

ALMA MATER STUDIORUM
UNIVERSITÀ DEGLI STUDI DI BOLOGNA

DIPARTIMENTO DI FISICA E ASTRONOMIA
Dottorato di Ricerca in Astronomia
CICLO XXVII

**SIMULATIONS FOR GRAVITATIONAL LENSING OF 21 CM
RADIATION AT EOR AND POST-EOR REDSHIFTS**

Dottorando
Alessandro Romeo

Supervisore
Prof. Lauro Moscardini

Co-Supervisori
Dr. Robert Benton Metcalf
Dr. Alkistis Pourtsidou

Coordinatore
Prof. Lauro Moscardini

Esame finale anno 2015

Settore Concorsuale: 02/C1 – Astronomia, Astrofisica, Fisica della Terra e dei Pianeti
Settore Scientifico-Disciplinare: FIS/05 – Astronomia e Astrofisica

Abstract

21 cm cosmology opens an observational window to previously unexplored cosmological epochs such as the Epoch of Reionization (EoR), the Cosmic Dawn and the Dark Ages using powerful radio interferometers such as the planned Square Kilometer Array (SKA). Among all the other applications which can potentially improve the understanding of standard cosmology, we study the promising opportunity given by measuring the weak gravitational lensing sourced by 21 cm radiation. We performed this study in two different cosmological epochs, at a typical EoR redshift and successively at a post-EoR redshift, when HI is mostly found within galaxies.

As already showed by early works, an SKA-like instrument could map the distribution of matter with high fidelity if EoR occurred at $z=8$. Moreover, if a large enough fraction of the sky could be observed, a high precision power spectrum could be obtained. The lensing signal is measurable because the 21 cm source can be divided up into multiple statistically independent maps while the lensing signal is the same for each. We will show how the lensing signal can be reconstructed using a three dimensional optimal quadratic lensing estimator in Fourier space, using single frequency band or combining multiple frequency band measurements.

To this purpose, we implemented a simulation pipeline capable of dealing with issues that can not be treated analytically, like the simulation of a telescope beam, the non-uniform visibility space coverage, the non-linearity of the lensing source field, the discreteness of visibility measurements, the non-Gaussianity of the 21 cm radiation source field, and the cross-correlation among frequency bands left by foreground cleaning techniques.

Considering the current SKA plans, in the first part of this work we studied the performance of the quadratic estimator at typical EoR redshifts, for different survey strategies and comparing two thermal noise models for the SKA-Low array: the first is widely used in the literature but assumes a uniform visibility space distribution, while the second takes into account a more realistic array density distribution. The simulation we performed takes into account the beam of the telescope (set by its maximum dimensions), and the discreteness of visibility measurements, preparing the way for future numerical studies aimed to investigate more realistic issues.

We found that an SKA-Low interferometer should obtain high-fidelity images of the underlying mass distribution in its phase 1 only if several bands are stacked up together, covering a redshift range that goes from $z=7$ to $z=11.5$ and with a total resolution of 1.6 arcmin. We also implemented a simple de-noising

procedure in order to filter out the small-scale noise which is likely to strongly contaminate the estimated signal. The SKA-Low phase 2, modeled in order to improve the sensitivity of the instrument by almost an order of magnitude, should be capable of providing images with good quality even when the signal is detected within a single frequency band. In this case the reconstructed image has a resolution of 1.15 arcmin at $z = 8$, within a field of view of 13 deg^2 .

Considering also the serious effect that foregrounds could have on this detections, we discussed the limits of these results and also the possibility provided by these models of measuring an accurate lensing power spectrum. In the case of multi-band detection of the lensed 21 cm signal made with an SKA2-Low telescope model we found constraints close to the sample variance ones in the range $L < 1000$, even for a small field of view such as a 25 deg^2 survey area.

In the second part of this work we adopted the discrete grid theoretical framework constructed at EoR redshifts in order to develop a weak lensing reconstruction formalism at post-EoR redshifts, namely $z \sim 2.5$. With Intensity Mapping, discrete point sources need to be resolved only in frequency and can be added incoherently to the clustered 21 cm signal. Following the study performed by Pourtsidou & Metcalf (2015), we included their contribution as an additive discrete Poisson noise to a clustering Gaussian three-dimensional signal, and demonstrated that they contribute to improve the lensing reconstruction signal-to-noise by computing a discrete non-optimal quadratic estimator, for clustered and unclustered sources.

Unclustered point sources have been simulated in order to study the behaviour of this estimator and to move towards more realistic clustered signal reconstruction simulations. The estimator for unclustered and clustered sources can be tested within our simulation code, and, as made for the EoR observations, we have modified the estimator by including the beam and considering a non-uniform antennae distribution in our simulation framework. In order to explore the post-EoR epoch, we have also modeled in our code a more accurate power spectrum for the 21 cm brightness temperature fluctuation field, with a formalism which will allow for studying different HI evolutionary models. We developed a thermal noise model for the lower frequency band SKA-Mid in interferometer mode, including the possibility to observe different parts of the sky. With SKA-Mid larger survey areas become indeed available and better constraints on power spectrum accuracy can be obtained.

Contents

Abstract	iii
1 Introduction	1
1.1 Observing the 21 cm Universe with SKA	3
1.1.1 SKA-Low Overview	4
1.1.2 SKA-Mid Overview	6
1.1.3 Comparison with Other Radio Telescopes	8
1.2 HI Intensity Mapping	10
1.2.1 Searching for Baryonic Acoustic Oscillations at Low Redshifts	11
1.3 The Need of Simulations	11
1.4 The Foregrounds Contamination Problem	12
1.4.1 Foregrounds Classification	13
1.4.2 To Be or Not to Be Blind?	14
1.4.3 Polarization Leakage	16
2 Cosmology and Weak Gravitational Lensing	19
2.1 Theoretical Cosmology Background	19
2.1.1 The Friedmann-Robertson-Walker-Lemaitre Model	20
2.1.2 The Expansion of the Universe	21
2.1.3 The Λ CDM Universe and its Evolutionary Phases	23
2.1.4 Distances in Cosmology	24
2.1.5 Large Scale Structure Formation	27
2.1.6 Poisson's Equation	32
2.2 Principles of Gravitational Lensing Theory	33
2.2.1 Lensing Convergence	35
2.2.2 Differential Effects	36
2.2.3 Second Order Effects	39
2.3 Lensing in General Relativity	39
2.3.1 Light Ray Bundle Propagation	40
2.3.2 Application to a FRW Metric	41
2.3.3 Weak Cosmological Lensing from LSS	43
2.4 Convergence Power Spectrum from Limber's Equation	50
2.4.1 Shear Power Spectrum	52

3	21 cm Radiation Cosmology	55
3.1	Physics of the 21 cm Line of Atomic Hydrogen	56
3.2	Determining the Spin Temperature	58
	3.2.1 Collisional Coupling	59
	3.2.2 The Wouthuysen-Field Effect	62
3.3	21 cm Radiation Global Evolution	64
	3.3.1 Evolution of the IGM	66
	3.3.2 Sources for Heating and Ionization	68
	3.3.3 Evolution of Ly α Flux	69
	3.3.4 Heating by Exotic Mechanisms	69
3.4	21 cm Brightness Fluctuation Power Spectrum	70
	3.4.1 Redshift Space Distortions	71
	3.4.2 The Alcock-Paczynski Effect	73
	3.4.3 Ionization Fluctuations	74
	3.4.4 Ly α Coupling Fluctuations	75
	3.4.5 Temperature and Ionization Fluctuations from X-rays	75
	3.4.6 Full Power Spectrum Evolution and other source of fluctuations	77
	3.4.7 Simulations: State of the Art	78
	3.4.8 Higher Order Statistics	79
3.5	Constraining Cosmology with 21 cm Radiation	79
4	21 cm Radiation Weak Lensing Reconstruction	83
4.1	Lensing Preliminaries	84
4.2	21 cm Brightness Temperature Fluctuation Field	85
4.3	Lensing Reconstruction in Fourier Space	88
	4.3.1 Faster Lensing Estimator	90
4.4	Including the Beam	91
4.5	Measuring 21 cm Lensing Power Spectrum	93
4.6	Cosmological Parameter Estimation	94
5	Simulations for 21 cm Lensing Imaging at EoR	97
5.1	Discrete 21 cm Field Modeling	97
5.2	Lensing Maps	100
5.3	Modeling the Beam	102
	5.3.1 Aliasing and the Beam	103
5.4	Uniform Thermal Noise Component	104
5.5	Results for the Uniform Thermal Noise Model	106
	5.5.1 Single-Band Reconstruction	107
	5.5.2 Testing Aliasing Contamination	113
	5.5.3 Other Tests and Multi-Realisation Reconstruction	116
	5.5.4 Performance for different SKA configurations	119
	5.5.5 Multi-Band Reconstruction	121
	5.5.6 Lensing Power Spectrum Measurement	128
	5.5.7 Cluster Detection	130
5.6	Adopting a More Realistic Thermal Noise Model	132
	5.6.1 Non-Uniform Thermal Noise Power Spectrum	132

5.6.2	SKA1-Low and SKA2-Low Specifications	134
5.6.3	Single-Band Results for non-Uniform Noise Model	134
5.6.4	De-noising of the Reconstructed Image	136
5.6.5	Multi-Band Results for non-Uniform Noise Model	143
5.6.6	Power Spectrum Measurement Estimate	146
5.7	Resuming of Results and Future Developments (pt. I)	148
6	Reconstruction at Low Redshifts	151
6.1	Weak Lensing Estimator for Unclustered Sources	152
6.1.1	Pure Poisson Noise Power Spectrum	152
6.1.2	Quadratic Estimator for Magnification	154
6.1.3	Including the Beam for Unclustered Sources	155
6.2	Simulating Unclustered Point Sources	155
6.3	Weak Lensing Estimator for Clustered Sources	156
6.3.1	Including the Beam for Clustered Sources	160
6.4	Telescope Model	161
6.4.1	Single-Dish Mode Thermal Noise	161
6.4.2	Interferometer Mode Thermal Noise	162
6.4.3	SKA1 and SKA2-Mid Specifications	162
6.5	The 21 cm Power Spectrum Model at Post-EoR Redshifts	163
6.5.1	HI Evolution	165
6.5.2	A Bias for HI Tracing	166
6.5.3	The Total Clustered Angular Power Spectrum	167
6.6	Resuming of Results and Future Developments (pt. II)	168
7	Conclusions	171
A	Gaussian Random Fields	175
A.1	Fourier Transforms Convention	175
A.1.1	Hermiticity Condition	175
A.1.2	Dirac Delta Function	175
A.1.3	Convolution Theorem	176
A.2	GRF Statistics	176
A.3	The Wiener-Khintchine Theorem	178
A.4	Higher-Order Statistics	179
A.5	Limber's Equation in Fourier Space	180
A.5.1	About Limber's Approximation Accuracy	182
B	Toy Models for Gravitational Lensing	183
B.1	Axially Symmetric Lenses	183
B.2	Point Mass	185
B.3	Singular Isothermal Sphere	186
B.3.1	Softened Isothermal Sphere	188
B.4	The Navarro-Frenk-White Lens Model	189

C	Discrete Fourier Transforms and FFTW Storing	191
C.1	Fourier Series	191
C.2	Sampling and Aliasing	192
C.2.1	The Nyquist Sampling Theorem	193
C.3	Discrete Fourier Transforms	193
C.4	Using FFTW	193
C.4.1	Two-dimensional arrays	195
C.4.2	Three-dimensional arrays	196
D	21 cm Weak Lensing Reconstruction	199
D.1	Line-of-Sight Discretisation in Fourier Space	199
D.2	Continuous 3D Quadratic Estimator	200
D.2.1	Comparison with 2D CMB Analogous and Comments	203
D.3	Discrete 21 cm Weak Lensing 3D Estimator	204
D.4	Discrete Estimator with a Beam Function	206
D.5	Fast Quadratic Estimator Derivation	207
E	Non-Uniform Thermal Noise Derivation	211
E.1	Noise Angular Power Spectrum	211
E.2	The Uniform Limit	214
E.2.1	Comparing Uniform and non-Uniform Cases	215
F	The Schechter Luminosity Function	217
F.1	Schechter Function Definition	217
F.2	Schechter Function Moments	218
G	High-Order Poissonian Statistics Computations	219
G.1	Statistics for Unclustered Sources	219
G.1.1	Estimator Variance for Unclustered Sources	220
G.2	Statistics of Clustered Sources	223
G.2.1	Power Spectrum	223
G.2.2	Bispectrum	224
G.2.3	Trispectrum	226
	Acknowledgements	231
	Bibliography	233

Chapter 1

Introduction

21 cm cosmology opens an observational window to previously unexplored cosmological epochs such as the Epoch of Reionization (EoR), the Cosmic Dawn and the Dark Ages (Furlanetto et al., 2006) using powerful radio interferometers such as the planned Square Kilometer Array (SKA)¹ (Pritchard et al., 2015). The same radiation from lower redshifts can be used to measure cosmological parameters through the Baryonic Acoustic Oscillations (BAO) (Chang et al., 2008; Ansari et al., 2012; Battye et al., 2013a; Smoot & Debono, 2014). In this work we aim to investigate numerically the possibility to use this radiation to measure weak gravitational lensing.

21 cm radiation is generated by the hyperfine, spin flip, transition of neutral hydrogen. When the CMB photons and neutral hydrogen spin temperature become thermally decoupled the radiation is potentially observable in absorption or emission depending on whether the spin temperature is lower or higher than CMB temperature. In principle, the 21 cm line gives us access to a huge volume of the currently unobserved Universe covering the redshift range $z \sim 6 - 200$ during which the neutral fraction is high, as well as more recent, post-reionization, epochs where neutral hydrogen (HI) is found only within galaxies. The redshifted 21 cm line allows us to obtain a 3D map of the Universe, across the sky and along cosmic time by observing in a range of frequencies.

Current and planned experiments like the SKA (Square Kilometer Array), LOFAR² (Low Frequency Array), 21CMA³, PAPER⁴ (Precision Array for Probing Epoch of Reionization), and MWA⁵ (Murchinson Widefield Array) have investigated the high-redshift Universe through HI as their primary or one of their primary goals. 21 cm EoR studies can be combined with other measurements coming from CMB, HI intensity mapping at lower redshift and/or galaxy surveys to study the dark sector of the Universe (Dark Matter and Dark Energy) and investigate whether exotic dark energy or modified gravity effects are present (Copeland et al., 2006; Clifton et al., 2012).

¹<http://www.skatelescope.org/>

²<http://www.lofar.org/>

³<http://21cma.bao.ac.cn/>

⁴<http://http://eor.berkeley.edu/>

⁵<http://www.haystack.mit.edu/ast/arrays/mwa/>

These observations will be a unique opportunity to study the EoR. In fact, the Wouthuysen-Field effect couples Ly- α photons to HI temperature, generating a 21 cm signal that only depends on hydrogen density and neutral fraction. 21 cm tomography is also able to provide the 3D 21 cm temperature brightness fluctuation power spectrum (McQuinn et al., 2006; Mao et al., 2008; Barkana & Loeb, 2005) and constrain the cosmological parameters. The technique of HI intensity mapping, which treats the 21 cm temperature field as a continuous, unresolved background and thus does not rely on detecting individual galaxies, can be used to measure the Baryonic Acoustic Oscillations (BAO) at low redshifts (Chang et al., 2008; Ansari et al., 2012; Battye et al., 2013a; Smoot & Debono, 2014), map the mass distribution in the pre-EoR era (Lu & Pen, 2008), or perform weak lensing studies in post-reionization redshifts (Pourtsidou & Metcalf, 2014).

The 21 cm radiation can be used as a source for gravitational lensing studies. Early works (Zahn & Zaldarriaga, 2006; Metcalf & White, 2009; Pourtsidou & Metcalf, 2015) showed that if the EoR is at redshift $z \sim 8$ or later, an SKA-like instrument could map the distribution of matter with high fidelity and if a large enough fraction of the sky could be observed a high precision power spectrum could be obtained. The weak lensing effect is measurable because the 21 cm source can be divided up into multiple statistically independent maps while the lensing signal is the same for each. This can be extracted from the data using a Fourier space quadratic lensing estimator, which was originally developed for the CMB case by (Hu & Okamoto, 2002) and then extended in 3D for the 21 cm case by (Zahn & Zaldarriaga, 2006). Here we make further modifications of the estimator to explicitly take into effect the beam of the telescope and gridding of the visibility measurements.

Such an estimator would allow for a high-fidelity imaging of the underlying mass distribution, mainly composed by Dark Matter, whose gravitational effects drive the linear growth and the subsequent non-linear collapse of the fluctuations measured in the Cosmic Microwave Background (CMB). Imaging the Dark Matter would be a big step forward in understanding its distribution and having an observational verification of its existence. This means that at low redshifts we could even see with high-fidelity the true matter distribution, allowing the dark matter haloes of individual galaxies to be directly observed. This would give a priceless statistical and morphological information about the relative distributions of mass and light (Hilbert et al., 2007).

Observing the 21 cm signal from the EoR is made challenging by several factors. At such low frequencies foreground contamination (mainly synchrotron emission) poses a particular problem. Foregrounds dominate over the cosmological signal by about four orders of magnitude, but studies indicate that they can be successfully removed by taking advantage of their relative coherence in frequency in comparison to the 21 cm signal from structure in the HI distribution (Liu & Tegmark, 2012; Chapman et al., 2012; Dillon et al., 2013).

In order to assess how well gravitational lensing could actually be measured in realistic observations it is crucial to perform numerical simulations. Previous assessments have been based on simplifying assumptions that make predicting the noise analytically tractable. In particular, the 21 cm emission has

been treated as a Gaussian random field (or a Gaussian clustered field with a discrete Poissonian contribution in the case of Pourtsidou & Metcalf (2015)) and it has been assumed that foreground subtraction is done perfectly with no residual effects that might affect the lensing results. These are both important factors that cannot be handled analytically. Incomplete and uneven u - v coverage is another issue that is best treated numerically. Here we introduce a numerical tool that can be used to do more realistic studies that will include these effects.

