following sections and throughout this thesis.

1.2 Weak gravitational lensing

The deflection of light by massive objects due to the curvature of space-time induced by their gravity is a common phenomenon in modern astronomical observations and is well-grounded in the theory of general relativity. Cosmological lensing effects can be broadly divided into two categories depending on the prominence of the lensing effect, as illustrated in Fig. 1.5: strong lensing, which produces substantial arcs or multiple images; and weak lensing, which causes gravitational distortions to the source galaxy images that are not visually detectable and thus require a statistical analysis.

This section provides a concise overview of gravitational lensing theory, emphasising weak lensing effects. We begin with a brief introduction to the fundamental concepts of lensing theory in Sect. 1.2.1, which is followed by discussions on the two primary weak lensing configurations: galaxy-galaxy lensing and cosmic shear in Sect. 1.2.2 and Sect. 1.2.3, respectively. The section concludes with a summary of major challenges related to the measurement and modelling of weak lensing for cosmological applications, presented in Sect. 1.2.4.

1.2.1 Fundamentals of gravitational lensing

For the study of gravitational lensing, we typically deal with lens objects that are much smaller than the Hubble length cH_0^{-1} and have peculiar velocities much smaller than the speed of light. In such cases, the lensing effects can be represented by introducing small

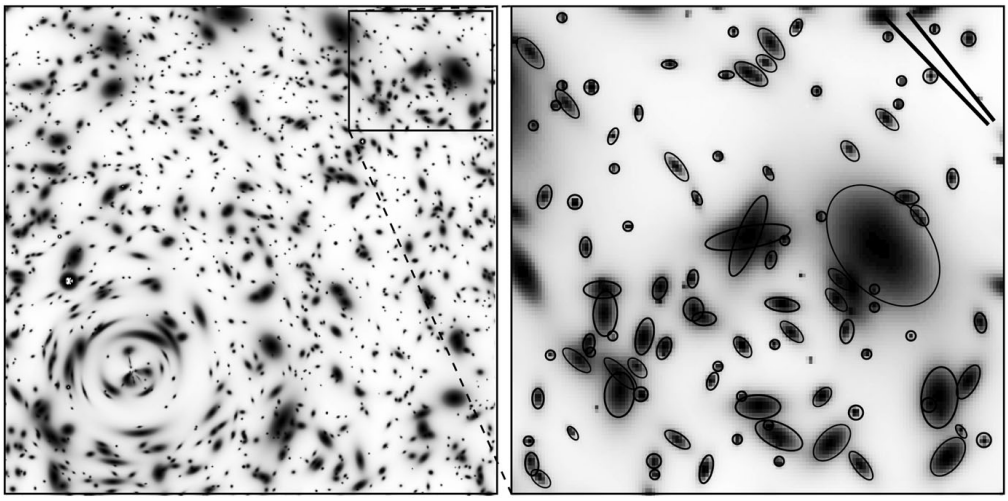

Figure 1.5: Simulated images illustrating the two lensing regimes. The bottom left part of the left panel plot concerns the strong lensing regime, where giant arcs and multiple images are produced. As the radial distance of source galaxies from the lens centre increases, the lensing effect decreases, and eventually lensing-induced distortions become smaller than the intrinsic ellipticity of the source galaxies. The right panel offers a zoomed-in view of this weak lensing regime, where the contours represent galaxy shapes as determined from their second moments. [Credit: Mellier (1999)]

perturbations to the FLRW metric (Eq. 1.1) as

$$\mathrm{d}s^2 = \left(1 + \frac{2\Phi}{c^2}\right) c^2 \mathrm{d}t^2 - \left(1 - \frac{2\Phi}{c^2}\right) a^2(t) \mathrm{d}l^2 \,, \tag{1.15}$$

where Φ is the Newtonian potential, which is assumed to be weak ($\Phi \ll c^2$) and satisfies the Poisson equation:

$$\nabla^2 \Phi = 4\pi G \rho \,. \tag{1.16}$$

Here, ρ represents the mass density of the lens objects, and the gradient operator, ∇, operates on the physical, or proper, coordinates.

This perturbed FLRW metric, in combination with the light geodesic equation ($\mathrm{d}s = 0$), forms the theoretical foundation for solving gravitational lensing problems that meet the weak gravitational field requirements. However, deriving a general analytical solution from these equations is challenging and often requires approximations. Therefore, a more practical approach would be to find specific solutions for individual problems based on their unique conditions. We shall adopt this approach, starting with a simple case known as the geometrically-thin lens to introduce some general concepts associated with gravitational lensing theory. More practical examples concerning the cosmological applications of weak gravitational lensing will be discussed in subsequent sections.

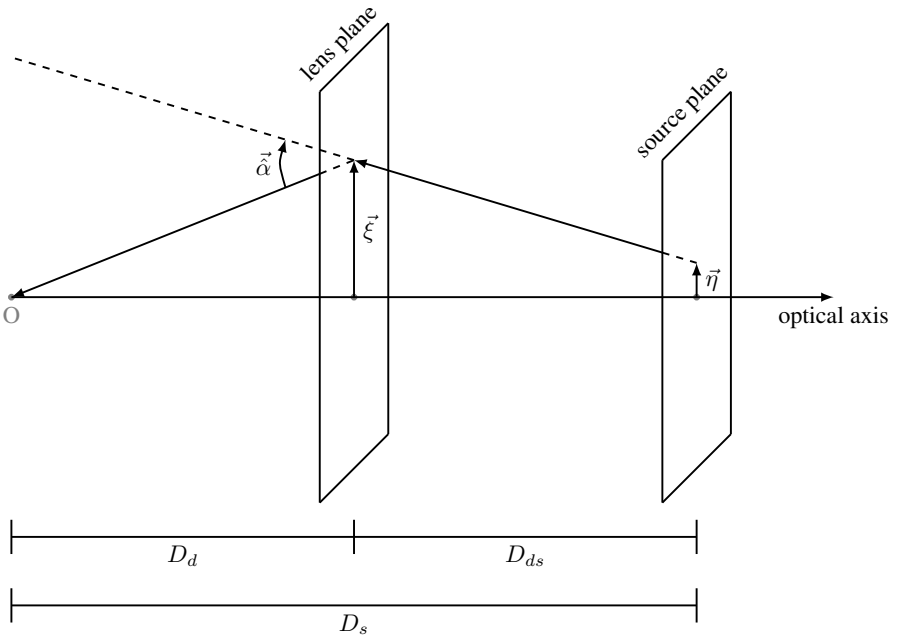

Figure 1.6: A sketch illustrating the geometry of a thin gravitational lensing system.

The deflection angle and lens equation

A geometrically-thin lens corresponds to a lens whose size is much smaller than the distances between the observer, lens, and source. This is a valid assumption for most astrophysical objects, such as galaxies and clusters of galaxies. Under such circumstances, we can use a geometric relationship to describe the effects of gravitational lensing on light propagation, as illustrated in Fig. 1.6. Here, the lens is located at an angular diameter distance $D_{\rm d}$, while the source object is located at an angular diameter distance $D_{\rm s}$. The angular diameter distance between the lens and source is denoted as $D_{\rm ds}$.

Locally, we can construct source and lens planes that are perpendicular to a straight line (referred to as the optical axis in the figure) extending from the observer to the lens, and onwards to infinity. These planes intersect the optical axis at the centres of the source and lens, respectively. Then, the smoothly curved trajectory of light, which is the reality under gravitational lensing, can be approximated by two straight lines that intersect at a point in the lens plane, forming a kink. The magnitude and direction of this kink are described by a deflection angle, $\vec{\hat{\alpha}}$, which quantifies the extent of light deflection due to the gravitational lensing effect.

The simplest lens model under this configuration involves a point mass M serving as the lens. As long as the impact parameter—defined as the distance between the kink and the lens—is much larger than the Schwarzschild radius of the lens ($|\vec{\xi}| \gg R_{\rm s} \equiv 2GMc^{-2}$), the weak gravitational field approximation remains valid, and the deflection angle can

be found as (see e.g. Narayan & Bartelmann 1997):

$$\vec{\hat{\alpha}}_{\rm p}(\vec{\xi}) = \frac{4GM}{c^2\vec{\xi}}. \tag{1.17}$$

An extended lens object can be treated as a collection of point masses. The resulting deflection angle is then calculated as the vector sum of the deflections due to each individual point mass. Each point mass element has a mass $\mathrm{d}M = \rho \mathrm{d}V$, where ρ is the mass density of the lens and $\mathrm{d}V$ is a volume element. If we denote the three-dimensional location of the kink as (ξ_1, ξ_2, r_3) and that of the mass element $\mathrm{d}M$ as (ξ_1', ξ_2', r_3'), then the impact vector is $\vec{\xi} - \vec{\xi}'$, independent of r_3 due to our assumption of a geometrically-thin lens. The total deflection angle can be obtained as (see e.g. Bartelmann & Schneider 2001):

$$\vec{\hat{\alpha}}(\vec{\xi}) = \frac{4G}{c^2}\int \mathrm{d}^2\xi'\ \Sigma(\vec{\xi}')\ \frac{\vec{\xi} - \vec{\xi}'}{|\vec{\xi} - \vec{\xi}'|^2}\ ; \qquad \Sigma(\vec{\xi}) \equiv \int \mathrm{d}r_3\ \rho(\xi_1, \xi_2, r_3)\ , \tag{1.18}$$

where $\Sigma(\vec{\xi})$ is the surface mass density.