In addition to developing the simulation technique and code, the aim of this work is to study how well the lensing signal from 21 cm sources at typical EoR redshifts ($z \sim 8$) can be reconstructed using the quadratic estimator technique, so that we can investigate its effective performance given the current SKA plans.

Moreover we will provide the basic formalism to implement within this code the quadratic estimator reconstruction for post-EoR redshifts, such as $z \sim 2.5$, including the point source signal of single galaxies and modeling a clustered HI density in the Universe. This reproduces more realistically the observational environment measured by the intensity mapping technique which treats the detected 21 cm radiation as a continuous background, without resolving the individual point sources.

1.1 Observing the 21 cm Universe with SKA

The SKA will be the world's largest and most sensitive radio telescope and its findings are expected to push the boundaries of our understanding of astrophysics and cosmology. The potential explorable volume by the SKA is impressive. If foregrounds are taken under control, we could detect signal from an unprecedented redshift range, spanning a survey volume that goes from $z = 0$ to $z \sim 30$, as shown in Figure 1.1. Compare this range with that of the Sloan Digital Sky Survey SDSS, whose luminous red galaxies sample has a mean redshift of $z \approx 0.3$, or with the one provided by CMB, a shell around $z \sim 1100$. If a large volume is probed, we could constrain also neutrino mass and the running of the primordial power spectrum (Visbal et al., 2009).

The design of the SKA is not permanent yet, so there is still some freedom in assuming the instrumental specifications. The main document which describes the most complete specifications for the SKA is the official baseline design written by Dewdney (2013). More recently the SKA has been de-scoped, due to economical budget restrictions (McPherson, 2015). The original baselines have been considerably modified, leading to important modifications of the telescope models adopted throughout this work. This issue will be treated more deeply in Section 5.6 and Appendix E.

Excluding the preliminary phase SKA0, in which limited prototypes and pathfinders of the SKA will be used to provide the first results (like MeerKAT or ASKAP), the SKA will be built in at least two phases (SKA1 and SKA2) and will have arrays for low frequency (Low, 50 - 350 MHz) and medium frequency (Mid, 350-13800 MHz), which are built in Australia and South Africa, respec-

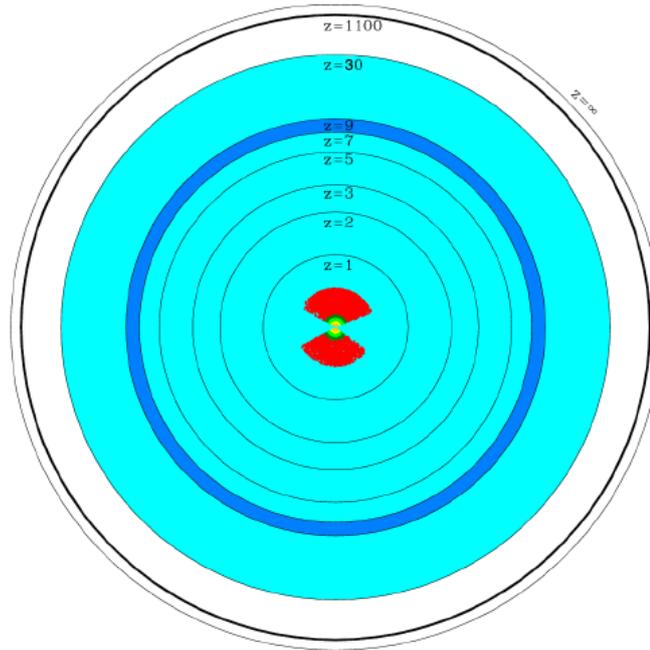


Figure 1.1: The 21 cm tomography could probe a vast part of our cosmological horizon (in teal) up to $z \sim 30$ and potentially it could explore further regions. To make a comparison we also show a tiny shell at $z \sim 1100$ probed by CMB, and the red region indicating the small volume probed by the Sloan Digital Sky Survey. The EoR region is the dark blue shell around $z \sim 8$. Taken from Mao et al. (2008).

tively. The design of Phase 2 is still quite uncertain, but the expectation is that it will be four and ten times more sensitive with respect to SKA1 phase, for Low and Mid arrays respectively, having at least four times the collecting area of the one used for phase 1. A further planned Sur (Survey) experiment was considered into the original design. This would have been operating in the range from 650 to 1670 MHz, providing surveys of large fractions of the sky, but it has been suppressed by the already mentioned re-baselining of the SKA.

It has been found that both SKA-Low and SKA-Mid intensity mapping should be able to detect weak lensing using the intensity mapping method. In this work we will mainly consider SKA-Low working in interferometer mode in Chapter 5, while SKA-Mid will be briefly addressed in Chapter 6 for post-EoR redshifts. As can be seen, our results will drastically depend on the assumptions made on telescope's design, and these will be exhaustively discussed in Sections 5.5 and 5.6 for SKA-Low. Here we will describe the main features of SKA phases, referring all detailed specifications like antennae distribution, total number of antennae, collecting area, etc., to the next chapters of this work, when we will need to model the telescope for our simulation sets.

1.1.1 SKA-Low Overview

The primary science objective of SKA-Low is to observe the reionization of the universe at high redshift through its signatures in the 21 cm radiation

(Pritchard et al., 2015). Another SKA-Low target is to detect the HI-line absorption against continuum sources over the redshift range $z \sim 6 - 20$.



Figure 1.2: Artist's impression of an SKA-Low antennae station. Image credits: Swinburne Astronomy Productions/ICRAR/U. Cambridge/ASTRON.

SKA-Low is a low-frequency aperture array (AA) to be used mainly in interferometer mode, *i.e.* using large baselines made of several 35 m diameter stations. These are a collection of hundreds of log-periodic dual-polarised dipole antennae arrays, like the ones shown in Figure 1.2. Log-periodic antennae are half-wavelength dipoles of different dimensions spaced along a boom and fed by a transmission line. The target frequency is determined by the ratio of the lengths of the longest to the shortest dipoles. Most of the elements are arranged in a compact core with a diameter of 1 Km, forming a dense AA equipped with Phased Array Feeds (PAF), for which the elements are capable of forming multiple independent beams within a restrained integration time. PAFs are inherited from phase 0 experiment, ASKAP.

The key design parameters, like the Field of View (FoV), the frequency range, or the collecting area through the elements filling factor, are highly coupled, principally because the antennae are quasi-resonant, *i.e.* the area falls off as λ^2 .

SKA-Low can also observe larger scales using the auto-correlation of the primary beam from each station, the so-called dish mode, or, better for observing large scales, it could use dipoles as correlation elements instead of stations.

The rest of the configuration is arranged more sparsely, forming three equally spaced spiral arms, up to a maximum radius of 45 Km. The whole stations configuration is pictured on Figure 1.3, while the old core configuration is shown in Figure 1.4. Consider that the old core configuration planned to have 911 stations of which 855 in the core. The recent rebaselining has halved these numbers. The signal coming from each station's primary beam will be cross-correlated with the others in a central signal processing building. Data processing with SKA-Low is challenging for various reasons. First, the dipoles will

see the whole sky with strong polarisation features. So, the stations will always change shape as seen from the sky. Second, the ionosphere is non-isoplanatic for low frequencies and large baselines. Third, the sky is bright at all frequencies, while the EoR signal is very weak.

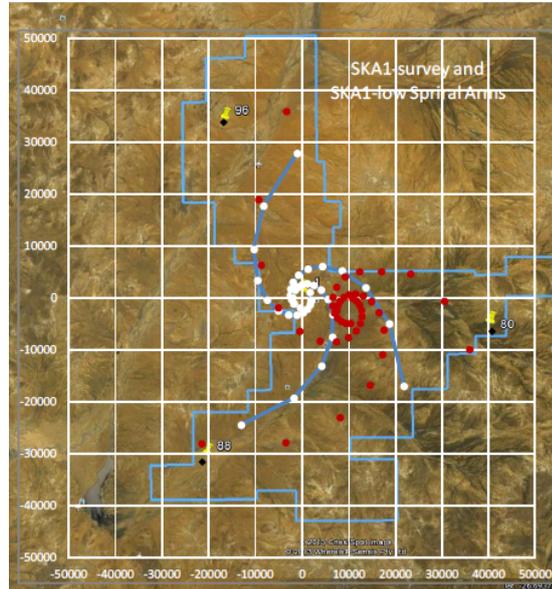


Figure 1.3: SKA1-Low configuration of stations (white dots) on a Google Earth image. Red dots are the previous configuration for the deleted SKA-Sur array, which incorporated the pathfinder ASKAP. Taken from Dewdney (2013).

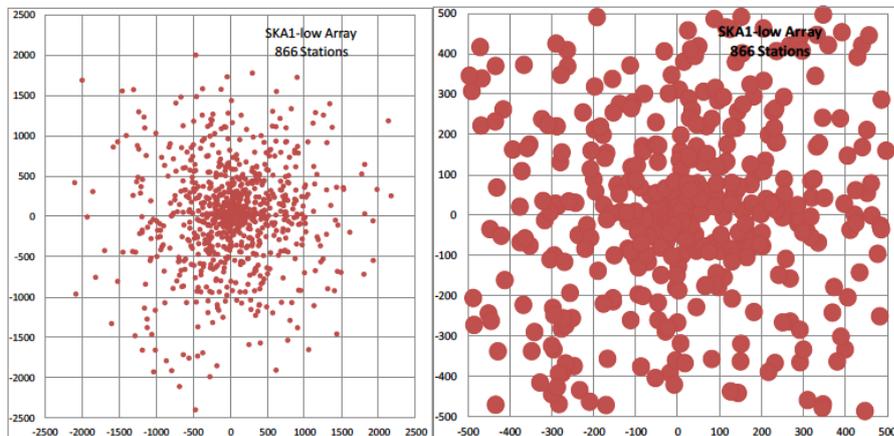


Figure 1.4: The SKA1-Low original core configuration including 866 35 m diameter stations. This number is currently halved, following the recent rebaselining procedure. Taken from Dewdney (2013).

1.1.2 SKA-Mid Overview

SKA-Mid will be able to measure 21 cm emission from HI in galaxies at lower redshifts ($\lesssim 3$) and will allow for a study of their evolution. Moreover, it

can provide a survey of the entire visible sky for pulsars with luminosity depth of 0.1 mJy Kpc^2 at 1400 MHz out to a distance of 10 Kpc. An interesting scientific opportunity is given by the possibility of performing 21 cm intensity mapping at redshifts of $z \sim 1 - 2$ (Santos et al., 2015) to detect BAOs at scales unprobed by SKA-Low because of the small FoV and the low frequency range explored.



Figure 1.5: Illustration showing the expected SKA-Mid South Africa side. Credits: <http://www.skatelescope.org/>

Other than in interferometer mode which is optimal to detect small scales, SKA-Mid can potentially observe in single-dish mode, where auto-correlations are used to probe large cosmological scales. The dishes are parabolic reflectors, like the ones in Figure 1.5. Each pointing made with dishes or feeds will provide a single pixel on the sky, giving more large scales modes.

The configuration will include the MeerKAT pathfinder, by sharing the same core and filling it with SKA antennae. Then there are the spiral arms which will be five in phase 2, as shown in Figure 1.6. The coverage in visibility space up to a maximum baseline of 100 Km has a good sensitivity but is patchy and incomplete, due to the Earth's rotation. On the other hand, this is very dense out to the 6 Km core, and the coverage pattern in visibility space has no holes within this radius.

The initial configuration of 250 dual-polarised antennae of 15 meters diameter each has been reduced by the 30% due to de-scoping. This includes 64 antennae from MeerKAT pathfinder. The distribution of antennae will place 40%, 55%, 70%, 81% and 100% of the total antennae within 0.4 Km, 1 Km, 2.5 Km, 4 Km, and 100 Km radius respectively. The frequency observation range is divided in five bands which can achieve different science targets and can still be re-defined (Bull, 2015). These will overlap with MeerKAT bands. The dishes can be equipped with PAFs, so a certain number of pointings can be simultaneously performed to cover a given sky area. In this case the beams will overlap below a certain critical frequency, resulting in a loss of sensitivity.

Villaescusa-Navarro et al. (2014) found that SKA-Mid should be able to resolve the 21 cm power spectrum up to scales of at least $1 h \text{ Mpc}^{-1}$ with 100 hours of observation.

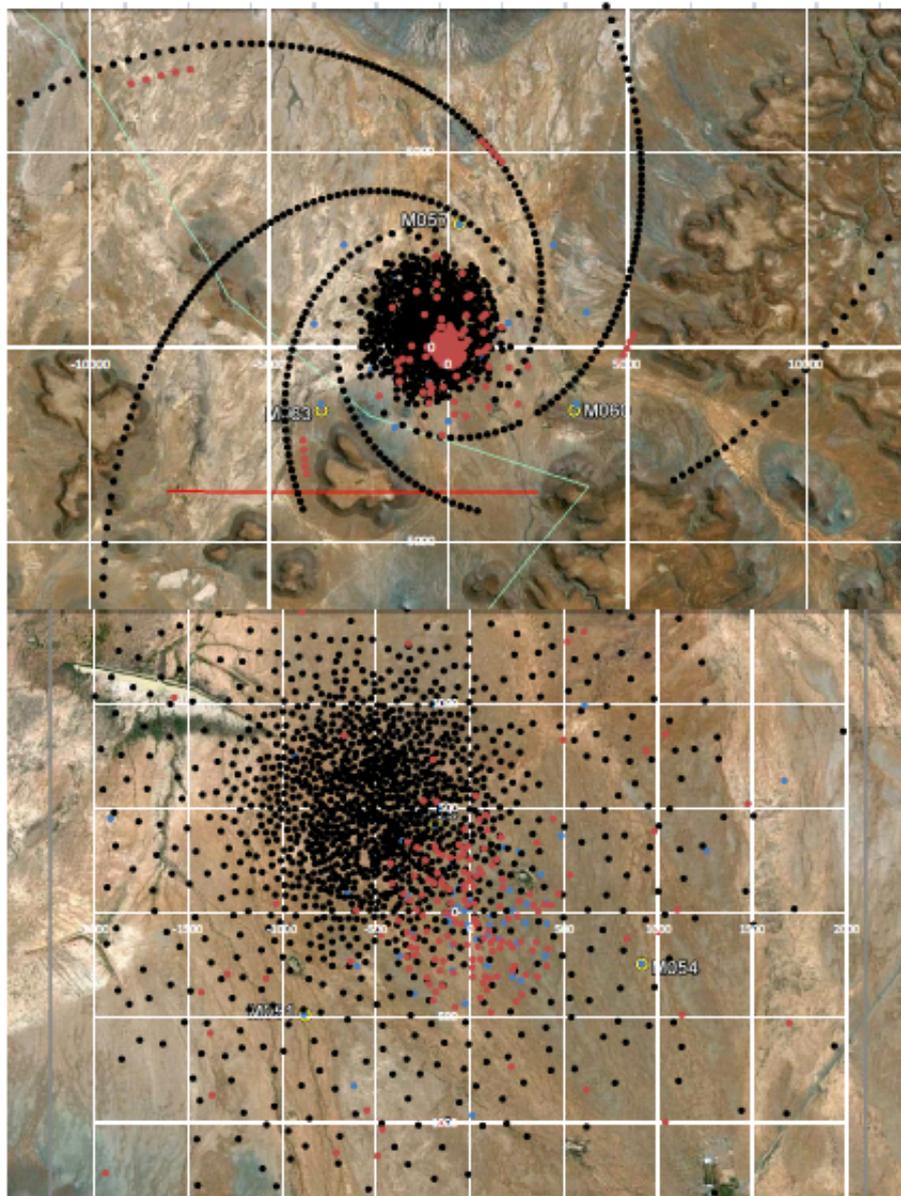


Figure 1.6: The generic SKA2-Mid configuration inglobing the SKA1-Mid core. Black dots are SKA2 antennae, red dots are SKA1-Mid antennae, blue dots are MeerKAT antennae. Taken from Dewdney (2013).

1.1.3 Comparison with Other Radio Telescopes

Currently, the planned SKA telescopes are the only radio telescopes with enough collecting area and sufficient resolution to observe 21 cm lensing.

Table xxx: Parameters for Comparable Telescopes															
		eMERLIN	JVLA	GBT	GMRT	Parkes MB	LOFAR	FAST	MeerKAT	WSRT	Arecibo	ASKAP	SKA1-survey	SKA1-low	SKA-mid
$A_{\text{eff}}/T_{\text{sys}}$	m^2/K	60	265	276	250	100	61	1250	321	124	1150	65	391	1000	1630
FoV	deg^2	0.25	0.25	0.015	0.13	0.65	14	0.0017	0.86	0.25	0.003	30	18	27	0.49
Receptor Size	m	25	25	101	45	64	39	300	13.5	25	225	12	15	35	15
Fiducial frequency	GHz	1.4	1.4	1.4	1.4	1.4	0.12	1.4	1.4	1.4	1.4	1.4	1.67	0.11	1.67
Survey Speed FoM	$\text{deg}^2 \text{m}^4 \text{K}^{-2}$	9.00×10^7	1.76×10^8	1.14×10^8	8.13×10^8	6.50×10^8	5.21×10^8	2.66×10^8	8.86×10^8	3.84×10^8	3.97×10^8	1.27×10^8	2.75×10^8	2.70×10^7	1.30×10^8
Resolution	arcsec	$10-150 \times 10^3$	1.4 - 44	420	2	660	5	88	11	16	192	7	0.9	11	0.22
Baseline or Size	km	217	1-35	0.1	27	0.064	100	0.5	4	2.7	225	6	50	50	200
Frequency Range	GHz	1.3-1.8, 4-8, 22-24	1-50	0.2-50+	0.15, 0.23, 0.33, 0.61, 1.4	0.44 to 24	0.03-0.22	0.1-3	0.7-2.5, 0.7-10	0.3-8.6	0.3-10	0.7-1.8	0.65-1.67	0.050-0.350	0.35-14
Bandwidth	MHz	400	1000	400	450	400	4	800	1000	160	1000	300	500	250	770
Cont. Sensitivity	$\mu\text{Jy}\cdot\text{hr}^{-1/2}$	27.11	3.88	5.89	6.13	16.26	266.61	0.92	3.20	20.74	0.89	28.89	3.72	2.06	0.72
Sensitivity, 100 kHz	$\mu\text{Jy}\cdot\text{hr}^{-1/2}$	1714	388	373	411	1029	1686	82	320	830	89	1582	263	103	63
SEFD	Jy	46.0	10.4	10.0	11.0	27.6	45.2	2.2	8.6	22.3	2.4	42.5	7.1	2.8	1.7

Notes to Table				
eMERLIN	Frequencies non-contiguous			
JVLA	Multiple antenna configurations			
GBT	Single dish			
GMRT	Frequencies non-contiguous			
Parkes MB	Multi-beam (13)		Frequencies non-contiguous	
LOFAR	Parameters for all NL stations		Frequencies non-contiguous	
FAST	Single dish		Under construction	
MeerKAT	SKA Precursor		Under construction	
WSRT	Frequencies non-contiguous			
Arecibo	Single dish			
ASKAP	SKA Precursor		Multi-beam (36)	Under construction
SKA1-survey	Multi-beam (36)		Mixed 12-m & 15-m dishes	FoV based on 15-m dishes
SKA1-low	Planned			
SKA-mid			Mixed 13.5-m & 15-m dishes	FoV based on 15-m dishes
Notes: All	Fiducial frequency: Most Parameters		$\Omega_{\text{FoV}} = (\pi/4)(66\lambda_r/D_{\text{dish}})^2$	Gray shading: <400 MHz capable
(cont'd)	SEFD derived from $A_{\text{eff}}/T_{\text{sys}}$		Sensitivity derived from SEFD & BW	System efficiency assumed 100%.