Once we determined the deflection angle, mapping from the source plane to the lens plane becomes straightforward using the geometric relations illustrated in Fig. 1.6. Let $\vec{\eta}$ represent the two-dimensional source position in the source plane, we have

$$\vec{\eta} = \frac{D_{\rm s}}{D_{\rm d}}\vec{\xi} - D_{\rm ds}\vec{\hat{\alpha}}(\vec{\xi})\ . \tag{1.19}$$

We can simplify this equation by introducing angular coordinates, defined as $\vec{\beta} \equiv \vec{\eta}/D_{\rm s}$ and $\vec{\theta} \equiv \vec{\xi}/D_{\rm d}$, and a scaled deflection angle

$$\vec{\alpha}(\vec{\theta}) \equiv \frac{D_{\rm ds}}{D_{\rm s}}\vec{\hat{\alpha}}(D_{\rm d}\vec{\theta}) = \frac{1}{\pi}\int \mathrm{d}^2\theta'\ \kappa(\vec{\theta}')\frac{\vec{\theta} - \vec{\theta}'}{|\vec{\theta} - \vec{\theta}'|^2}\ , \tag{1.20}$$

where $\kappa(\vec{\theta}) \equiv \Sigma(D_{\rm d}\vec{\theta})/\Sigma_{\rm cr}$ is the dimensionless surface mass density, often called convergence, with $\Sigma_{\rm cr} \equiv c^2 D_{\rm s}/(4\pi G D_{\rm d} D_{\rm ds})$ representing the critical surface mass density. With these definitions, Eq. (1.19) simplifies to

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta})\ . \tag{1.21}$$

This relation, known as the lens equation, provides the mapping from a sky position in the source plane to a position in the lens plane. In physical terms, this means that a source situated at a position $\vec{\beta}$ in the sky will appear at position $\vec{\theta}$ due to the gravitational lensing effect. Notably, this mapping from the source plane to the lens plane may not be one-to-one. Indeed, Eq. (1.21) can yield more than one solution for certain $\vec{\beta}$. In physical terms, these multiple solutions correspond to multiple images of the same source object being detected by the observer, a phenomenon commonly referred to as

strong lensing.

The Fermat potential and time delay

Equation (1.20) implies the existence of a deflection potential, defined as

$$\psi(\vec{\theta}) \equiv \frac{1}{\pi} \int \mathrm{d}^2\theta' \, \kappa(\vec{\theta}') \ln|\vec{\theta} - \vec{\theta}'| \,, \tag{1.22}$$

which satisfies $\alpha = \nabla_\theta \psi$. Here, ∇_θ represents the two-dimensional gradient operator in respect to θ, acting on the lens plane. This deflection potential can be interpreted as a two-dimensional analogue of the Newtonian potential, given that it also satisfies the two-dimensional Poisson equation:

$$\nabla^2_\theta \psi = 2\kappa = \frac{2\Sigma}{\Sigma_{\mathrm{cr}}} \,. \tag{1.23}$$

Building on this, we can define a Fermat potential

$$\tau(\vec{\theta}; \vec{\beta}) \equiv \frac{1}{2}(\vec{\theta} - \vec{\beta})^2 - \psi(\vec{\theta}) \,, \tag{1.24}$$

and the condition that its gradient $\nabla_\theta \tau$ equals to zero is equivalent to the lens equation given in Eq. (1.21).

As demonstrated by Schneider (1985), the Fermat potential $\tau(\vec{\theta}; \vec{\beta})$, subject to an affine transformation, represents the light travel time from the source to the observer, accounting for the lensing effect. Thus, the lens equation can be interpreted as another manifestation of Fermat's principle, which states that light follows a path where the travel time is stationary, i.e., a local minimum, maximum or saddle point. Given the generality of the Fermat potential and its connection to the light travel time, it proves valuable in understanding the properties of lens systems, including the study of multiple images and time delays in strong lensing (e.g., Blandford & Narayan 1986).

The Jacobian matrix and image distortion

Until now, we have focused on the mapping of a single source position. In reality, source objects like galaxies have an extended shape, and light emitted from different parts of these sources will be deflected differently. This results in distortions in the observed images. In general, quantifying this image distortion involves solving the lens equation for each point within an extended source. However, when a source is relatively small and the lens properties do not change significantly within its angular scale, the image distortion can be linearly represented by the Jacobian matrix:

$$\mathcal{A}(\vec{\theta}) = \frac{\partial \beta}{\partial \theta} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} \,, \tag{1.25}$$

where

$$\gamma_1 = \frac{1}{2}\left(\frac{\partial^2\psi}{\partial\theta_1^2} - \frac{\partial^2\psi}{\partial\theta_2^2}\right); \qquad \gamma_2 = \frac{\partial^2\psi}{\partial\theta_1\partial\theta_2}\,, \tag{1.26}$$

represent the two components of the shear parameter $\gamma \equiv \gamma_1 + \mathrm{i}\gamma_2$. The convergence κ, introduced earlier, is connected to the deflection potential through Eq. (1.23).

According to Liouville's theorem, surface brightness is conserved before and after gravitational lensing, assuming no photon emission or absorption occurs during the process. As a result, the Jacobian matrix contains all information about changes in the local shape and flux of the source pre- and post-lensing distortion. To better grasp how the parameters in the Jacobian matrix relate to the distortions of the source image, we can re-formulate the Jacobian matrix as

$$\mathcal{A}(\vec{\theta}) = (1-\kappa)\begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix}\,, \tag{1.27}$$

by introducing the reduced shear parameter: $g = \gamma/(1-\kappa)$.

With Eq. (1.27), it is evident that the factor $(1-\kappa)$ influences the image isotropically, affecting only the size and apparent flux of the observed image without altering its shape. In the context of weak lensing, since the intrinsic properties of source objects are unknown, this isotropic magnification effect cannot be directly measured for individual lens systems[1]. On the other hand, the anisotropic distortions are entirely determined by the reduced shear g, making it the central quantity in the study of weak gravitational lensing.

1.2.2 Galaxy-galaxy lensing

After establishing the general principles of gravitational lensing, we can now explore some practical examples of weak lensing that are vital in observational cosmology. Our first example involves lens systems where individual massive objects, such as galaxies, or groups/clusters of galaxies, serve as the lens. Such lens systems are commonly referred to as 'galaxy-galaxy lensing' in contemporary weak lensing studies. Considering that a typical galaxy cluster is only a few Mpc in size, while the cosmological distances involved are usually several hundred to a few thousand Mpc, the geometrically-thin lens condition is well satisfied. Thus, the results derived in the previous section can be directly implemented in galaxy-galaxy lensing studies.

The key element in galaxy-galaxy lensing studies involves the correlation between the surface mass density and the lensing effect, as illustrated in Eq. (1.18). This correlation enables a range of applications such as estimating the total lens mass, mapping the matter distribution, and more. Although the convergence κ is directly linked to the surface mass density by its definition, it is challenging to measure in practice. Conversely, the

[1]However, we can statistically measure the magnification effect, for example, by evaluating changes in the local source number density (e.g. Broadhurst et al. 1995; Schneider et al. 2000; Hildebrandt et al. 2009) or in the size-magnitude-redshift relation of galaxies (e.g. Alsing et al. 2015).

shear parameter γ is relatively easier to measure from observed source images. Hence, establishing a connection between convergence and shear is crucial.

Assuming the surface mass distribution of the lens is axisymmetric, this connection can be readily derived from their relations to the deflection potentials (Eqs. 1.23 and 1.26):

$$\gamma_{\rm t}(\theta) = \overline{\kappa(<\theta)} - \kappa(\theta) \ , \tag{1.28}$$

where $\overline{\kappa(<\theta)}$ represents the average value of κ within a radius of θ. We also introduced the tangential shear: $\gamma_{\rm t} \equiv -\gamma_1 \cos(2\varphi) - \gamma_2 \sin(2\varphi)$, where φ represents the relative position angle of the source with respect to the lens centre. Remarkably, this relationship holds for general matter distributions as long as we replace local $\gamma_{\rm t}(\theta)$ and $\kappa(\theta)$ estimations with their azimuthally-averaged counterparts $\langle\gamma_{\rm t}(\theta)\rangle$ and $\langle\kappa(\theta)\rangle$ (Bartelmann 1995).