Figure 1.7: Table comparing various performance parameters for actual planned or built radio-telescopes. Taken from Dewdney (2013).

The SKA design has been driven by the need for high sensitivity and high survey speed. The table pictured in Figure 1.7 collects a number of performance parameters for some planned radio telescope. We can note that SKA1-Low covers a similar frequency range to LOFAR or MWA, but its sensitivity is more than an order of magnitude larger, mainly because most of the collecting area is in a very compact array. The PAFs will provide a two orders of magnitude larger FoV than the one provided by JVLA, despite the comparable sensitivity indicated by the SEFD, the System Equivalent Flux which is the ratio between the effective collecting area and the system temperature. SKA-Mid can cover a contiguous frequency coverage from 350 MHz to 20 GHz, like the upgraded JVLA, but with a much bigger spanned survey area.

1.2 HI Intensity Mapping

The redshift information available from 21 cm surveys would provide further improvements on cosmology and its constraints Jarvis et al. (2015), like Dark Energy (DE) evolution for example. Telescopes probing the sky between a rest frequency of 1420 MHz and 250 MHz would be able to detect galaxies up to redshift 5 (Santos et al., 2015). The problem is that this emission line is usually quite weak: at $z = 1.5$, most galaxies with a HI mass of $10^9 M_\odot$ will be observed with a flux density of $\sim 1 \mu\text{Jy}$ using the HI line. As showed by Alonso et al. (2014), experiments with sensitivities better than $10 \mu\text{Jy}$ over 10 kHz channels are required to provide enough galaxies to beat shot noise and become cosmic variance dominated. Although “near term” radio telescopes such as ASKAP and MeerKAT should be able to achieve such sensitivities on deep single pointings, it will require a much more powerful telescope such as SKA2 phase, to integrate down to the required sensitivity over the visible sky in a reasonable amount of time. This would imply that one would need to wait until then to use radio telescopes for cosmology.

Galaxy surveys are threshold surveys in that they set a minimum flux above which galaxies can be individually detected. Instead we could consider measuring the integrated 21 cm emission of several galaxies in one angular pixel on the sky and for a given frequency resolution. For a reasonably large 3D pixel we expect to have several HI galaxies in each pixel so that their combined emission will provide a larger signal. Moreover, as will be done in this work, we can use statistical techniques similar to those that have been applied for instance to CMB experiments, in order to measure quantities in the low signal to noise regime. By not requiring the detection of individual galaxies, the specification requirements imposed on the telescope will be much less demanding. This is what has been commonly called an Intensity Mapping (IM) experiment, which treats the 21 cm emission as a continuous three-dimensional background field, without detecting individual galaxies, which are unresolved in angle coordinates, not in frequency.

Hence, galaxies are resolved in frequency but not in angular position. This means that a reasonably sized telescope with a relatively large beam (equivalently, relatively low resolution) can perform high precision cosmological mea-

surements (Santos et al., 2015). Historically, IM has been introduced in seminal papers (Loeb & Wyithe, 2008; Wyithe & Loeb, 2007) and is based on the presence of neutral hydrogen before and during EoR. In fact, reionization eliminated most of the HI in the Universe, but there could be overdense regions in which it still resides in galaxies. This is seen from damped Ly α systems in quasar spectra. So, if pre-reionization signal can give insights on the topology of ionized regions and on HI bubble sizes, the post-EoR signal is intrinsically connected to the clustering of collapsed haloes based on the underlying matter density field⁶.

1.2.1 Searching for Baryonic Acoustic Oscillations at Low Redshifts

Other than 21 cm lensing, which will be explored through this whole work, the other important application for IM at low redshifts is Baryonic Acoustic Oscillations (BAOs) detection. In this short paragraph we will give a basic idea about the relevance of this measurement. BAOs are generated by the same physics producing the acoustic peaks on the CMB and have a characteristic wavelength which is set by the sound horizon at recombination. This makes them a perfect standard ruler that allows for testing the geometry of the Universe and its matter content, constraining also the equation of state of Dark Energy. To have an idea of the resolution requirements to measure BAOs, consider that the third peak (the last available one before non-linearities destroy the signal) has a comoving wavelength of $35h^{-1}$ Mpc, which requires at $z \sim 1.5$ a resolution of 20 arcmin. Such an instrument could provide very good DE constraints, making it a potential alternative to more traditional galaxy surveys. Examples of experiments alternative and complementary to SKA aimed to detect uniquely BAOs are, among the others, CHIME⁷ and BINGO (Battye et al., 2013b). Chang et al. (2008) shown that the Green Bank Telescope (GBT) measurements, with removing the frequency smooth foreground component, demonstrate that such experiments are feasible.

1.3 The Need of Simulations

As already stated in the general introduction, there are a lot of practical issues that must be addressed in order to maximize the amount and quality of the scientific information that can be extracted. Quantifying the exact statistical and systematic uncertainties for a practical experiment analytically can be an unsurmountable task: simulations must be used, describing both the cosmological signal we expect to measure and all other processes

⁶Arguments about the robustness of the tracing HI-density matter have been made. It is believed that damped Ly α systems are relatively low-mass systems and so should have a low bias (Wyithe & Loeb, 2007). Moreover fluctuations in the ionizing background lead to spatial variation in the HI content only at the percent level, since after reionization the mean free path for ionizing photons increases, making the background uniform (Wyithe & Loeb, 2009). This avoids the generation of signal that would mimic cosmological signatures.

⁷<http://chime.phas.ubc.ca/>

(e.g. foregrounds, instrumental effects, etc.) that may have a significant effect on the recovered data (Santos et al., 2015). Ideally we would want these simulations to yield the most realistic state-of-the-art description possible, however this is not always feasible if a large number of independent realizations need to be generated in order to quantify the aforementioned uncertainties. A compromise between computational speed and complexity must be met, so that enough simulations can be run, while correctly reproducing the relevant physics.

There is a number of exhaustive works on this topic, which have the scope to provide fast and accurate methods to generate mock galaxy distribution (Tassev et al., 2013; White et al., 2014). In order to simulate a proper HI distribution populating a given Dark Matter halo N-body or hydrodynamical simulations are required, and these need to be performed at relatively high resolution to resolve the smallest haloes that can host HI (Villaescusa-Navarro et al., 2014).

In general a good IM simulation has to be simple and needs to be complemented by a Halo Occupation Distribution model in order to cross-correlate HI properties with those of a realistic galaxy population. Moreover it should be bound to reproduce observable quantities at high and low redshifts, and have to address the scale dependence of the bias⁸ existing between HI and matter distribution and the amount of HI residing outside galaxies. Moreover a good simulation code must provide realistic HI observed maps, once the instrument and noise specifications are known.

1.4 The Foregrounds Contamination Problem

The results of this work will depend mainly on the capability of performing a foreground removal which does not compromise too much the detected 21 cm brightness temperature field with intensity mapping. The problem consists in how to develop cleaning methods to remove everything that is not the HI signal at a given frequency. This in turn also impacts on the calibration requirements of the instrument. For example, the situation is slightly complicated by the frequency dependence of beam shape for any real radio instrument: this would cause mode mixing, and its impact on foreground subtraction has been discussed in Morales et al. (2006).

Foregrounds can be several orders of magnitude bigger than the 21 cm signal, and extracting the Large Scale Structure (LSS) signal using intensity mapping, without identifying the HI point sources, represents a valuable challenge. This would depend mainly on the smoothness in frequency of the foregrounds and on how well these are modeled with respect to the true disturbing emis-

⁸We will see in Chapters 3 and 4 that the HI traces the matter density field, and this is true up to a given bias which is scale-dependent. It can assumed to be constant at high redshifts (Villaescusa-Navarro et al., 2014), while it is not constant at lower redshifts, since we are averaging over many galaxies. (Font-Ribera et al., 2012) with the SDSS-BOSS survey, and using the properties of damped Ly α systems, found that HI mainly resides in high-mass haloes. This generates a tension between simulations and models, causing the bias factor to be a factor 2 higher than expected.

sion. After reionization the HI density is low. This would produce a much weaker signal with respect to the one produced by a fully neutral Inter-Galactic Medium (IGM). At lower frequencies the foregrounds are expected to be huge and various orders of magnitude bigger than the HI signal. It is hence extremely important for observing EoR redshifts to develop foreground removal techniques which will clean the observed HI. At the frequencies corresponding to a 21 cm emission at $z \sim 1 - 3$ the amplitude of the foregrounds (like Galactic synchrotron emission) is smaller than or comparable to such a signal at the same frequencies, and this represents a "sweet spot" for 21 cm observations.

Simulating foreground cleaning is not easy: while it is possible to use a few datasets and certain empirical models to produce conservative realizations of the radio foregrounds (de Oliveira-Costa et al., 2008), the lack of full-sky multi-frequency data prevents us from developing truly realistic simulations of the radio sky.

The basic idea lying behind foreground cleaning consists in taking advantage of the smooth frequency structure as well as other statistical properties of the foregrounds which are significantly different from those of the cosmological signal (Di Matteo et al., 2002; Wang et al., 2006; Morales et al., 2006; Jelić et al., 2008). Although a lot of work in terms of simulations and testing cleaning techniques has already been done, we still face huge challenges ahead, in particular if we want to use this signal for high precision cosmology (Santos et al., 2015). Increasingly realistic large simulations should be developed to try to test the limitations of the intensity mapping measurements. This should include as many instrumental effects as possible in order to account for possible contamination from the calibration process. We will need to start analysing real data in order to improve and build up our knowledge towards the SKA.

1.4.1 Foregrounds Classification

At low frequencies, the main contribution is due to synchrotron emission from our Galaxy, while free-free emission has a minor importance (Liu & Tegmark, 2012). Here we will briefly review the most used foreground models found in literature.

Let us consider a given discrete grid whose pixel values represent the intensity of our 21 cm field. The Galactic synchrotron emission at each pixel is modeled as a power-law with spectral index α that may vary from pixel to pixel, namely

$$x(\nu) = A_{\text{sync}} \left(\frac{\nu}{\nu^*} \right)^{-\alpha} \quad (1.1)$$

(Liu & Tegmark, 2012), in which $A_{\text{sync}} = 335.4$ K, α is Gaussian distributed at each pixel with mean $\alpha_{\text{sync}} = 2.8$ and deviation $\Delta\alpha_{\text{sync}} = 0.1$, and ν^* is the pivot frequency (Wang et al., 2006). The mean value of this field, *i.e.* the averaged field on the pixels map, is

$$m^{\text{sync}}(\nu) = \langle x(\nu) \rangle = A_{\text{sync}} \left(\frac{\nu}{\nu^*} \right)^{-\alpha_{\text{sync}} + \Delta\alpha_{\text{sync}}^2 \ln(\nu/\nu^*)/2} . \quad (1.2)$$

The foreground covariance is defined without considering the subtraction of the mean term, since that is added to the covariance of the cosmological signal. This is

$$C^{\text{sync}}(\nu, \nu') = \langle x(\nu)x(\nu') \rangle = A_{\text{sync}}^2 \left[\frac{\nu\nu'}{(\nu^*)^2} \right]^{-\alpha_{\text{sync}} + \Delta\alpha_{\text{sync}}^2 \ln[\nu\nu' / (\nu^*)^2] / 2}. \quad (1.3)$$

The synchrotron radiation is partially linearly polarised. Its polarised part will be affected by the Faraday effect, a rotation of the polarization angle caused by the Galactic magnetic field and the optically thick interstellar medium (Rybicki & Lightman, 1986; Waelkens et al., 2009). Since the dependence of this effect on frequency is quite strong for the frequency range that pertains to IM, any leakage from the polarized part into the unpolarized measurements of the cosmological signal would generate a very troublesome foreground to subtract (Jelić et al., 2010; Moore et al., 2013), as will be better seen in Section 1.4.3.

The free-free emission is modeled analogously to the synchrotron emission, so

$$\begin{aligned} m^{\text{ff}}(\nu) &= A_{\text{ff}} \left(\frac{\nu}{\nu^*} \right)^{-\alpha_{\text{ff}} + \Delta\alpha_{\text{ff}}^2 \ln(\nu/\nu^*)/2}, \\ C^{\text{ff}}(\nu, \nu') &= A_{\text{ff}}^2 \left[\frac{\nu\nu'}{(\nu^*)^2} \right]^{-\alpha_{\text{ff}} + \Delta\alpha_{\text{ff}}^2 \ln[\nu\nu' / (\nu^*)^2] / 2}, \end{aligned} \quad (1.4)$$

with $A_{\text{ff}} = 33.5$ K, $\alpha_{\text{ff}} = 2.15$, and $\Delta\alpha_{\text{ff}} = 0.01$. Free-free emission is caused by free electrons accelerated by ions, and thus traces the warm ionised medium.

Other than Galactic Foregrounds we will have to deal with extragalactic radio sources, which are historically defined as a source of further noise. These can be classified into two different categories: bright radio galaxies, such as Active Galactic Nuclei (AGN), and “normal” star-forming galaxies. The spatial distributions of these two types should be qualitatively different, the former being less dominated by gravitational clustering and more by Poisson noise. We will see how we can take advantage of this signal in Chapter 6, when we will study lensing reconstruction at low redshifts.

Other possible foreground sources are atmospheric noise, radio frequency interference and line foregrounds, caused by line emission from astrophysical sources in other frequencies. Due to the spectral isolation of the 21 cm line, together with the expected low intensity of the most potentially harmful lines (such as OH at $\nu_{\text{OH}} = 1600$ MHz), the HI signal should be very robust against line confusion (Santos et al., 2015).

1.4.2 To Be or Not to Be Blind?

In this little section we will shortly review the most adopted foreground cleaning methods, in order to give a more precise idea of the problem.

21 cm observation will provide high-resolved data in frequency, and the best foreground subtraction techniques should operate primarily using line-of-sight (LoS) information, given also the nature of the foregrounds, as already

stated in the introduction of this section. Most LoS techniques are blind, in the sense that they do not require any knowledge about the foreground model. This approach takes into account the smooth nature of the foreground spectra, separating the rapidly fluctuating cosmological signal by subtracting off a predetermined set of low-order polynomials (Liu et al., 2009; Bowman et al., 2009), or by imposing a predetermined filter in Fourier space (Petrovic & Oh, 2011; Paciga et al., 2011). It seems that this schemes could be advantageous especially at EoR redshifts, since foregrounds are fairly unconstrained observationally at these frequencies and most models are based on extrapolations and interpolations from other frequencies, where the instruments are optimized for different science goals. These methods leave an important level of post-foreground subtraction residuals and noise, which can be suppressed to a level smaller than the expected amplitude of the cosmological signal. Recently Wolz et al. (2014) studied the effectiveness of independent component analysis (in particular the implementation of FastICA⁹) for intensity mapping. By propagating the foreground removal residuals into the cosmological analysis, they showed that, while foreground cleaning may induce a residual bias on large angular scales which could prevent a full analysis based on the shape of the temperature power spectrum, robust features like the BAO scale should remain unaffected. This result is reasonable: most relevant foregrounds are (fortunately) exceptionally smooth and therefore it should be possible to distinguish them from the much “noisier” cosmological signal. Any foreground residual will probably be dominated by galactic synchrotron emission, which is most relevant on large angular scales.

(Liu & Tegmark, 2012) found that interesting constraints on foreground physics can not be found even with the most careful foreground spectrum measurements, because of the high levels of degeneracies between different foreground parameters. On the other hand the foreground spectra are so featureless and smooth that they could be characterized to a greater accuracy than necessary for foreground subtraction using only three or four independent parameters, and this bodes well for 21 cm cosmology and astrophysics¹⁰.

Non-blind schemes on the other hand could be performed iteratively until the models converge to the true measured foregrounds, and continually improve as measurements place increasingly strong constraints on foreground models. (Liu & Tegmark, 2011) proposed a non-blind inverse variance scheme in which the foregrounds are not treated as an additional noise to be removed but as a form of correlated noise. This method seems to achieve several advantages over a LoS polynomial subtraction methods often proposed in literature. First of all LoS polynomial foreground subtraction is non-optimal, and gives rise to larger error bars in final power spectrum than inverse-variance does. This is due to the fact that LoS methods project out low-order poly-

⁹<http://research.ics.aalto.fi/ica/fastica/>

¹⁰To be more precise, the authors performed a Principal Component Analysis (PCA) on foreground total covariance. The great accuracy obtained for 21 cm cosmology comes from the properties of the Eigenforegrounds, *i.e.* the vectors which satisfy the eigenvalue equation for the correlation matrix of the covariances computed at different frequencies. In fact, only a few Eigenforegrounds modes are enough to specify completely the foregrounds power spectra. See the cited paper for details.

mial modes destroying information. The inverse variance method on the other hand preserves all modes, even if it does downweight some of them substantially. Secondly, LoS methods contain residual noise and foreground biases in estimates of the power spectrum (which however can be easily quantified and removed for an inverse variance method). Third, foreground subtraction along the LoS may lead to correlated errors in power spectrum measurements (both horizontal and vertical bars), which can limit their usefulness for cosmological parameter estimation, being most effective on a narrow parallel to LoS Fourier modes (k_{\parallel}) range. The only disadvantage of the inverse variance method is its computational cost, since it requires the inversion ($\mathcal{O}(n_{\text{pix}}^3)$) of a $n_{\text{pix}} \times n_{\text{pix}}$ matrix, where $n_{\text{pix}} \sim n_{\perp}^2 n_{\parallel}$ is the grid volume of a data cube. LoS polynomial algorithm only requires the inversion of a $n_{\parallel} \times n_{\parallel}$ matrix.