In practical terms, this implies that we can determine the azimuthally-averaged mass profile of a lens object by measuring the azimuthally-averaged tangential shear, irrespective of whether the density distribution is axisymmetric. This method can be used to map the projected matter distribution of lens objects or to estimate the total mass of the lens, provided the density profile of the lens is available. This application of galaxy-galaxy lensing with real-world data will be the topic of Chapter 5.

1.2.3 Cosmic shear

Another primary application of the weak lensing effect is its use in directly probing the large-scale structure of the late-time Universe. This method, known as 'cosmic shear', considers all matter distributions between the source and the observer as the lens. As mentioned in Sect. 1.1.3, the study of large-scale structures provides powerful constraints on cosmological models, and cosmic shear is one of the few methods capable of such examination. Therefore, cosmic shear plays a central role in modern cosmological surveys.

On the theoretical side, 'the lens' of cosmic shear, spanning the entire space between the source and the observer, breaks the geometrically-thin lens condition and requires an extension of the classical lens theory. Furthermore, given the lack of a concrete lens in this context, cosmic shear measurements depend on statistical analyses to establish a connection with the statistical properties of the cosmic density field. Covering all these aspects with careful derivations would exceed the scope of this introductory section. Therefore, we will only outline the key elements of cosmic shear without detailed derivation, focusing on the second-order statistical measures, namely, the two-point correlation function and power spectrum as defined in Sect. 1.1.2. For a more comprehensive discussion, we refer to Bartelmann & Schneider (2001) and Kilbinger (2015).

Although the geometrically-thin lens condition no longer holds, the assumption of weak gravitational perturbations remains applicable. Thus, the perturbed FLRW metric and the Poisson equation, as defined in Eqs. (1.15) and (1.16), are still valid. Considering $2\Phi/c^2$ as a small parameter, we can identify an effective deflection potential expressed

as

$$\psi_{\rm eff}(\vec{\theta}, w) = \frac{2}{c^2}\int_0^w \mathrm{d}w' \frac{f_K(w-w')}{f_K(w)f_K(w')}\Phi[f_K(w')\vec{\theta},\ w']. \tag{1.29}$$

Here, w represents the comoving distance, $f_K(w)$ is the comoving angular diameter distance, which equals w in a spatially flat Universe, and the integral is performed from the observer to a target source.

Using this definition of $\psi_{\rm eff}$, we can apply Eq. (1.23) to introduce an effective convergence parameter, formulated as

$$\kappa_{\rm eff}(\vec{\theta}, w) = \frac{3H_0^2\Omega_{\rm m,0}}{2c^2}\int_0^w \mathrm{d}w'\ \frac{f_K(w')f_K(w-w')}{f_K(w)}\frac{\delta[f_K(w')\vec{\theta},\ w']}{a(w')}\ , \tag{1.30}$$

where we used the Poisson equation (1.16) and replaced ρ with the overdensity δ, as defined in Eq. (1.7). This equation reveals the correlation between the cosmic shear effect and cosmological parameters. The effective convergence parameter is essentially a line-of-sight integral of the cosmic perturbation field, with the scale factor $a(w')$ and a combination of comoving angular diameter distance factors acting as a weighting function. The combination of the Hubble constant and matter density parameter only influences the amplitude of $\kappa_{\rm eff}(\vec{\theta}, w)$.

In practice, cosmic shear measurements from a collection of source galaxies are averaged to enhance the signal-to-noise ratio. If these sources have a redshift distribution such that $p_z(z)dz = G(w)dw$, we can derive an averaged effective convergence, expressed as

$$\overline{\kappa}_{\rm eff} = \frac{3H_0^2\Omega_{\rm m,0}}{2c^2}\int_0^{w_{\rm H}} \mathrm{d}w'\ \overline{W}(w')f_K(w')\frac{\delta[f_K(w')\vec{\theta},\ w']}{a(w')}\ . \tag{1.31}$$

Here, the weighting function, also known as the lensing efficiency factor, is defined as

$$\overline{W}(w') \equiv \int_{w'}^{w_{\rm H}} \mathrm{d}w\ G(w)\frac{f_K(w-w')}{f_K(w)}\ . \tag{1.32}$$

In these equations, the upper limit of integration corresponds to the horizon distance, equivalent to infinite redshift.

To connect the cosmic shear measurements to the statistical properties of the cosmic overdensity field, we use the power spectrum and apply the Limber's equation, which links the two-point statistical measures of a projected field to its original three-dimensional field (Limber 1953). Following this approach, we can derive

$$P_\kappa(k) = \frac{9H_0^4\Omega_{\rm m,0}^2}{4c^4}\int_0^{w_{\rm H}} \mathrm{d}w\ \frac{\overline{W}^2(w)}{a^2(w)}P_\delta\left(\frac{k}{f_K(w)}, w\right)\ , \tag{1.33}$$

which links the power spectrum of convergence field to the power spectrum of the cosmic matter distribution.

However, as mentioned earlier, convergence is challenging to measure in weak

lensing studies, necessitating a connection to the shear parameter. Fortunately, in the weak lensing limit, the two-point statistical properties of convergence and shear are identical: $P_\gamma = P_\kappa$ (Blandford et al. 1991). This relationship can be straightforwardly demonstrated in Fourier space by considering the relations of convergence and shear to the deflection potential, as defined in Eqs. (1.23) and (1.26). Hence, we can use the observed shear statistics to directly study the cosmic matter distribution, adhering to the same relationship as in Eq. (1.33).

The line-of-sight integral in Eq. (1.31) implies that cosmic shear primarily measures two-dimensional projections of the three-dimensional cosmic matter distributions. To mitigate the signal loss from this integration, contemporary cosmic shear analyses often include an additional layer of correlations along the line of sight by binning galaxies based on their redshifts, a method known as tomographic analysis. This technique has proven highly effective in enhancing the statistical power of cosmic shear for constraining cosmological parameters (e.g. Hu 1999). Furthermore, it enables cosmic shear to constrain properties of dark energy by facilitating measurements of the evolution of cosmic structure (e.g. Huterer 2002). Chapter 4 presents an application of this cosmic shear tomography using real-world data.

1.2.4 Challenges in weak lensing analysis

By now, it should be evident that the theory of weak lensing is well-established and has shown itself to be extremely powerful in studying the matter distribution of both individual objects and large-scale cosmic structures. However, applying it to real-world data presents substantial challenges that warrant further investigation. This is particularly true given the advancements of current and upcoming cosmological surveys, and the intriguing disagreements found among different cosmological probes, as discussed in Sect. 1.1.3. This section provides a brief overview of some key systematic issues encountered during the measurement and modelling of weak lensing. Chapters 2 and 3 will delve deeper into these systematics, exploring their impact and discussing strategies for their control in real-world applications.

Shear measurement and selection biases

The initial challenge in weak lensing studies arises when attempting to measure subtle weak lensing distortions from noisy, faint galaxy images. This complexity is amplified by distortions caused by the point spread function (PSF), resulting from instrumental and observational conditions (e.g. Paulin-Henriksson et al. 2008; Massey et al. 2013), the detection or selection bias during the identification of faint objects (e.g. Hartlap et al. 2011; Chang et al. 2013; Hoekstra et al. 2021), and the blending effects that occur when two or more objects are in close proximity on the sky (e.g. Hoekstra et al. 2015; Dawson et al. 2016; Hoekstra et al. 2017; Mandelbaum et al. 2018; Samuroff et al. 2018; Euclid Collaboration et al. 2019). These factors not only introduce shear measurement bias but also modify the selection function of the source sample, complicating the interpretation of the measured shear signals. Therefore, managing shear measurement and selection

biases is a crucial first step in any weak lensing studies.

Following two decades of development, the state-of-the-art shear measurement methods are capable of controlling measurement biases at a sub-percent level (e.g. Heymans et al. 2006; Massey et al. 2007; Bridle et al. 2010; Kitching et al. 2012; Mandelbaum et al. 2015). In addition, the use of realistic pixel-level image simulations helps address residual biases and account for selection biases (see Chapter 3). This progress paints a promising picture for controlling shear biases in future high-requirement weak lensing surveys. Further investigation into blending effects and the interplay between shear and redshift measurements could still provide benefits (e.g. MacCrann et al. 2022; Li et al. 2023b). However, the shear measurement alone does not appear to pose the major challenge for future surveys. More demanding are tasks such as careful modelling of the PSF and improving the realism of image simulations (e.g. Li et al. 2023a,b). These improvements, related to the shear measurement and calibration, seem to be the more challenging aspects and warrant further study.

Photometric redshift uncertainties

Determining the distances of source galaxies presents a significant challenge for current and future weak lensing surveys. This distance estimation relies on observing the electromagnetic spectral energy distribution (SED) of a galaxy. The cosmological expansion stretches the observed SED towards longer wavelengths by a factor $(1 + z)$, where z is the redshift. Therefore, the redshift can be related to a proper distance, assuming a cosmological model. To measure the stretch of the SED, we need to identify some characteristic features in it, such as emission and absorption lines. These features are easy to identify if the observed wavelength resolution is high, i.e., if we obtain the spectra of galaxies.