(Dillon et al., 2013) presented an highly parallelizable method which reduces considerably the computational cost of the inverse variance scheme, since it scales as $\mathcal{O}(n_{\text{pix}} \log n_{\text{pix}})$.

Another non-blind method is the parametric eigenvalue algorithm developed by Shaw et al. (2014) for the CHIME experiment. It decomposes the data with help of statistical models for both foregrounds and 21cm signal. This algorithm leaves minor foreground residuals in the large modes of the power spectrum. Instrumental errors such as polarization leakage, beam deformations and calibration uncertainties can significantly affect the foreground removal by mode mixing effects. Shaw et al. (2015) advanced their description to polarized data considering a number of instrumental errors in their tests. For future SKA experiments, detailed studies including varying instrumental settings and the impact of the residuals on the power spectrum are required in order to minimize bias on cosmological results. Foreground subtraction for the SKA is treated in detail in Wolz et al. (2015) including realistic simulations.

1.4.3 Polarization Leakage

Sky polarization is an additional foreground source which has been well described in (Santos et al., 2015). The problem can be described by using the measurement equation formalism (Hamaker et al., 1996; Smirnov, 2011) that describes the propagation of the signal through an interferometric array. A pedagogical view can be presented by using the scalar form of the measurement equation (Sault et al., 1996), which relates the measured visibilities to the (I, Q, U, V) Stokes parameters that describe the true sky brightness distribution:

$$\begin{aligned} V_{ij}^{pp} &= \frac{1}{2} g_i^p g_j^p (I + Q) & V_{ij}^{qq} &= \frac{1}{2} g_i^q g_j^q (I - Q) \\ V_{ij}^{pq} &= \frac{1}{2} g_i^p g_j^q (U + iV) & V_{ij}^{qp} &= \frac{1}{2} g_i^q g_j^p (U - iV). \end{aligned} \tag{1.5}$$

Here V is the visibility between antennae placed at the position (i, j) and with polarizations (p, q) : seen under another point of view this is the cross-correlation

of electric fields measured by each antenna¹¹. g is the gain of the system, *i.e.* the actual response of the telescope to the input sky.

If the instrument is calibrated perfectly, the gains can be renormalised $g^p = g^q = 1$, so that the measured intensity is $\tilde{I}_{ij} = V_{ij}^{pp} + V_{ij}^{qq} = I$. This ideal case will not happen in reality, so the gains will have a little amount of uncertainty due to calibration errors, $g_{i,j}^{p,q} = 1 + dg_{i,j}^{p,q}$. This will cause an error in the measured intensity, namely $\tilde{I}_{i,j} = I + dI + dQ$ for small calibration errors, with $dI = I(dg_i^p + dg_i^q + dg_j^p + dg_j^q)/2$ the error due to instrumental calibration, and $dQ = Q(dg_i^p - dg_i^q + dg_j^p - dg_j^q)/2$ the polarization term which leaks the measured intensity. This will be zero if the error is the same for both p, q polarizations.

Although the polarization leakage is different for different instruments, typical values for leakages are below 1% and they tend to be reasonably stable over time scales of hours. In this case, the greatest contamination could come from off-axis leakage, *i.e.* signals entering the telescope from directions other than the pointing one. Their magnitude can be far greater (*i.e.* up to 30%), depending upon the observing frequency and their time variability. With respect to the other foregrounds, this leakage is not smooth in frequency, since the Q, U Stokes parameters undergo a Faraday rotation when the radiation goes through an ionized medium. In principle, this contamination can be modeled by looking at polarised point sources with the telescope, but in reality it represents the main limiting factor in IM measurements (Switzer et al., 2013).

Unlike the EoR case, both point-like and diffuse Galactic polarized emission may be problematic for intensity mapping at $z \sim 1 - 2$. The average polarization fraction of extragalactic radio sources is $\sim 5\%$ at 1.4 GHz (Tucci et al., 2004) with RM values up to a few tens of rad m^{-2} at high Galactic latitude where HI intensity mapping is carried out (Simard-Normandin et al., 1981; Taylor et al., 2009). The properties of Galactic synchrotron polarization are much less known at the frequencies relevant to HI intensity mapping. It is fairly observationally established that the spatial distribution of polarized intensity poorly correlates with total intensity at 1.4 GHz due to small scale structure present in the ionized interstellar medium (Gaensler et al., 2001; Bernardi et al., 2003; Sun et al., 2011). Observations of supernova remnants also show that objects further away than a few Kpc are completely depolarized at 1.4 GHz, indicating the presence of a polarization horizon beyond which diffuse polarization is no longer observable (Sun et al., 2011). The distance of such polarization horizon decreases at lower frequencies, down to a few hundreds pc at 150-300 MHz (Haverkorn et al., 2004; Bernardi et al., 2013), indicating that relativistic and thermal plasma are co-located in the interstellar medium (Burn, 1966). Typical RM values for Galactic polarization also decrease with decreasing frequencies. Given the complex spatial and frequency properties of Galactic polarization, extrapolations to the frequencies relevant for HI IM observations are fairly uncertain, although we expect that a significant improvement will happen in the next years due to new surveys.

¹¹For single dish $i = j$.

Chapter 2

Cosmology and Weak Gravitational Lensing

Since we are going to investigate how 21cm radiation coming from the Epoch of Reionization (and after it) is lensed and what cosmological information we can extract from it, we need to introduce what is gravitational lensing. In general, gravitational lensing happens when a light ray trajectory is deflected by a gravitational field placed between the radiation source and an observer. On cosmological scales the gravitational field is caused by the entire Large Scale Structures (LSS) mass distribution. A dimensionless quantity related to deflection, the convergence κ , is usually introduced to quantify the lensing effect. In the weak lensing approximation, *i.e.* when lensing is generated by LSS, this effect is small, so that $\kappa \ll 1$.

Weak gravitational lensing assumed a fundamental importance in modern cosmology, being a powerful tool to investigate the Λ CDM paradigm used to describe the observed Universe. It is indeed linked directly to the source LSS potential for lensing, and so it allows for a standard cosmological parameters measurement to be used additionally to Cosmic Microwave Background (CMB) or galaxy clusters measurements. This will break degeneracy of cosmological parameters. Moreover, it is directly linked to the gravitational potential, and thus is not sensitive to the light-mass biasing or to dynamics of complicated systems.

2.1 Theoretical Cosmology Background

Gravitational lensing is one of the strongest General Relativity's confirmations. So, in order to properly analyse it, we need to introduce briefly the relevant aspects of the standard cosmological model mostly adopted by the scientific community. In this section we will follow Bartelmann & Schneider (2001). Interested readers can consult Coles & Lucchin (2002); Dodelson (2003); Peacock (1999); Schneider (2015); Rowan-Robinson (2004) for further details.

2.1.1 The Friedmann-Robertson-Walker-Lemaitre Model

The standard cosmological model relies on Albert Einstein's General Relativity (GR) theory. GR describes the quadri-dimensional space-time as a dynamical field whose evolution is fully represented by Einstein's Field equations, which relate the source of the gravitational field, the local matter stress-energy tensor contained within the space-time, to local curved space-time geometry encapsulated into the Einstein tensor. This is

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu}. \quad (2.1)$$

Here G is Newton's Gravitational constant and c is the light speed in the vacuum, $g_{\mu\nu}$ is the metric tensor, which describes the relation among events in the space-time and it defines fundamental notions as distance, volume or curvature. The motion of a point in the space-time is represented by geodesics, whose dynamics is described by $G_{\mu\nu}$, the Einstein tensor, which is related to the Ricci tensor and the Ricci scalar. The first is a tensor containing derivatives of the metric, which describes how the space-time is curved respect to an Euclidean space-time expressed by the Minkowski metric $\eta_{\mu\nu}$; the second is the Ricci scalar, namely Ricci tensor's trace $R = g^{\mu\nu}R_{\mu\nu}$, and describes how much the space-time geodesic volume is modified respect to the Euclidean case. $T_{\mu\nu}$ is the stress-energy tensor and represents the source of space-time distortion. This is a symmetric tensor related to matter, radiation and non-gravitational force fields. Greek indices run from 0 to 3, while latin ones run only for spatial components, from 1 to 3.

Often, these equations are alternatively written as:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu} + \Lambda g_{\mu\nu}, \quad (2.2)$$

where Λ is the cosmological constant.

As already stated, the metric tensor defines the separation of two events, with coordinates dx^μ in space-time, namely $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$. The standard cosmological model constrains the form of this space-time element, by imposing isotropy and homogeneity in the background and on large scales. This implies that $ds^2 = c^2 dt^2 - a^2(t)dl^2$, where the scale function $a(t)$ ensures that every spatial region contracts or expands in the same way for every point in the Universe, so that background quantities describing the matter content of the Universe are a function of time only. The spatial element dl^2 can be flat or curved, with arbitrary coordinate origin and spherically symmetric. So, it can be generally written in spherical coordinates as

$$dl^2 = dw^2 + f_K^2(w) (d\phi^2 + \sin^2 \theta d\theta^2) = dw^2 + f_K^2(w) d\Omega^2. \quad (2.3)$$

$f_K(w)$ is the curvature function that, depending on whether the curvature K is positive, null, or negative, can be a trigonometric, linear or hyperbolic function of a general radial coordinate w . This is

$$f_K(w) = \begin{cases} K^{-1/2} \sin(K^{1/2}w) & (K > 0) \\ w & (K = 0) \\ (-K)^{-1/2} \sinh[(-K)^{1/2}w] & (K < 0) \end{cases}. \quad (2.4)$$

Now, if we define r to be the radius of the three-spheres by $f_K(w) = r$, the metric will be

$$dl^2 = \frac{dr^2}{1 - Kr^2} + r^2 d\Omega^2. \quad (2.5)$$

We note how this line element will define a different constant-curvature manifold for describing our Universe. If $K = 0$ we will have a flat manifold, *i.e.* a Minkowski space with an extra a^2 factor due to Universe evolution; if $K < 0$ the Universe will take the form of a three-dimensional saddle, or hyperboloid, at any given time; if $K > 0$ we will have a three-sphere manifold. Now we can apply this metric, known as Friedmann-Robertson-Walker (FRW) metric, to Einstein equations Eq. 2.2. $T_{\mu\nu}$ has to have the form of the stress-energy tensor of a homogeneous perfect fluid, characterised by its density $\rho(t)$ and its pressure $P(t)$. We then have two resulting independent equations, called Friedmann's equations, that express how the scale factor depends on time and how the curvature depends on the matter which fills space-time. These are

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} + \frac{\Lambda}{3}, \quad (2.6)$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right) + \frac{\Lambda}{3}. \quad (2.7)$$

These two equations can be combined in order to obtain another independent equation, which expresses the continuity of mass-energy $T^{\mu\nu}_{;\nu} = 0$, namely:

$$\dot{\rho} = -3\frac{\dot{a}}{a}\left(\rho + \frac{P}{c^2}\right). \quad (2.8)$$

Here we can see that, being a manifestation of the mass-energy, the pressure acts as a source of gravity. For a non-relativistic particle this pressure term, $P \propto nkT$, will be negligible compared to the density term, which is proportional to mass $\rho \propto nmc^2$, with n the number density. $\rho + P/c^2$ is the trace of the stress-energy tensor.

2.1.2 The Expansion of the Universe

Universe's expansion has been confirmed by Edwin Hubble, who discovered that neighbour galaxies are drifting apart with acceleration rate H , called Hubble constant. Scale factor's expansion rate $\dot{a}/a = H(z)$ is the Hubble parameter. Recent PLANCK observations (Planck Collaboration et al., 2015) have determined that the present day value is $H(t_0) = H_0 \simeq 67.8 \pm 0.77$ Km/s/Mpc. Usually the expression for the Hubble constant is presented as $H_0 = 100h$ Km/s/Mpc, so that $h = 0.67$.

Type Ia Supernova measurement made by Riess et al. (1998) and Perlmutter et al. (1999) confirmed that the expansion is accelerated, namely $\ddot{a} > 0$: this means that the distance between us and distant galaxies increases with time. The source of this acceleration is an energy source defined by the Cosmological constant term within the Λ CDM model. This term has negative pressure and it is the manifestation of the so-called Dark Energy. Scientists are making huge

efforts to investigate this component, since its nature and relation with the rest of known physics is unknown.

Hence, the Universe is expanding and photons, for which $ds^2 = 0$, are redshifted while they propagate from the source to the observer by Doppler effect. This means that the comoving coordinate distance between an emitting source and a comoving observer placed at the coordinate origin is constant:

$$w = \int_o^e dw = \int_{t_o}^{t_e} \frac{cdt}{a} = \text{const.} \quad (2.9)$$

From this we can define the redshift z as the ratio between emitted and observed light, namely

$$1 + z = \frac{\lambda_0}{\lambda_e} = \frac{a(t_0)}{a(t_e)} = a^{-1}(t). \quad (2.10)$$

So, higher redshifts represent earlier epochs since light travels with finite velocity c . By convention the scale factor at the present epoch t_0 is $a(t_0) = 1$.

If we consider the first Friedmann equation for a flat Universe with no cosmological constant, it is possible to define the critical density of the Universe $\rho_c = 3H_0^2/8\pi G = 1.9h^2 \times 10^{-29} \text{ g/cm}^3$. The parameter that defines the ratio of the actual density for given i species respect to the critical one is the density parameter $\Omega_i = \rho_i/\rho_c$. Now, we can write down the first Friedmann equation as

$$H^2(z) = H_0^2 \left[\frac{\rho_i(z)}{\rho_c} - \frac{Kc^2}{H_0^2} (1+z)^2 + \frac{\Lambda}{3H_0^2} \right]. \quad (2.11)$$

Thus we have to define what is the density for every considered component of the Universe. Usually these are resumed in matter (baryons and Dark Matter) and radiation (photons and neutrinos). The equation of state for density can be written as

$$\rho(t) = a^{-n_i}(t)\rho_0, \quad (2.12)$$

where we have $n_i = 3$ or $n_i = 4$ for matter and radiation respectively. For non-relativistic matter $P = 0$, while for relativistic matter $P = \rho c^2/3$. We could absorb such notation for the curvature and DE part in Eq. (2.11), so that they are described by an equation of state with $n_i = 2$ and $n_i = 0$ respectively. Note that the DE term acts like a vacuum energy term. So we have that $\rho_i(z) = \rho_c \Omega_i (1+z)^{n_i}$, $\Omega_K = -Kc^2/H_0^2$, and $\Omega_\Lambda = \Lambda/3H_0^2$. We note that because of this the DE term can be related a proper energy density $\rho_\Lambda c^2$, but this is not the same for the curvature term. Finally we can write the evolution equation for the Hubble parameter:

$$H(z) = H_0 \sqrt{\Omega_r (1+z)^4 + \Omega_m (1+z)^3 + \Omega_K (1+z)^2 + \Omega_\Lambda}. \quad (2.13)$$

Usually this result is expressed in terms of the Hubble function $E(z) = H(z)/H_0$. In particular, at $z = 0$, $E(0) = 1$ and so

$$\sum_i \Omega_i = 1, \quad (2.14)$$

where the index i stands for m, r, K, Λ , *i.e.* matter, radiation, curvature and DE. Usually because of this, one can often find the curvature density parameter written as $-\Omega_K = 1 - \Omega_i$. In order to parametrise our ignorance on the nature of DE, the Λ term in Eq. (2.13) can be written as a function of $(1 + z)$, namely $\Omega_\Lambda(1 + z)^{3f(z)}$. Here we have defined

$$f(z) = -\frac{1}{\ln(1+z)} \int_{-\ln(1+z)}^0 [1 + \omega(a)] d \ln a, \quad (2.15)$$

where $\omega(a)$ is the Dark Energy equation of state parameter, *i.e.* the ratio between DE pressure and density. When $\omega(a) = -1$ we get again the usual Λ term.

Friedmann equations allow us to explore several kind of possible universes if we vary the fundamental cosmological parameters contained in Friedmann's equations for the evolution of the scale factor. For example an open ($K < 0$) Universe, with no matter, radiation or Λ is known as Milne Universe, and has $H(t) \propto t^{-1}$. The flat $K = 0$, matter-dominated $\Omega_m = 1$ case is the Einstein-de Sitter Universe and has $H(t) \propto 2t^{-1}/3$. So this model corresponds to early times for a Universe with non-relativistic matter. If we consider $\Omega_K < 0$, $\dot{a} = 0$ we have a closed Universe model. If we go back in time in a flat Universe dominated by radiation we have that $H(t) \propto t^{-1}/2$, and this applies to sufficiently early times. Note that a flat Universe does not asymptote to a constant value, but simply grows more slowly than an open case. An Universe dominated by DE is called de Sitter Universe, and grows exponentially since $a(t) \propto e^{t\sqrt{\Lambda/3}}$.

2.1.3 The Λ CDM Universe and its Evolutionary Phases

Eq. (2.13) implies that the Universe went through several evolutionary phases with $T = (1 + z)T_0$, *i.e.* with always higher temperatures, or energies, as we go back in Universe's history. The standard cosmological model describes our real Universe with the actual observed values for the density parameters. Following PLANCK results (Planck Collaboration et al., 2015), $\Omega_K \simeq 0$, $\Omega_m \simeq 0.31$, $\Omega_\Lambda \simeq 0.69$, and $\Omega_r \simeq 10^{-5}$. Notice that nearly 5% of the matter contribution constitutes the baryonic matter studied by known physics. The rest is made of Dark Matter (DM). Without involving any controversial modification to gravity's laws, evidences for its existence have been found analysing galaxies light-curves, as seen in Rubin & Coyne (1988), or from simply applying the Virial theorem to some observed galaxy cluster (Zwicky, 1933). More strongly, this dominating contribution to the total Universe matter is measured from the distances between two CMB peaks, since this is sensitive to the matter density at recombination. DM gives the main contribution to the gravitational potential and its elementary nature is still unknown. It is usually thought as a cold, *i.e.* non-relativistic, and collisionless component, while all the other Universe components interact via electromagnetic, gravitational or weak nuclear forces.