However, for deep imaging surveys, the fraction of galaxies with observed spectra is very low. Fortunately, weak lensing analyses do not require precise redshift estimates for each source galaxy. The source redshifts enter the calculation through an integral, as shown in Eq. (1.32). Therefore, as long as the average redshift of the ensemble of source galaxies is accurately estimated, the analysis yields unbiased results. This allows for the use of so-called photometric redshifts, which are less precise.

This method measures the flux of distant galaxies in a few broader wavelength filters, thereby obtaining a sparse sampling of the SED. It estimates the redshift based on broad features such as the Lyman and Balmer breaks. Nevertheless, the raw estimates from this process have large uncertainties that exceed the requirements of current and future weak lensing analyses (e.g. Hildebrandt et al. 2010; Salvato et al. 2019). Therefore, an additional step of redshift calibration, using more precise spectroscopic reference samples, is necessary. However, this calibration process introduces its own biases, for instance caused by incompleteness or non-representative sampling of the samples with spectroscopic redshifts. This results in another layer of residual bias correction that also relies on realistic simulations.

The first-step calibration is currently undergoing rapid development with the rise

of machine learning-based techniques, such as self-organising maps (e.g. Kohonen 1982; Masters et al. 2015). The focus now is on enhancing the spectroscopic reference samples and gaining a better understanding of the transfer function in both samples, which describes how intrinsic photometry maps to measured ones and its correlation with redshifts. This progress is crucial as colour-based redshift calibration methods rely on the assumption that galaxy colour and redshift are closely correlated, and that the spectroscopic calibration sample and the faint photometric target sample share similar underlying properties (e.g. Hildebrandt et al. 2021; Myles et al. 2021; Rau et al. 2022). Simulations can aid this understanding, but the current absence of simulations for the spectroscopic reference sample, and the complexity of the transfer function in spectroscopic measurements, make this a challenging endeavour. Therefore, further explorations in this direction are critical.

For the second step, residual bias correction, realism in simulations is pivotal, much like in shear calibration (e.g. Hoyle et al. 2018; van den Busch et al. 2020; DeRose et al. 2022). Recent advancements in multi-band image simulations appear promising (e.g. Li et al. 2023b). These simulations naturally account for the transfer function in the photometric sample and the blending effect, an aspect previously overlooked in catalogue-level simulations. Furthermore, multi-band image simulations enable a joint calibration of both shear and redshift estimates, leading to a full end-to-end calibration, that is beneficial for both processes. The development of such multi-band image simulations will be the focus of Chapter 3.

Baryonic effects and Intrinsic alignment

Even after achieving robust shear and redshift measurements, interpreting the weak lensing signal, particularly the cosmic shear signal, remains a complex task. This complexity arises from the fact that cosmic shear probes the matter distribution in the late-time Universe, which undergoes non-linear evolutions, as discussed in Sect. 1.1.2. Additionally, measured statistical signals include contamination from the intrinsic alignment (IA) of galaxies, a non-lensing-induced coherent alignment of galaxies, originating from tidal effects between neighbouring galaxies. Therefore, a crucial aspect of weak lensing studies involves enhancing modelling accuracy to account for these non-linear and astrophysical effects.

Regarding non-linear evolution, cosmological numerical simulations offer a promising solution. Advances in simulation algorithms and computational power now make it possible to create gravitational-only N-body simulations that meet the requirements for interpreting future weak lensing measurements (e.g. Heitmann et al. 2009; Nishimichi et al. 2019; Euclid Collaboration et al. 2021). However, baryonic processes, such as feedback from star formation and active galactic nuclei, also alter matter distributions on scales relevant to cosmic shear studies (e.g. van Daalen et al. 2011; Semboloni et al. 2011). These non-gravitational processes cannot be captured by N-body simulations, necessitating hydrodynamic simulations that include baryonic processes. However, constructing hydrodynamic simulations that capture all astrophysical processes is compu-

tationally demanding. More cost-effective alternatives, such as simplified hydrodynamic simulations with fixed subgrid physics or semi-analytical models, require a sufficient understanding of the astrophysical processes that are not included in the original simulations, and this remains an active research area (e.g. Schaye et al. 2023; Kugel et al. 2023).