We can investigate Λ CDM Universe evolutionary phases by studying the behavior of the scalar factor with cosmic time. This well agrees with the widely-adopted Big Bang Theory (BBT). The Big Bang existence proof, *i.e.* a time in which $a(t) = 0$, is usually inferred from the existence of Cosmic Microwave

Background (CMB), a relic of the time in which the Universe become transparent respect to the photons. This time is called Recombination and it is explained by an early, hot phase of the Universe. The Big Bang Theory is confirmed by several astrophysical evidences (Boerner & Ehlers, 1988) and agrees with the Standard Model (SM) of elementary particles, since every elementary particle that composes the known matter has been formed in well determined Universe evolution phases. As the temperature rises, the average energy of particles increases, allowing more and more interactions to take place. In fact, any interaction among particles will depend on interaction rates and particles masses. Moreover, BBT is also able to predict how the nucleosynthesis of lightest elements occurred during Universe's history. Interested readers can find detailed discussion about Big Bang Theory and its history in several books and references, as the ones mentioned at the beginning of this chapter. In order to solve some problems due to the existence of an initial singularity and to explain why the Universe is so homogenous and isotropic, we need also another phase called Inflation. We will see, in a few paragraphs, that this is a crucially important accelerated phase, because it is putting the seeds of all structures we observe today.

We could take a look at every evolutionary phase by studying the behaviour of the density parameter with time. In fact,

$$\Omega_i(z) = \frac{\rho_i(z)}{\rho_c(z)} = \frac{\rho_c \Omega_i (1+z)^{n_i}}{\rho_c E^2(z)} = \Omega_i \frac{(1+z)^{n_i}}{E^2(z)}. \quad (2.16)$$

We can see on Figure 2.1 the thermal evolution for a flat Universe with radiation, matter and DE. After Inflation the Universe went through a Radiation-dominated era (RDE), which is really important to understand elementary particles formations and how their interaction decoupled one from each other, agreeing with SM. This is also the period in which happened the primordial Nucleosynthesis, where the first light elements have been formed. The point at $z \sim 3700$ in which matter and radiation densities are the same is called Equality. Then the Universe experienced a Matter-dominated era (MDE), during which Recombination happened. During current epoch we are evolving to a Universe which is exponentially accelerating due to Λ dominant term. Consider that at low redshifts, and generally for early times, radiation and curvature contributions are neglected.

Integrating the inverse of Eq. (2.13) we can also determine the age of the Universe, since $dt = \dot{a}^{-1} da = (aH)^{-1} da$. The Universe is nearly 13.7 billion years old. Notice that this age is different, smaller, from the one we should obtain by simply considering the Hubble time $1/H_0 \sim 14$ Gyr. This is because we are taking into account the presence of a component that accelerated Universe expansion and dominates its evolution in current epoch. So the Universe is not evolving as a simple flat and matter dominated FRW Universe.

2.1.4 Distances in Cosmology

The proper distance, which is the distance measured by the travel time of a light ray propagating from a source at z_e to an observer at z_o with $z_o < z_e$, is

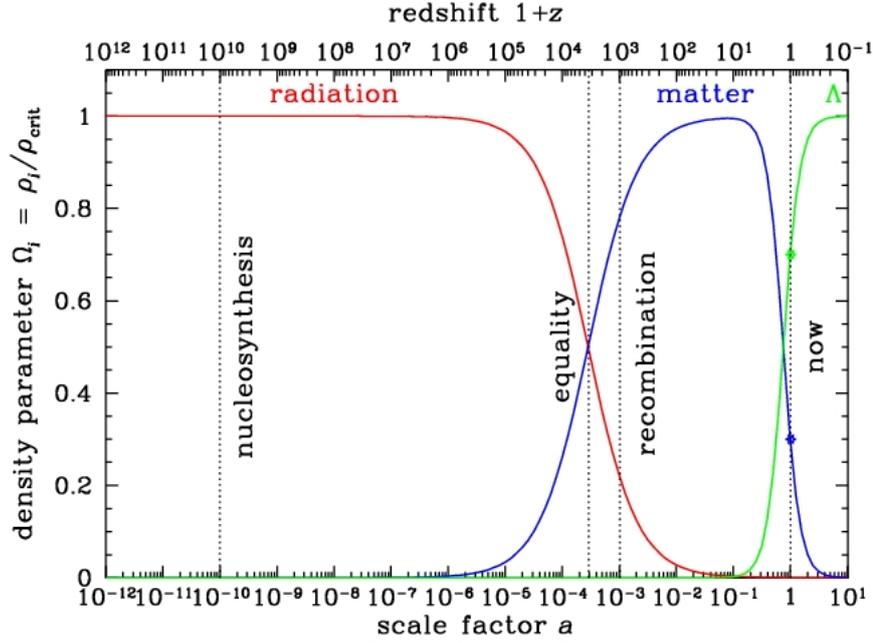


Figure 2.1: The density dependence of the various Ω_i components of the standard Λ CDM Universe.

defined as

$$dD_p = -cdt = -c\dot{a}^{-1}da = -c[aH(a)]^{-1},$$

since $ds^2 = 0$. Integrating it we get

$$\begin{aligned} D_p(z_o, z_e) &= c \int_{a(z_o)}^{a(z_e)} \frac{da}{aH(a)} \\ &= \frac{c}{H_0} \int_{a(z_o)}^{a(z_e)} \left[a^{-2}\Omega_r + a^{-1}\Omega_m + \Omega_K + a^2\Omega_\Lambda \right]^{-1/2} da. \end{aligned} \quad (2.17)$$

The comoving distance is the distance between a source and an observer that is comoving with the cosmic flow along the line-of-sight direction and is synchronised at $a(0)$. This is the proper distance scaled with a factor a : $cdt = -adw$. So, we have $dw = -cdt/a = -cda[a^2H(a)]^{-1}$. Then

$$\begin{aligned} \chi_c(z_o, z_e) &= \frac{c}{H_0} \int_{z_e}^{z_o} \left[\Omega_r + a\Omega_m + a^2\Omega_K + a^4\Omega_\Lambda \right]^{-1/2} da \\ &= d_H \int_0^{z_e} \frac{dz}{E(z)}, \end{aligned} \quad (2.18)$$

where we translated this integral in redshift coordinates and used Eq. (2.13). $d_H = c/H_0 \sim 3h^{-1}$ Gpc is the Hubble distance. The comoving distance can be easily generalised into transverse comoving distance, *i.e.* the distance between two events at the same redshift and separated on the sky by $d\theta$. If we apply

$cdt = -adw$ for a FRW metric and the definition of the trigonometric function $f_K(w)$ Eq. (2.4), we get

$$\chi = f_K(\chi_c) = \begin{cases} \Omega_K^{-1/2} \sin(\Omega_K^{1/2} \chi_c/d_H) & (K > 0) \\ \chi_c & (K = 0) \\ (-\Omega_K)^{-1/2} \sinh[(-\Omega_K)^{1/2} \chi_c/d_H] & (K < 0) \end{cases} \quad (2.19)$$

The angular-diameter distance is defined to be the ratio between its proper physical diameter and the angle $d\theta$ that it subtends on the sky. Namely,

$$\mathcal{D}(z) = \frac{D_p}{d\theta} = \frac{\chi}{(1+z)}, \quad (2.20)$$

where for small angles $D_p = \chi a_e d\theta$. This means

$$d_p = a\chi = \int_0^r \frac{adr'}{\sqrt{1-Kr'^2}}. \quad (2.21)$$

The luminosity distance links the bolometric luminosity of a source to its bolometric flux. This latter is defined as $L/[4\pi d_L^2(z)]$. If we consider a spherical coordinate system with the origin at the source at radial coordinate χ , emitting at time t_e , and observed at t_0 , the total area of the sphere subtended by it is

$$A = \int a_0^2 \chi^2 d\Omega = 4\pi\chi^2. \quad (2.22)$$

So, considering a $(1+z)^2$ factor due to Doppler shift and to decreasing of the number of photons, we have $F = L/[4\pi\chi^2(1+z)^2]$. So,

$$d_L = (1+z)\chi = (1+z)^2 \mathcal{D}. \quad (2.23)$$

Now we can finally define the distance to horizon, *i.e.* the maximum distance from which an information could have traveled to the observer. This represents the size of the observable Universe, over which events are causally disconnected. Let us convert the comoving distance to r_h into a proper distance by multiplying it with a scale factor $a(t)$, so that

$$d_h(t) = a(t)\chi(t) = a(t) \int_0^t \frac{cdt'}{a(t')} = a(t)c\eta.$$

Here we have defined the conformal time η that has passed since $a(t) = 0$. This can be written as

$$d_h(z) = \frac{d_H}{1+z} \int_z^\infty \frac{dz'}{E(z')} = \frac{\chi(z)}{1+z}. \quad (2.24)$$

Note that this distance is bigger than d_H : today $d_H = 3h^{-1} \text{ Gpc} \simeq 3.48 \text{ Gpc}$, while $d_h = 14.6 \text{ Gpc}$. This is due to the expansion of the Universe.

2.1.5 Large Scale Structure Formation

The real Universe is not exactly homogeneous and isotropic. At small scales, clustering effects are non-negligible: for example, near a black hole, the more appropriate metric is Schwarzschild metric. First-order perturbation theory can be used to find approximate solutions to Einstein equations. These perturbations are the seeds of the Large Scale Structure (LSS) we see and observe every day. They have been originated by quantum fluctuations blown up during the Inflation. They are uncorrelated and the distribution of their amplitudes is Gaussian. The growth of their amplitudes has been caused by gravitational instability, and, since the relative density contrast is much smaller than unity, these deviations from isotropy and homogeneity can be described using linear perturbation theory. These perturbations can be visualised in CMB anisotropy map, since these latter are a direct tracer of density fluctuations.

There are three kind of perturbations that evolve independently: scalar, such as the density perturbations, vector, and tensor, which are the gravitational waves. In order to describe the behaviour of an ensemble of interacting particles, one can apply Boltzmann's equation for the phase space distribution function

$$\frac{Df_i}{Dt} = C[f_j], \quad (2.25)$$

where on the left-hand side we have the total time derivative of this distribution function for particle species i , while on the right-hand side we have the effect of all the possible interactions, which can possibly depend on the distribution function of all other species j . Hence, perturbations of matter and radiation can be computed by numerically solving a set of second-order partial differential equations in a certain *gauge*, that is the correspondancy between physical space-time and FRW one. Working in a certain gauge can produce different results from the ones obtained from another gauge. For example, working in a fixed gauge could create some fake modes, called gauge modes, that are non-physical, since they disappear if we move on another gauge. Another choice consists in working with gauge-invariant objects which are linear combinations of perturbations. Interested readers can find a useful reference in Bartolo et al. (2007). Codes like CAMB¹ can solve numerically this set of differential equations, and are used by the majority of the scientific community.

Linear Growth Function

We are mainly interested in studying scalar perturbations, *i.e.* matter density perturbations. We define the density contrast:

$$\delta(\mathbf{x}, z) = \frac{\rho(\mathbf{x}, z)}{\bar{\rho}(z)} - 1, \quad (2.26)$$

where $\bar{\rho}(a) = \rho_0 a^{-3}$ is the average cosmic density. The overdensity always satisfies the inequality $-1 \leq \delta \leq \infty$. Applying relativistic and non-relativistic perturbation theory, we can see that perturbations with $\delta \ll 1$ evolve like a^{n-2} ,

¹<http://camb.info/>

with $n = 4$ before Equivalence and $n = 3$ after that time. In fact, the result of the second-order differential equation generally has decaying and growing solutions, and the growing ones will dominate over time. The key source term ensures them to grow for gravitational instability, identifying the source of perturbations evolution. The first derivative term in perturbation acts as a friction term, which causes fluctuations to grow slower than they would in a static universe. So perturbations do not evolve exponentially, but they rather grow as a power-law. Not all scales are allowed to grow in this way, but only modes larger than a characteristic scale, called Jeans scale, can grow. This Jeans length corresponds to the minimum mass that a structure can have to gravitationally collapse and be formed. This mass is $M_J \propto c_s^3 \bar{\rho}^{-1/2}$, where c_s is the speed of sound. So the Jeans scale λ_J acts as a sound horizon for matter density perturbations. For example before Recombination, photons and baryons are tightly coupled, causing the speed of sound to be close to that of radiation, namely $c_s \simeq c/\sqrt{3}$. In this regime the Jeans mass is way larger than that of any other known object in the Universe, of order $10^{19} M_\odot$. Only CDM can collapse, while baryonic structures can not be formed by gravitational collapse. After Recombination, this mass is of order $10^5 M_\odot$, and recombination radiation pressure is not able to block the formation of cosmologically interesting structures.

It can be proved that during RDE, perturbations grow extremely slowly (logarithmically). During MDE, at redshifts $z \gg z_{\text{eq}}$, where the Universe has $\Omega_m \neq 1$ or $\Omega_\Lambda \neq 0$, density perturbations evolve proportionally to the scale factor, namely $\delta(t) \propto a(t)$. During DE era, which will presumably happen in the future, structures will not grow, allowing for static solutions. In general, when $a \gg a_{\text{eq}}$ with $\Omega_m \neq 1$ or $\Omega_\Lambda \neq 0$, perturbations evolve according to

$$\delta(a) = \delta_0 a \frac{g'(a)}{g'(1)} = \delta_0 a g(a), \quad (2.27)$$

where δ_0 is the density contrast extrapolated to the present epoch and $g(a)$ is the so-called growth function (Carroll et al., 1992; Van Waerbeke & Mellier, 2003) which is a function of redshift (or scale factor), namely

$$g(z) = \frac{5}{2} \Omega_m H(z) \int_0^a \frac{da'}{a'^3 H^3(a')}. \quad (2.28)$$

This represents the growth relative to that in an Einstein-de Sitter Universe. The growth function has to satisfy, for any DE equation of state $\omega(z)$, the dimensionless form of the following second order ordinary differential equation

$$2 \frac{d^2 g(a)}{d \ln a^2} + [5 - 3\omega(a)\Omega_{DE}(a)] \frac{dg(a)}{d \ln a} + 3 [1 - \omega(a)] \Omega(a)g(a) = 0. \quad (2.29)$$

We conclude this section demonstrating how DM can be argued also using the growth of linear perturbations. CMB shows temperature fluctuations of order 10^{-5} that are linked to density fluctuations of the same order of magnitude by Sachs-Wolfe effect. CMB originated at redshift well after the Universe became matter-dominated, where its fluctuations should reach a level of 10^{-2} , and we observe structures with $\delta \gg 1$. This discrepancy is solved by arguing

that CMB displays only baryonic matter component fluctuations, but $\delta(a)$ is made of baryonic and Dark Matter components. We know that this additional matter component interacts only through weak gravitational interactions with the rest of known physics. We already realised that DM component will decouple from cosmic plasma well before baryonic matter, and their fluctuations have all the time to reach the amplitudes observed today. These fluctuations, which will evolve independently in Fourier space, do not collapse because of a growth suppression experienced during the RDE: the expansion scale, dominated by radiation, will be smaller than the typical DM collapse scale. This suppression is restricted to scales that are smaller than the horizon, *i.e.* the size of causally connected regions in the Universe. A scale will enter the horizon if $k = d_h^{-1}(a)$. Large-scales perturbations remain unaffected preventing dark matter fluctuations from collapsing. This suppression factor is $f = (k_0/k)^2$, where $k = 1/\lambda$ is the independent Fourier mode and $k_0 = 1/d_h(a_{\text{eq}})$ is the horizon size scale at Equality. This is roughly $0.083(\Omega_m h^2) \text{ Mpc}^{-1} \simeq 10^{-2} \text{ Mpc}^{-1}$. Remember that the horizon distance is different from the Hubble radius c/H_0 and a scale will enter the Horizon when its length is comparable to the Hubble length $1/H$. So, if the perturbation is smaller than the horizon at Equality, it enters the sound horizon at $z_{\text{enter}} < z_{\text{eq}}$ during RDE. This prevents perturbations from growing further, while perturbations bigger than the horizon at the Equality will keep growing.

The Primordial Power Spectrum of Density Fluctuations

As previously stated, in order to solve some BBT inconsistency and to explain why CMB is so isotropic and presents super-horizon correlations, we need an early period called Inflation. We will not give a full review for it (which can be found in Baumann (2009) for example), but we will only mention its consequences and how these are linked to LSS formation. Inflation is basically a period of accelerated expansion ($\ddot{a} < 0$) driven by a slow-rolling scalar field, in which the early Universe occupied a very small volume. This blown up so largely and so quickly that any inhomogeneities or curvature in such volume are smoothed out, diluting the density of non-relativistic particles. During this period, the horizon scale is much larger than the Hubble length at that time (where H was constant), and small-scale quantum fluctuations are blown up very quickly. Once they are larger than the apparent horizon set by the Hubble length, they are frozen in and behave as completely classical fluctuations. So, when a fluctuation re-enters the Horizon, overdensities begin to collapse.