The issue of IA contamination is notably challenging to resolve due to our currently limited understanding of its properties (see e.g. Joachimi et al. 2015; Troxel & Ishak 2015 for reviews). The prevailing approach introduces empirical models with nuisance parameters to mitigate the IA impact (e.g. Hirata & Seljak 2004; Bridle & King 2007; Blazek et al. 2019). However, these IA parameters were found to correlate with other nuisance parameters, such as those related to redshift calibration (see Chapter 2). This issue is further complicated by the fact that different types of galaxies exhibit varying IA behaviours, which complicates the interpretation of a mixed-colour weak lensing sample (e.g. Johnston et al. 2019; Fortuna et al. 2021b; Samuroff et al. 2022). In this regard, a strategy of splitting the source samples based on their colours would be beneficial, a method we employ in our Chapter 2 study. Moreover, advancements in cosmological simulations allow for the study of IA within these simulations, providing valuable insights into the underlying mechanisms of IA signals and aiding in refining our modelling approach (e.g. Tenneti et al. 2014; Chisari et al. 2015; Hilbert et al. 2017; Kurita et al. 2021).

1.3 This thesis

The rest of this thesis delves into the real-world applications of weak lensing as a cosmological probe. We focus on two main aspects of weak lensing applications: cosmic shear and galaxy-galaxy lensing, as previously introduced. Our goal is to enhance constraints on cosmological parameters and deepen our understanding of the connection between dark matter halos and galaxies. Throughout this work, we devote considerable attention to identifying, testing, and controlling systematic effects that arise during weak lensing measurements and modelling, recognising this aspect as critical for current and future weak lensing studies.

In **Chapter 2**, we conduct an internal consistency test for cosmic shear analysis using data from the Kilo-Degree Survey (KiDS). We split source galaxies based on their colours and assess the robustness of the current KiDS pipeline when handling samples with varied galaxy properties. Our findings confirm that the existing KiDS pipeline effectively accounts for various systematic residuals. However, we identify a correlation between IA parameters and nuisance parameters associated with redshift calibrations, emphasising the need for improved IA models. Notably, the consistency test method we use is insensitive to assumed cosmological models, making it an effective tool for verifying the robustness of cosmic shear analysis before drawing any cosmological conclusions.

Chapter 3 introduces a suite of multi-band image simulations for shear and redshift calibration of KiDS weak lensing analysis. These newly developed simulations enable a

joint calibration of shear and redshift, enhancing the realism and accuracy over previous efforts. By integrating cosmological simulations with high-quality imaging data, we generate a large volume of simulated galaxies, accurately mirroring realistic galaxy properties to an adequate depth. We improve the realism of image simulations by carefully addressing various observational and instrumental effects, including noise background, point spread function, and stellar density. Furthermore, we study blended systems at different redshifts by introducing variable shear fields into the simulations. Our study reveals a correlation between shear and redshift biases due to blending effects, underscoring the importance of joint shear and redshift calibration. We also identify a minor but noteworthy impact of PSF modelling errors on shear bias. After conducting sensitivity tests, we confirm the robustness of the current KiDS fiducial shape measurement algorithm, within the requirements of the KiDS weak lensing analysis. For future weak lensing surveys with stricter requirements, we recommend further studies on blending effects, refinement of PSF modelling, and improvements to shape measurement techniques to reduce their sensitivity to variations in galaxy properties.

In **Chapter 4**, we conduct a cosmic shear analysis, incorporating improvements in cosmic shear measurements and calibrations facilitated by the work conducted in Chapter 3. Moreover, we take into account recent advancements in cosmological inference and investigate the influence of prior information on IA modelling on the final cosmological results. We propose treating the statistical and systematic uncertainties from the shear calibration separately, given their distinct origins. Our final cosmological results align with previous KiDS studies and other weak lensing surveys, showing a tension of approximately 2.3σ with the *Planck* cosmic microwave background constraints on S_8.

Finally, in **Chapter 5**, we investigate the connection between dark matter haloes and galaxies using galaxy-galaxy lensing. We measure galaxy-galaxy lensing signals around galaxy groups identified by the Galaxy And Mass Assembly (GAMA) project using KiDS data. We interpret these statistical lensing signals through the halo model formalism to infer the total halo mass associated with these galaxy groups. By adopting optical observables provided by GAMA, we constrain the scaling relation between the halo mass and the total luminosity of the groups. We assess the robustness of our halo mass estimates by varying the treatment of halo model parameters not well-constrained by the current data. These sensitivity tests help us identify critical model components and guide future improvements. Given the increased statistical power of the measured signals, this exploration and future enhancement of our current model are considered necessary.