One of the most powerful predictions coming from Inflation theory, other than the creation of a potentially detectable primordial gravitational waves background, is that primordial density perturbations can be modeled as Gaussian Random Fields (GRF) having scale-invariant power spectrum. This has also been confirmed with CMB measurements, since the temperature anisotropy power spectrum is proportional to the matter density fluctuation one. We give a review of GRF statistics in Appendix A. The Fourier transform of density fluc-

tuations will be

$$\delta(\mathbf{k}) = \int d^3x \delta(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}, \quad (2.30)$$

and their phases are random and these modes are uncorrelated with each other. All the statistical properties of a GRF are described by its variance, *i.e.* its correlation function at the comoving position \mathbf{x} , $\langle \delta(\mathbf{x})\delta(\mathbf{x}') \rangle = \xi(|\mathbf{x} - \mathbf{x}'|)$. The Fourier transform of the correlation function is the Power Spectrum $P(k)$, whose definition is

$$\langle \delta(\mathbf{k})\delta^*(\mathbf{k}') \rangle = (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{k}') P_\delta(k), \quad (2.31)$$

where $k = |\mathbf{k}|$. We can safely apply Fourier decomposition because perturbations scales are much smaller than the curvature radius of the Universe. The scale-invariant power spectrum of primordial density fluctuations is called Harrison-Zel'dovich-Peebles spectrum and scales as $P_\delta(k) \propto k$. Usually this is written modifying it with a scale-dependent transfer function $T(k)$,

$$P_\delta(k, z) = P_\delta(k) g^2(z) T^2(k), \quad (2.32)$$

where $g(z)$ is the growth function. The transfer function encodes modes-growth information when the Universe is radiation or matter dominated. It has to be constant for $k < k_0$ and $T(k) \propto k^{-2}$ for $k > k_0$, with k_0 the Horizon scale at equality. So, growth perturbation modes whose wavelength was small enough to have entered the horizon during RDE could not grow until the Universe became matter dominated. Longer wavelengths, which entered the horizon during MDE, have not been suppressed. Because of this, the power spectrum pic-

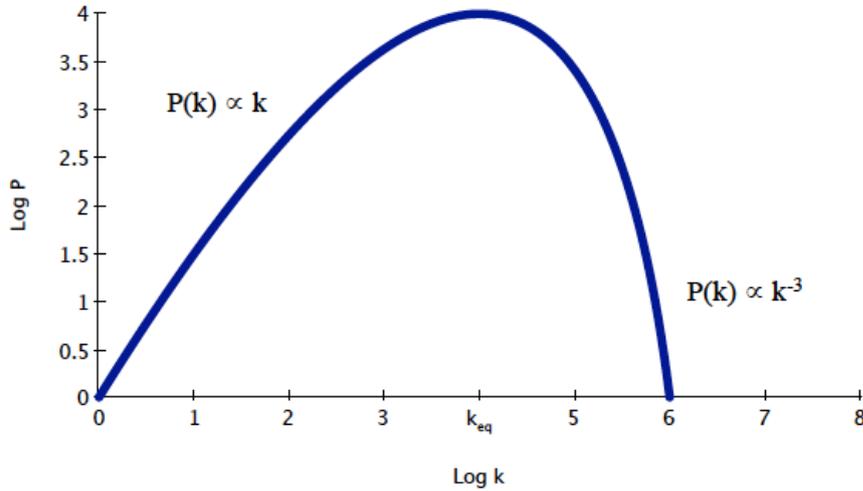


Figure 2.2: The processed power spectrum with scale-invariant Harrison-Zel'dovich-Peebles power spectrum $P_\delta(k) \propto k$.

tured on Figure 2.2, is parametrised as

$$P_\delta(k, z) = \begin{cases} Ak^{n_s}, & (k \ll k_0) \\ Ak^{n_s-4}, & (k \gg k_0) \end{cases}. \quad (2.33)$$

n_s is the spectral index of primordial power spectrum. This is measured, and predicted from slow-roll parameters, to be slightly smaller than 1, namely $n_s =$

0.9655 ± 0.0062 . A is the normalisation of the power spectrum. This scaling can be predicted by making a simple argument. Since the matter density contrast grows as $\delta \propto a^{n-2}$, its power spectrum will scale as $P_\delta(k) \propto a^{2(n-2)}$. When perturbations enter the horizon scale the total power spectrum is scale-invariant. Since $P_{\text{enter}} \propto a_{\text{enter}}^{2(n-2)} P_\delta(k) \propto k^{-4} P_\delta(k)$ for $k \gg k_0$. So, this implies that $P_{\text{enter}}(k) \propto k^{-3}$, while the primordial power spectrum scales as $P_\delta(k) \propto k$. One of the biggest confirmations for the validity of Λ CDM model comes from primordial power spectrum measurement, since this is well approximated by a power spectrum produced by CDM density fluctuation in an universe whose expansion is driven by Dark Energy. We can see the detection of this power spectrum from several techniques in Figure 2.3. We have to take into account that observation will detect smoothed fluctuations up to a certain characteristic cut scale k_s . Such a field will be the convolution of the field with a window function (it can be Gaussian or Top-Hat), and the resulting power spectrum is $P(k, k_s) = |W(k, k_s)|^2 P(k)$.

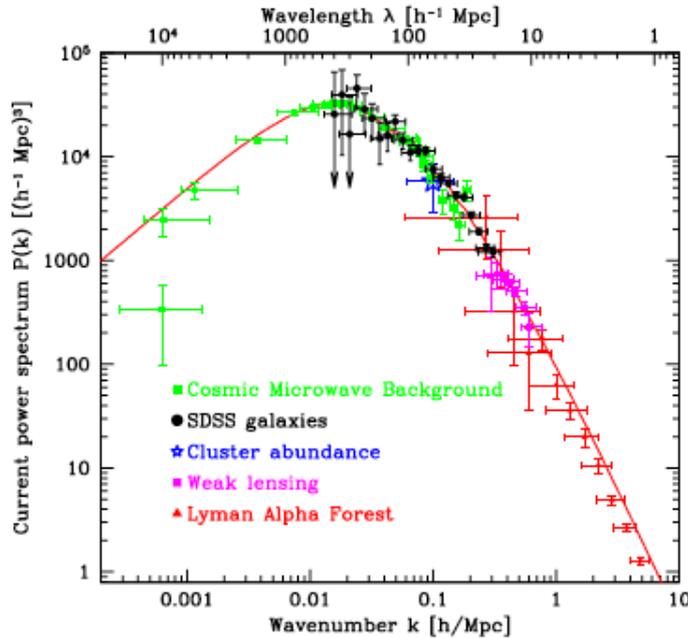


Figure 2.3: The observed power spectrum of density perturbations, as measured from a variety of techniques. Taken from Tegmark & Zaldarriaga (2002).

This spectrum is sensitive to Dark Matter nature: if DM is made of relativistic particles, the produced fluctuation λ needs to have a minimum size in order to keep them gravitationally bound. All the perturbations smaller than this size will be smoothed away by free-streaming of particles, and the power spectrum will show an exponential cut-off at large k . From this we can define the Hot Dark Matter (HDM) as particles that smooth small-scale perturbations, while the Cold Dark Matter (CDM) particles are slow enough to cause no significant damping. Without taking into account exotic particles such as SUSY ones, DM is well modelled by massive and collisionless particles which fell out of equilibrium at very early times, and Weakly Interacting Massive Particles (WIMPs) are among the leading candidates, disfavouring a top-down

HDM scenario driven by relativistic particles like neutrinos. Cosmological observations clearly favor the CDM bottom-up scenario instead, where massive objects form first. Recently there have been attempts trying to implement into simulations an evolution made by Warm DM.

We can note that power spectrum normalisation has to be fixed. It is possible to measure it using several different ways which have different results. We can have it from CMB anisotropies, but the computed normalisation is good only for small scales and CMB is sensitive to the amplitude of scalar and tensor perturbation modes, and not to scalar only, which determine the fluctuation growth. We can measure it from local variance of galaxy counts, since they trace DM fluctuations, but the result will be biased because of its dependence from galaxy formation mechanism and statistical sampling. The normalisation can be also measured from local abundance of galaxy clusters, since the cluster formation will sensitively depend on the amplitude of dark-matter fluctuations. In this case the normalisation can be measured only on scales of order $10h^{-1}$ Mpc. Finally it can be measured from gravitational lensing by large-scale structures, since this detection is sensitive to scales of order $k_0^{-1} \sim 12(\Omega_0 h^2)$ Mpc. Usually this factor is denoted as σ_8 , which is the variance of mass for a sphere with radius $8h^{-1}$ Mpc and with a top hat smoothing window such that

$$\begin{aligned}\sigma^2(R) &= \langle \delta_R^2 \rangle = \int d^3k P(k) W^2(kR) \\ &= \int \Delta^2(k) \left[\frac{3j_1(kR)}{kR} \right]^2 d \ln k,\end{aligned}\tag{2.34}$$

where $\Delta^2(k) = k^3 P(k)/2\pi^2$ is the dimensionless power spectrum usually implemented in numerical codes, and j_1 is the type 1 Bessel function. This is the rms amplitude of mass fluctuations smoothed over a scale R . Planck analysis gives a value $\sigma_8 = 0.829 \pm 0.014$.

There is another effect to consider on primordial power spectrum. At small scales, the growth of density fluctuations begins to depart from linear behaviour, and so fluctuations of different sizes begin to interact. In this case the computation of $P(k, z)$ becomes complicated and numerical methods, such as N-body simulations, are generally required. In dimensionless notation this happens when $\Delta^2 \geq 1$. There is a semi-analytic derivation that works well but it is made under the ansatz that the two-point correlation functions in linear and non-linear regimes are related by a general scaling relation. However these approaches can provide only expectation values of non-linear power spectrum. The most used fitting formula adopted for this non-linear transfer function is the one derived by Eisenstein & Hu (1998).

2.1.6 Poisson's Equation

We conclude this brief cosmological introduction by noting how density perturbations which are smaller than horizon scale and whose velocities are

not relativistic can be described by Newtonian gravity. They will obey to Poisson's equation in proper coordinates

$$\nabla_r^2 \Phi' = 4\pi G \rho_{\text{tot}} = 4\pi G(1 + \delta)\bar{\rho}, \quad (2.35)$$

where ρ_{tot} is the total matter density, and Φ' is the sum of the potentials of the smooth background $\bar{\Phi}$ and the potential of fluctuations Φ . Using Poisson's equation linearity, we can subtract the background contribution and have

$$\nabla_x^2 \Phi = 4\pi G a^2 \bar{\rho} \delta, \quad (2.36)$$

where we introduced the gradient in comoving coordinates $\nabla_x = a\nabla_r$. In matter-dominated era, where $\bar{\rho} = a^{-3}\bar{\rho}_m$, Poisson's equation is

$$\nabla_x^2 \Phi = \frac{3H_0^2}{2a} \Omega_m \delta. \quad (2.37)$$

2.2 Principles of Gravitational Lensing Theory

Here we will introduce the gravitational lensing effect. Gravitational lensing describes how light rays are deflected by gravitational fields. Depending on whether this effect is big or not, *i.e.* so depending on the lensing source and on the lens itself, we can use some approximations to treat it. While weak lensing theory is generally used to describe gravitational lensing by LSS, a more approximated theory can be used to illustrate lensing due to local mass distributions such as clusters and galaxies.

In order to introduce the main features of gravitational lensing, we first describe this theory in a simpler way following (Bartelmann & Schneider, 2001). This comes from the more general description found by perturbing the space-time metric and linearising the propagation of the photon along these geodesics. The deflection of light rays passing through an inhomogeneous medium is a well known phenomenon studied in geometric optics and well illustrated by Fermat's principle. The quantity that determines the change in photon's path is the index of refraction. A typical sketch of a cosmological gravitational lens system is given in Figure 2.4, where a mass concentration, whose gravitational potential plays the same role of the optical refraction index, is placed at a redshift z_L , corresponding to an angular diameter distance \mathcal{D}_L .

We assume that its gravitational field is weak, and that the angles by which light rays are deflected are small. This corresponds to the *Born approximation*, so the actual light ray is approximated as a straight line in the neighbourhood of the lens. This lens deflects the light rays coming from a source redshift z_s , corresponding to an angular diameter distance \mathcal{D}_s . We also consider the *thin lens approximation*: the deflection occurs on scales much smaller than the size of the Universe, and so the thickness of the lens is negligible with respect to the distance between observer, lens, and source. To prove the validity of this statement, let us consider that a typical galaxy cluster has a typical size of a few Mpc, while the comoving distances \mathcal{D} are fractions of the Hubble length.

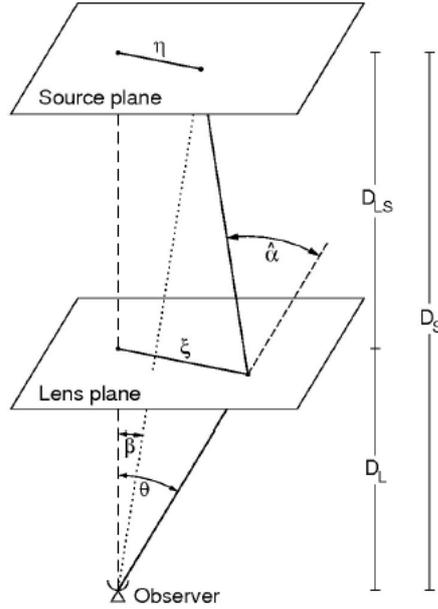


Figure 2.4: Geometry of a gravitational lensing system. Figure from (Bartelmann & Schneider, 2001).

Moreover the lens is static, because the light crossing time is short compared to the dynamical time scale of the lens.

The relation between physical size and the product of the angular size with the distance is not valid if the space-time is curved. Distances are defined in a curved space-time such that this relation holds, but these distances will not be additive, *i.e.* $\mathcal{D}_L + \mathcal{D}_{L_S} \neq \mathcal{D}_S$. Our reference direction will be perpendicular to both lens and source planes, and passing through the observer. Source angular positions on the sky will hence be related to proper physical distances on the plane through $\eta = \beta \mathcal{D}_S$. Due to deflection α , the observer detects the light coming from the source as if it was emitted from the angular distance $\xi = \theta \mathcal{D}_L$, that is the impact parameter on the lens plane.

If the lens is assumed to be a point mass M , the deflection angle produced by it is twice the classical Newtonian value, namely $\alpha = 4GM/(c^2\xi)$. Small deflections like $\alpha \ll 1$ hence correspond to an impact parameter that is $\xi \gg R_S = 2GM/c^2$, the Schwarzschild radius of the lens. In reality we have to deal with a 3D mass distribution with density $\rho(\mathbf{r})$, so that in a volume cell $dm = \rho(\mathbf{r})dV$. For weak deflections the total deflection angle is the sum of the single deflections due to every element dm . Because of Born approximation, the light ray spatial trajectory (ξ, r_3) will not depend on affine parameter, namely on the way we parametrise the light geodesics. This is because the incoming light ray is supposed to propagate along the line of sight direction r_3 which is perpendicular to ξ . So, the impact vector of light ray relative to mass element dm at

position $r = (\xi', r'_3)$ is $\xi - \xi'$, and the total deflection angle is

$$\begin{aligned}\alpha(\xi) &= \frac{4G}{c^2} \sum dm(\xi', r'_3) \frac{\xi - \xi'}{|\xi - \xi'|^2} \\ &= \frac{4G}{c^2} \int d^2\xi' \int dr'_3 \rho(\xi', r'_3) \frac{\xi - \xi'}{|\xi - \xi'|^2} \\ &= \frac{4G}{c^2} \int d^2\xi' \Sigma(\xi') \frac{\xi - \xi'}{|\xi - \xi'|^2},\end{aligned}\quad (2.38)$$

where we have defined the surface mass density

$$\Sigma(\xi') = \int dr'_3 \rho(\xi', r'_3), \quad (2.39)$$

which is the mass density projected onto a plane perpendicular to the incoming light ray. This relation holds as long as the deviation of the light ray from a straight line within the mass distribution is small compared to the scale on which the mass distribution changes significantly. While this is the proper approximation for galaxy and galaxy cluster lensing, this is not true for the LSS case, in which the mass extends all the way from the source to the observer. As already stated, this is only an introductory treatment useful to define lensing observables.

The true position of the lensed object on the sky is related to the apparent one through the *Lens Equation*. From Figure 2.4 we can see that by geometrical construction we have:

$$\eta = \frac{\mathcal{D}_s}{\mathcal{D}_L} \xi - \mathcal{D}_{L_s} \alpha(\xi). \quad (2.40)$$

This can be read as

$$\begin{aligned}\mathcal{D}_s \beta &= \frac{\mathcal{D}_s}{\mathcal{D}_L} \mathcal{D}_L \theta - \mathcal{D}_{L_s} \alpha(\mathcal{D}_L \theta) \\ \beta &= \theta - \frac{\mathcal{D}_{L_s}}{\mathcal{D}_s} \alpha(\mathcal{D}_L \theta) \\ \beta &= \theta - \alpha(\theta).\end{aligned}\quad (2.41)$$

In the last passage we have re-scaled the deflection angle $\alpha(\theta)$.

2.2.1 Lensing Convergence

We can see that Eq.(2.41) has multiple solutions, and so the lens produces multiple images on the sky. The number of these images is always odd and they show up when lensing effect is strong. Usually one can quantify this effect by defining the convergence, or dimensionless surface mass density

$$\kappa(\theta) = \left(\frac{c^2}{4\pi G} \frac{\mathcal{D}_s}{\mathcal{D}_L \mathcal{D}_{L_s}} \right)^{-1} \Sigma(\mathcal{D}\theta) = \frac{\Sigma(\mathcal{D}\theta)}{\Sigma_{cr}}, \quad (2.42)$$

where we have defined the critical surface mass density Σ_{cr} , which is a characteristic value for the surface mass density. In practice, Σ is a projection of the

mass. So, when $\kappa(\boldsymbol{\theta}) \geq 1$ we are in the strong lensing regime. One can write the deflection angle Eq.(2.38) in terms of convergence, namely

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{\mathcal{D}_{L_s}}{\mathcal{D}_s} \boldsymbol{\alpha}(\mathcal{D}_L \boldsymbol{\theta}) = \frac{1}{\pi} \int d^2 \theta' \kappa(\boldsymbol{\theta}') \frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2}. \quad (2.43)$$

This latter equation implies that, as can be proved in geometric optics for the refraction index, the deflection angle is the gradient of the lensing potential

$$\Phi(\boldsymbol{\theta}) = \frac{1}{\pi} \int d^2 \theta' \kappa(\boldsymbol{\theta}') \ln |\boldsymbol{\xi} - \boldsymbol{\xi}'|, \quad (2.44)$$

which is the 2D analogue of the Newtonian gravitational potential of GR. Hence, this potential is related to the convergence via the Poisson equation

$$\kappa(\boldsymbol{\theta}) = \frac{1}{2} \nabla^2 \Phi(\boldsymbol{\theta}) = \frac{1}{2} \nabla \cdot \boldsymbol{\alpha}. \quad (2.45)$$

The lensing potential Eq.(2.44), does not only generate an observed deflection, but also a delay in photons arrival time. This is

$$\tau = \frac{1 + z_L}{H_0} \frac{\mathcal{D}_L \mathcal{D}_s}{\mathcal{D}_{L_s}} \left[\frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \Phi(\boldsymbol{\theta}) \right] \quad (2.46)$$

(Blandford & Narayan, 1985), where the first geometrical term quantifies the increase in the path length due to deflection angle $\boldsymbol{\alpha}(\boldsymbol{\theta}) = \boldsymbol{\theta} - \boldsymbol{\beta}$, while the second term, known as Shapiro delay, decrease it by the photons travelling through the lensing potential. This is another manifestation of Fermat's principle: images are formed at the extrema of the light arrival surface, namely $\nabla \tau = 0$, and this leads to the Lens Equation Eq.(2.41), giving the stationary points where the images are formed. We can note how time delays can provide an interesting estimate on the Hubble parameter H_0 (Kochanek, 2002), once we have accurate measures of time delays and models for the potential. The only limit consists in the uncertainty for the mass distribution. An example of such a measurement can be found in Kochanek & Schechter (2004).

2.2.2 Differential Effects

The lensed images are not only moved, but will present a differential distortion in their shape. This must be derived by solving the Lens Equation for every possible point within an extended source. At this point we remember that gravitational lensing conserves surface brightness, or specific intensity. This is a consequence of Liouville's theorem in absence of emission or absorption. So, the observed surface brightness distribution in the lens plane is related to the one in the source plane via

$$I_{obs}(\boldsymbol{\theta}) = I_s[\boldsymbol{\beta}(\boldsymbol{\theta})] \approx I_s(\mathcal{J}\boldsymbol{\theta}). \quad (2.47)$$

In the last passage we have locally linearised this mapping with a Taylor expansion around the unlensed position, since deflections are small and the angular scale of the source is much smaller than typical lensing scale. In absence

of any rotational component, the distortion of the images is described by the Jacobian matrix $\mathcal{J}(\boldsymbol{\theta})$ which represents how the unlensed position changes respect to the lensed one, namely

$$\mathcal{J}(\boldsymbol{\theta}) = \frac{\partial \beta}{\partial \boldsymbol{\theta}} = \delta_{ij} - \frac{\partial^2 \Phi(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} = \delta_{ij} - \frac{\partial \alpha_i(\boldsymbol{\theta})}{\partial \theta_j}. \quad (2.48)$$

Let us call $\partial^2 \Phi(\boldsymbol{\theta}) / \partial \theta_i \partial \theta_j = \Phi_{,ij}$ the derivatives of the potential respect to the angular positions.

If we split off an isotropic part from the Jacobian, we recover an antisymmetric, trace-free matrix called *shear matrix*, which quantifies the projection of the gravitational tidal field and describes the distortions of background sources. This is

$$\begin{aligned} \left[\mathcal{J} - \frac{1}{2} \text{tr}(\mathcal{J} \cdot \mathcal{I}) \right]_{ij} &= \delta_{ij} - \Phi_{,ij} - \frac{1}{2} (1 - \Phi_{,11} + 1 - \Phi_{,22}) \delta_{ij} \\ &= -\Phi_{,ij} + \frac{1}{2} (\Phi_{,11} + \Phi_{,22}) \delta_{ij} \\ &= \begin{pmatrix} -\frac{1}{2} (\Phi_{,11} - \Phi_{,22}) & -\Phi_{,12} \\ -\Phi_{,12} & \frac{1}{2} (\Phi_{,11} - \Phi_{,22}) \end{pmatrix} \\ &= \begin{pmatrix} -\gamma_1(\boldsymbol{\theta}) & -\gamma_2(\boldsymbol{\theta}) \\ -\gamma_2(\boldsymbol{\theta}) & \gamma_1(\boldsymbol{\theta}) \end{pmatrix}. \end{aligned} \quad (2.49)$$

In the last passage we have defined the components of *shear* tensor, $\gamma = \gamma_1 + i\gamma_2 = |\gamma| e^{2i\phi_\gamma}$. Since it is defined as the trace-free part of the Jacobian matrix \mathcal{J} , it has two independent components. This is not a vector, since it transforms under $e^{2i\phi_\gamma}$ rotations of the coordinate frame. The trace of the Jacobian matrix \mathcal{J} is

$$\begin{aligned} \frac{1}{2} \text{tr}[\mathcal{J}(\boldsymbol{\theta})] &= \left[1 - \frac{1}{2} (\Phi_{,11} + \Phi_{,22}) \right] \delta_{ij} \\ &= \left(1 - \frac{1}{2} \nabla^2 \Phi \right) \delta_{ij} \\ &= (1 - \kappa) \delta_{ij}. \end{aligned} \quad (2.50)$$

So, the Jacobian matrix can be re-written as

$$\mathcal{J}(\boldsymbol{\theta}) = \begin{pmatrix} 1 - \kappa - \gamma_1(\boldsymbol{\theta}) & -\gamma_2(\boldsymbol{\theta}) \\ -\gamma_2(\boldsymbol{\theta}) & 1 - \kappa + \gamma_1(\boldsymbol{\theta}) \end{pmatrix}. \quad (2.51)$$

The rotational component ω should appear in anti-diagonal terms in this latter expression, but this contribution is null at lowest order. So, as shown in Figure 2.5, the distortion induced by convergence is isotropic, being the image re-scaled by a constant factor in every direction. The shear stretches the intrinsic shape of the source along one particular direction. The measurement of the shear is different from convergence one, since it involves galaxies' shapes measurements. Weak lensing changes the ellipticity of a galaxy, but it is quite difficult to distinguish the intrinsic shear from the lensing induced one. But if

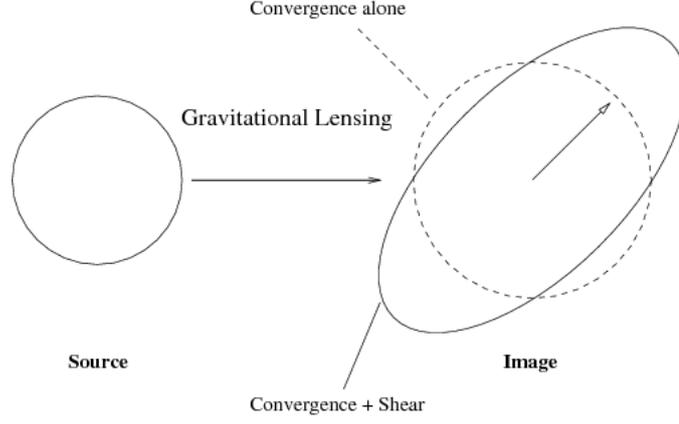


Figure 2.5: Distortion effects due to convergence and shear components on circular source. Figure taken from Narayan & Bartelmann (1996).

we consider that galaxies ellipticities are random, in average this contribution has to be null. Hence a lensing shear measurement comes directly from average ellipticity of background objects. Useful discussions on how to estimate lensing shear through cosmology can be found in Joachimi & Bridle (2010); Bernstein (2009) and various projects, like KIDS or Euclid, have the aim to go further (Cropper et al., 2013; LSST Science Collaboration et al., 2009).

Moreover, since the surface brightness is conserved, a reduction in volume corresponds to an increase in flux. If we define the *magnification* $\mu = |\mathcal{J}|^{-1}$ as the inverse of the determinant of the Jacobian matrix, we get

$$\mu = |\mathcal{J}|^{-1} = \frac{1}{(1 - \kappa)^2 - \gamma^2}. \quad (2.52)$$

The points in which the magnification is ideally infinite are called critical lines. The sources generating images around the critical lines are located along the caustics. In particular the tensor $\mathcal{M} = \mathcal{J}^{-1}$ is called the magnification tensor. Infinite magnification does not occur in reality, mainly because each astrophysical source is extended. Then even point sources would be magnified by a finite value, since a wave-optics description is needed to treat magnification on critical curves. We can note that $\mu \approx 1 + 2\kappa$, and so at lowest order the magnification is determined by the only convergence.

Some schematic models for lens haloes are discussed in Appendix B, whose theoretical properties will be explained. These models will be used to test the validity of the work presented in this thesis.

Since shear and convergence can both be derived from potential's second derivatives, we can link the shear as a convolution of the convergence with a kernel $\chi(\boldsymbol{\theta})$:

$$\gamma(\boldsymbol{\theta}) = \frac{1}{\pi} \int d^2\theta' \chi(\boldsymbol{\theta} - \boldsymbol{\theta}') \kappa(\boldsymbol{\theta}'), \quad (2.53)$$

with convolution kernel given by

$$\chi(\boldsymbol{\theta}) = \frac{\theta_2^2 - \theta_1^2 - 2i\theta_1\theta_2}{\theta^4} \quad (2.54)$$

(Kaiser & Squires, 1993; Squires & Kaiser, 1996). If we invert this expression, we have the resulting mass reconstruction $\hat{\kappa}$ given by

$$\hat{\kappa}(\boldsymbol{\theta}) - \kappa_0 = \frac{1}{\pi} \int d^2\theta' \chi^*(\boldsymbol{\theta} - \boldsymbol{\theta}') \gamma(\boldsymbol{\theta}'), \quad (2.55)$$

where the constant κ_0 shows that the surface density can only be recovered up to a mass sheet. This reflects the fact that a constant κ does not cause a shear: a transformation $\kappa' = \lambda\kappa + (1 - \lambda)$ leaves the reduced shear $g_i = \gamma_i/(1 - \kappa)$ unchanged, where λ is a multiplicative factor such that $\mathcal{J} \rightarrow \lambda\mathcal{J}$. This is the so-called *mass-sheet degeneracy*. Note that κ is a real function, while $\hat{\kappa}$ could have a non negligible imaginary part which could be a sign of systematics in real data.

2.2.3 Second Order Effects

We could use also higher-order terms in Eq.(2.47), the Taylor expansion around the undeflected point. For example, allowing the presence of a second order term, we have for the position

$$\beta_i \simeq \frac{\partial\beta_i}{\partial\theta_j} \theta_j + \frac{1}{2} \frac{\partial^2\beta_i}{\partial\theta_j\partial\theta_k} \theta_j\theta_k = \mathcal{J}_i^j \theta_j + \frac{1}{2} \mathcal{S}_i^{jk} \theta_j\theta_k, \quad (2.56)$$

where $\mathcal{S}_i^{jk} = \mathcal{J}_i^{j,k} = \partial\mathcal{J}_i^j/\partial\theta_k$. These matrices are related to shear derivatives (and so to third derivatives of potential) via

$$\mathcal{S}_i^{j1}(\boldsymbol{\theta}) = \begin{pmatrix} -2\gamma_1^1 - \gamma_2^2 & -\gamma_2^1 \\ -\gamma_2^1 & -\gamma_2^2 \end{pmatrix} \quad (2.57)$$

and

$$\mathcal{S}_i^{j2}(\boldsymbol{\theta}) = \begin{pmatrix} -\gamma_2^1 & -\gamma_2^2 \\ -\gamma_2^2 & 2\gamma_1^2 - \gamma_2^1 \end{pmatrix}. \quad (2.58)$$

The complex quantities which describe second order distortions are called first and second *flexion*. These are defined as

$$F = F_1 + iF_2 = (\gamma_1^1 + \gamma_2^2) + i(\gamma_2^1 - \gamma_1^2) \quad (2.59)$$

$$G = G_1 + iG_2 = (\gamma_1^1 - \gamma_2^2) + i(\gamma_2^1 + \gamma_1^2). \quad (2.60)$$

Flexion is responsible for introducing a curvature and other anisotropic distortions. We can see in Figure 2.6 the effect of these components. A vector \mathbf{F} can hence be formed, with components F_1 and F_2 , so that $\mathbf{F} = \nabla\kappa$. This tells us that the first flexion can be used to obtain the convergence field.

2.3 Lensing in General Relativity

We have just derived a weak lensing description in a heuristic way in order to better understand what are physical properties of a lens system. It is

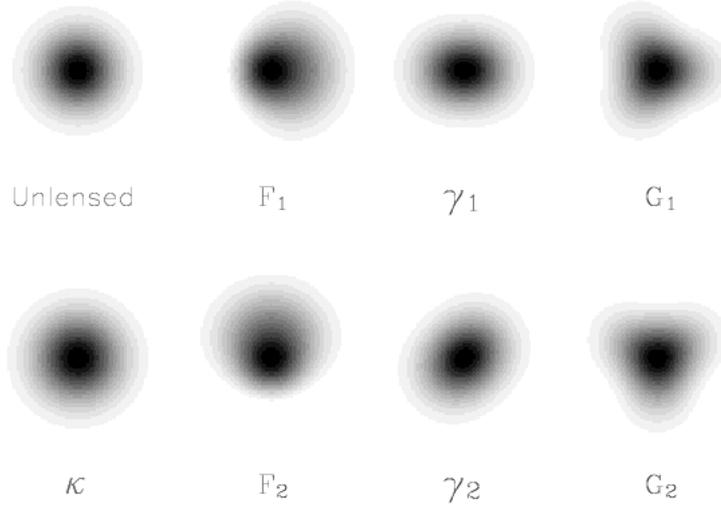


Figure 2.6: Effect of first and second order distortions on a circular source. While shear and convergence effects have been already displayed on Figure 2.5, the first and the second flexion introduce curvature and other distortion. Figure taken from Kneib & Natarajan (2011).

possible to derive gravitational lensing theory in more general terms using GR. The geodesic equation can be used to follow the evolution in the separation of nearby rays, describing gravitational lensing as a distortion of invariant cross-section of infinitesimal light rays bundles. We will treat this topic following (Bartelmann & Schneider, 2001) and (Lewis & Challinor, 2006). This approach deals with physically observable quantities, but global properties of the light-cone, such as the number of images formed, are not readily computable.

2.3.1 Light Ray Bundle Propagation

In an unperturbed space-time, light travels along null geodesic lines, and the transverse separation respect to line-of-sight (LoS) direction can be described by a 2D vector instead of a general 4D vector if we choose an appropriate reference system. We know that the propagation of a thin light bundle in an arbitrary space-time evolves according to the geodesic equation. This equation will depend on the choice made on its affine parameter. Let us consider an observer O with four-velocity U_o^μ satisfying $U_o^\mu U_{o\mu} = 1$ and a photon with measured past-directed dimensionless wavevector $\tilde{k}^\mu = \omega_o k^\mu / c$, which depends on its physical wavevector k^μ . We choose an affine parameter λ passing through O such that: $\lambda = 0$ at observer position; $\lambda > 0$ along the backward light-cone of O ; $U_o^\mu \tilde{k}_\mu = 1$ at O . It follows that $\tilde{k}_\mu = dx/d\lambda$ and λ represents the proper distance along light rays for events close to O .

Let us assume that θ is the infinitesimal separation between $\gamma^\mu(\theta, \lambda)$ light rays and a fiducial ray $\gamma_0^\mu(0, \lambda)$. If this separation is small, we can linearise the problem. Hence, the vector which connects the rays with the fiducial one for the same λ will be $Y^\mu(\theta, \lambda) = \gamma^\mu(\theta, \lambda) - \gamma_0^\mu(0, \lambda) = [\partial\gamma^\mu(\theta, \lambda)/\partial\theta_k] \theta_k$. This vector can

be decomposed in two transverse components $\xi_k(\boldsymbol{\theta}, \lambda)$ (with $k = 1, 2$) respect to LoS direction (and so perpendicular to U_o^μ and \tilde{k}^μ) and one along the beam direction, which is $\xi_0(\boldsymbol{\theta}, \lambda)$. The former components will belong to a plane which is tangential to the sphere of directions seen by the observer. This vector hence is

$$Y^\mu(\boldsymbol{\theta}, \lambda) = \xi_1(\boldsymbol{\theta}, \lambda)E_1^\mu - \xi_2(\boldsymbol{\theta}, \lambda)E_2^\mu - \xi_0(\boldsymbol{\theta}, \lambda)\tilde{k}^\mu, \quad (2.61)$$

where E_k are the orthonormal unit vectors belonging to the plane which is orthogonal to the LoS one, so that $E_k^\mu E_{k'\mu} = -\delta_{kk'}$ and $E_k^\mu \tilde{k}_\mu = E_k^\mu U_{o\mu} = 0$. So the transverse components describe the separation of two light rays at affine parameter λ , while the LoS component allows for deviations along the beam direction. Because of linearisation, ξ depends linearly on $\boldsymbol{\theta}$, so

$$\left. \frac{d\xi}{d\lambda} \right|_{\lambda=0} = \boldsymbol{\theta} \quad \Rightarrow \quad \xi(\lambda) = \mathcal{A}(\lambda)\boldsymbol{\theta}. \quad (2.62)$$

$\mathcal{A}(\lambda)$ is called *Jacobi map* and it is a 2×2 matrix which satisfies the Jacobi differential equation

$$\frac{d^2 \mathcal{A}(\lambda)}{d\lambda^2} = \mathcal{T}(\lambda)\mathcal{A}(\lambda), \quad (2.63)$$

with initial conditions $\mathcal{A}(0) = 0$ and $d\mathcal{A}(0)/d\lambda = I$, where I is the identity matrix. $\mathcal{T}(\lambda)$ is called *tidal matrix* and it is a symmetric matrix which depends on the metric, being a function of Ricci and Weyl tensors. This can be written as

$$\mathcal{T}(\lambda) = \begin{pmatrix} R(\lambda) + \mathcal{R}[F(\lambda)] & \mathcal{I}[F(\lambda)] \\ \mathcal{I}[F(\lambda)] & R(\lambda) - \mathcal{R}[F(\lambda)] \end{pmatrix}, \quad (2.64)$$

where $R(\lambda)$ is the Ricci tensor depending part called source of convergence, and leads to an isotropic focusing of light bundles (so light rays with circular cross-sections will keep having circular cross-sections). $F(\lambda)$ is a complex function called source of shear which causes anisotropic effects by changing the shape of the bundle. In general the tidal matrix describes the influence of space-time curvature on the propagation of light. The affine parameter λ has to be chosen such that it locally reproduces the proper distance and increases with decreasing time. So, the elements of the matrix \mathcal{T} have the dimensions of $[\text{length}]^{-2}$.

2.3.2 Application to a FRW Metric

The tidal matrix Eq.(2.64) in a FRW metric is proportional to the unit matrix I , since in a homogeneous and isotropic universe any source of anisotropy has to vanish. So the tidal matrix symmetric-free part has to be null and $\mathcal{T}(\lambda) = R(\lambda)I$. Thus the solution of Eq.(2.63) has the form $\mathcal{A}(\lambda) = A(\lambda)I$. We need to find the form of this function $A(\lambda)$.

As we already stated $R(\lambda)$ is a function of Ricci tensor, namely

$$R(\lambda) = -\frac{1}{2}R_{\mu\nu}(\lambda)\tilde{k}^\mu(\lambda)\tilde{k}^\nu(\lambda). \quad (2.65)$$

If we replace Ricci tensor with the Einster tensor Eq.(2.2), we are left only with the energy-momentum part of it, since the other two parts depending on scalar R and constant Λ are multiplied by $g_{\mu\nu}$ and \tilde{k}^μ is a null vector. We remember that $k^0 = \omega/c = \omega_o(1+z)/c$ and so $\tilde{k}^0 = -(1+z)$. Hence

$$\begin{aligned} R(\lambda) &= -\frac{4\pi G}{c^2} \left(\rho + \frac{P}{c^2} \right) (1+z)^2 \\ &= -\frac{3}{2} \left(\frac{H_0}{c} \right)^2 \Omega_m (1+z)^5, \end{aligned} \quad (2.66)$$

where in the last passage we have applied Eq.(2.12) for a MDE Universe. We can write the transport Jacobi equation as

$$\frac{d^2 A(\lambda)}{d\lambda^2} = -\frac{3}{2} \left(\frac{H_0}{c} \right)^2 \Omega_m (1+z)^5 A(\lambda), \quad (2.67)$$

From Eq.(2.62) we can note how the Jacobi map, and hence the solution for the Jacobi equation $A(\lambda)$, can be identified with the angular diameter distance Eq.(2.20). To prove this let us note that for $z_1 = 0$ we have an initial-value problem in $\mathcal{D}(z_1, z_2) = a(z_2) f_K[\chi_c(z_1, z_2)] = a(z_2) \chi(z_1, z_2)$ where we have identified $\chi_c(z_1, z_2) = w(z_1, z_2)$. So,

$$\frac{d^2(\mathcal{D}/a)}{d\chi_c^2} = -K \frac{\mathcal{D}}{a}, \quad (2.68)$$

with $\mathcal{D}(0) = 0$ and $d\mathcal{D} = d\chi_c$ for $\chi_c = 0$, because of properties of function $f_K(\chi_c) = \chi$. Now we need a relation between λ and χ_c . Since $x^0 = c(t_0 - t)$ and $dx^\mu = \tilde{k}^\mu d\lambda$, we obtain $d\lambda = -acdt$. But $dt = da/\dot{a}$, so

$$da = -\frac{\dot{a}}{ca} d\lambda \quad \Rightarrow \quad dz = \frac{\dot{a}}{ca^3} d\lambda. \quad (2.69)$$

For null rays $cdt = -ad\chi_c$, $dt = -ad\chi_c/c$. So finally

$$d\lambda = a^2 d\chi_c. \quad (2.70)$$

Applying this result for A in Eq.(2.68), we have

$$\frac{d^2}{d\chi_c^2} \left(\frac{A}{a} \right) = a^2 \frac{d}{d\lambda} \left[a^2 \frac{d}{d\lambda} \left(\frac{A}{a} \right) \right] = a^3 \frac{d^2 A}{d\lambda^2} - a^2 A \frac{d^2 a}{d\lambda^2}. \quad (2.71)$$

Since $da/d\lambda = -\dot{a}/(ac)$, the second derivative of the scale factor respect to the affine parameter is

$$\frac{d^2 a}{d\lambda^2} = \frac{1}{2} \frac{d(da/d\lambda)}{da} = \frac{1}{2c^2} \frac{d}{da} \left(\frac{\dot{a}}{a} \right)^2 = \frac{1}{2c^2} \frac{dH^2}{da}. \quad (2.72)$$

Substituting this together with Eq.(2.67) in Eq.(2.71), we have

$$a^3 \left[-\frac{3}{2} \left(\frac{H_0}{c} \right)^2 \Omega_m (1+z)^5 A(\lambda) \right] - a^2 A(\lambda) \frac{d}{da} \left(H_0^2 \Omega_m a^{-3} \right). \quad (2.73)$$

Here we can see that A satisfies the same equation for \mathcal{D} Eq.(2.68) with same initial conditions. So we can state that $A = \mathcal{D}$.

We can show that we can recover the Lens Equation Eq.(2.41) from such a general case. Let us consider a weak isolated mass inhomogeneity whose dimension is much smaller compared to Hubble radius c/H_0 , like a galaxy or a galaxies cluster. Its potential will be weak, namely $\Phi \ll c^2$, and the metric assumes the form

$$ds^2 = \left(1 + \frac{2\Phi}{c^2}\right)c^2 dt^2 - \left(1 - \frac{2\Phi}{c^2}\right)dx^2. \quad (2.74)$$

Applying this metric to compute the convergence and the shear sources, we have

$$R = -\frac{4\pi G}{c^2}\rho \quad F = -\frac{1}{c^2}(\Phi_{,11} - \Phi_{,22} + 2i\Phi_{,12}). \quad (2.75)$$

In order to apply this perturbation to a homogeneous and isotropic expanding background we Taylor expand the metric around the weak Newtonian potential with peculiar velocities $v \ll c$, namely,

$$ds^2 = a^2(\tau) \left[\left(1 + \frac{2\Phi}{c^2}\right)c^2 d\tau^2 - \left(1 - \frac{2\Phi}{c^2}\right)(d\chi_c^2 + f_K^2(\chi_c)d\Omega^2) \right], \quad (2.76)$$

where $d\tau = dt/a$ is the conformal time element and Φ satisfies Poisson's equation with source $\Delta\rho$, the density perturbation. We need to add a factor $(1+z)^2$ to R and F , since this contribution will come from the presence of the term $\tilde{k}^\mu \tilde{k}^\nu \propto (1+z)^2$. The tidal matrix is

$$\mathcal{T}_{ij}(\lambda) = -\frac{3}{2} \left(\frac{H_0}{c}\right)^2 \Omega_m (1+z)^5 \delta_{ij} - \frac{(1+z)^2}{c^2} (2\Phi_{,ij} + \delta_{ij}\Phi_{,33}), \quad (2.77)$$

where we assumed that light rays are propagating along the LoS direction x_3 . Assuming that the density perturbation $\Delta\rho$ is characterised by its surface mass density and has infinitely thin mass distribution along the LoS direction, its potential will depend on $\Sigma(\xi)$. Inserting this projected mass in the tidal matrix expression just derived, we can compute from Eq.(2.63) $d\mathcal{D}/d\lambda$ across the thin lens plane, whereas the components of \mathcal{D} far from the lens plane are given by a linear combination of solutions of the transport equation Eq.(2.67). We can fix the solution by imposing derivative's continuity at λ_L (corresponding to z_L). The solution will be $\mathcal{D}(\theta, \lambda_s) = \partial\eta/\partial\theta$, and its line integration gives the Lens Equation. For further details consult Seitz et al. (1994).

2.3.3 Weak Cosmological Lensing from LSS

In this work we will have to deal with the lensing signal generated by large scale structures, so we need to extend our previous treatment to a distribution of lenses whose size is comparable to the curvature scale of the Universe. This is commonly referred as *cosmic shear*. Our goal consists in providing a propagation equation for the transverse separation between the light rays of a thin light bundle, which leads to the deflection angle α of weakly deflected

light rays. Moreover we aim to refine lensing observables, showing how the two-point statistics of the convergence can be linked to the power spectrum of matter fluctuations.

The propagation of thin light bundles through an arbitrary space-time is described by the Jacobi equation for geodesic deviations Eq. (2.63). This equation implies that the transverse physical separation ξ between neighbouring rays in a thin light bundle is described by

$$\frac{d^2 \xi}{d\lambda^2} = \mathcal{T} \xi, \quad (2.78)$$

with $d\lambda = -cad t$. Using a FRW metric, the tidal matrix is proportional to the identity matrix and the proportionality factor is Eq.(2.66). We can assert that the closer the tidal matrix is to the unit matrix, the weaker the gravitational lensing effect will be. Since the affine parameter is related to the comoving distance Eq.(2.70), it is convenient to introduce the *comoving separation vector* $x = \xi/a$. Applying these definitions and quantities to the Jacobian equation Eq.(2.78), we have

$$\frac{d^2 x}{d\chi_c^2} + K x = 0, \quad (2.79)$$

where K is the usual spatial curvature term (for negligible relativistic matter)

$$K = \left(\frac{H_0}{c}\right)^2 (\Omega_m - \Omega_\Lambda - 1). \quad (2.80)$$

Eq.(2.79) has the form of an oscillator equation, hence its solutions are trigonometric or hyperbolic functions, depending on whether K is positive or negative. If $K = 0$, we have a flat Universe, and the comoving separation between light rays is a linear function of distance χ_c , namely $x = \chi_c \theta$. Like the tidal matrix, K has the dimensions of $[\text{length}]^{-2}$. Remembering that boundary conditions are $x(\chi_c = 0) = 0$ and $dx/d\chi_c|_{\chi_c=0} = \theta$, we have $x(\chi_c) = f_K(\chi_c)\theta = \chi\theta$, where f_K is the usual function defined in Eq.(2.4).

Perturbing Light Rays

Now we wish to add perturbations to Eq. (2.79). This can be easily done if we consider that lensing masses are typically smaller than the Hubble radius. These inhomogeneities are thus small, and they move with velocities much smaller than c . Moreover, the typical scale over their potential $|\Phi|$ changes is much smaller than the curvature scale of the FRW background model. This means that there is a local neighbourhood around each perturbation which is large enough to contain the perturbation, but is small enough to be considered flat. In other words we can use the perturbed space-time Minkowski metric, and the line element is Eq. (2.74). Fermat's principle demands that the light travel time along actual light paths $x(l)$ is stationary, namely

$$\delta \int_A^B n[x(l)] dl = 0, \quad (2.81)$$

where n is the local refraction index. This can be found by imposing that light propagates at zero eigentime $ds = 0$ and computing the velocity of light in the gravitational field. This latter is

$$c' = \frac{|d\mathbf{x}|}{dt} = c \sqrt{\frac{1 + 2\Phi/c^2}{1 - 2\Phi/c^2}} \approx c \left(1 + \frac{2\Phi}{c^2}\right), \quad (2.82)$$

so that $c' < c$. The index of refraction is

$$n = \frac{c}{c'} = \frac{1}{1 + \frac{2\Phi}{c^2}} \approx 1 - \frac{2\Phi}{c^2}. \quad (2.83)$$

If we write the line element Eq.(2.81) as a function of a curve parameter λ , namely $dl = |d\mathbf{x}/d\lambda|d\lambda$, we will recover the Euler equations

$$\frac{d}{d\lambda} \frac{\partial L}{\partial \dot{\mathbf{x}}} - \frac{\partial L}{\partial \mathbf{x}} = 0, \quad (2.84)$$

where the Lagrangian is $L(\dot{\mathbf{x}}, \mathbf{x}, \lambda) = n[(\lambda)]|d\mathbf{x}/d\lambda|$. With a bit of algebra, one can find that the total deflection angle of the light path is

$$\alpha = \frac{2}{c^2} \int_{\lambda_A}^{\lambda_B} \nabla_{\perp} \Phi d\lambda, \quad (2.85)$$

namely the integral over the perpendicular to LoS direction gradient of gravitational potential. This implies that light rays are locally deflected by

$$\frac{d^2 \mathbf{x}}{d\chi_c^2} = -\frac{2}{c^2} \nabla_{\perp} \Phi, \quad (2.86)$$

describing how an actual light ray is curved away from a straight line in unperturbed Minkowski space. The perpendicular gradient needs to be taken respect to comoving coordinates as well, so

$$\nabla_{\perp} \Phi = \frac{1}{\chi} \nabla_{\theta} \Phi. \quad (2.87)$$

We have to generalise this relation for LSS inhomogeneities that are inside an expanding FRW Universe. All light rays from a source are deflected by the same amount of distortion, so the image of the object can not be distinguished from the original. This means that the absolute deflection of light rays has no physical sense, and we need to introduce deflections relative to a fiducial light ray. For a bundle of rays that enclose this fiducial ray in the center and intersect at the observer, we can denote each ray by its initial angular separation θ from the fiducial ray (for which $\theta = 0$) corresponding to the comoving separation vector $\mathbf{x}(\theta, \chi_c)$. This means that we need to combine the cosmological contribution Eq.(2.79) with this local contribution, such that there will be a dependency on the relative difference of the perpendicular potential gradients between one ray and the fiducial one. So,

$$\frac{d^2 \mathbf{x}}{d\chi_c^2} + K\mathbf{x} = -\frac{2}{c^2} \left\{ \nabla_{\perp} \Phi [\mathbf{x}(\theta, \chi_c), \chi_c] - \nabla_{\perp} \Phi^0(\chi_c) \right\}. \quad (2.88)$$

We note that if we linearise the right-hand side of this latter equation we recover the geodesic deviation equation Eq.(2.78), with the tidal matrix which contains homogeneous cosmological and local perturbations contributions. To solve Eq.(2.88), we construct a Green's function $G(\chi_c, \chi'_c)$ defined on the square $0 \leq \chi_c \leq \chi_c(z_s)$, $0 \leq \chi'_c \leq \chi_c(z_s)$, where $\chi_c(z_s)$ is the comoving distance to the source. This function has to satisfy the following conditions:

- G is continuously differentiable in this square and satisfies the homogeneous differential equation Eq.(2.79);
- G is continuous on the entire square;
- the derivative of G respect to χ_c jumps by 1 on the diagonal boundary between the two triangles made by χ_c and χ'_c ;
- as a function of χ_c , G satisfies the homogeneous boundary conditions at $\chi_c = 0$, namely

$$x = 0 \quad \frac{dx}{d\chi_c} = \theta. \quad (2.89)$$

So,

$$G(\chi_c, \chi'_c) = \begin{cases} A(\chi'_c) \cos \sqrt{K}\chi_c + B(\chi'_c) \sin \sqrt{K}\chi_c, & 0 \leq \chi_c \leq \chi_c(z_s) \\ C(\chi'_c) \cos \sqrt{K}\chi_c + D(\chi'_c) \sin \sqrt{K}\chi_c, & 0 \leq \chi'_c \leq \chi_c(z_s). \end{cases} \quad (2.90)$$

For the homogeneous boundary conditions we find $A = B = 0$, while imposing continuity we have, for basic function and its derivative respectively,

$$\begin{aligned} C \cos \sqrt{K}\chi'_c + D \sin \sqrt{K}\chi'_c &= 0, \\ -C \sin \sqrt{K}\chi'_c + D \cos \sqrt{K}\chi'_c &= \frac{1}{\sqrt{K}}, \end{aligned}$$

from which

$$\begin{aligned} C &= -\frac{1}{\sqrt{K}} \sin \sqrt{K}\chi'_c, \\ D &= \frac{1}{\sqrt{K}} \cos \sqrt{K}\chi'_c. \end{aligned}$$

This implies

$$G(\chi_c, \chi'_c) = \begin{cases} 0, & (\chi_c < \chi'_c) \\ \frac{1}{\sqrt{K}} \sin \sqrt{K}(\chi_c - \chi'_c) & (\chi_c > \chi'_c) \end{cases} = \begin{cases} 0, & (\chi_c < \chi'_c) \\ f_K(\chi_c - \chi'_c) & (\chi_c > \chi'_c) \end{cases}. \quad (2.91)$$

Therefore, the general solution of the propagation equation Eq.(2.88) reads

$$\mathbf{x}(\boldsymbol{\theta}, \chi_c) = f_K(\chi_c) \boldsymbol{\theta} - \frac{2}{c^2} \int_0^{\chi_c} d\chi'_c f_K(\chi_c - \chi'_c) \left\{ \nabla_{\perp} \Phi [\mathbf{x}(\boldsymbol{\theta}, \chi'_c), \chi'_c] - \nabla_{\perp} \Phi^0(\chi'_c) \right\}. \quad (2.92)$$

We can see that the Born approximation is satisfied when the comoving separation between the two light rays due to light deflection is small compared to the comoving separation of two unperturbed rays, namely when

$$\frac{|\mathbf{x}(\boldsymbol{\theta}, \chi_c) - f_K(\chi_c)\boldsymbol{\theta}|}{|f_K(\chi_c)\boldsymbol{\theta}|} \ll 1. \quad (2.93)$$

When this is true, we are allowed to replace $x(\boldsymbol{\theta}, \chi'_c)$ with $f_K(\chi'_c)\boldsymbol{\theta}$ in the integrand. So the Born approximation allows us to replace the difference of the perpendicular potential gradients with the perpendicular gradient of the potential difference. Taking the potential difference then amounts to adding a term to the potential which depends on the comoving distance χ'_c from the observer only. Moreover, we can evaluate the integral in Eq.(2.92) along the unperturbed path.

We can notice that for an unperturbed space-time $\mathbf{x} = \boldsymbol{\theta}f_K(\chi_c)$. So, the relative deflection angle is defined as the deviation between the perturbed and the unperturbed path divided by the angular diameter distance to χ_c ,

$$\boldsymbol{\alpha}(\boldsymbol{\theta}, \chi_c) = \frac{\boldsymbol{\theta}f_K(\chi_c) - \mathbf{x}(\boldsymbol{\theta}, \chi_c)}{f_K(\chi_c)} = \frac{2}{c^2} \int_0^{\chi_c} d\chi'_c \frac{f_K(\chi_c - \chi'_c)}{f_K(\chi_c)} \nabla_{\perp} \Phi [\boldsymbol{\theta}f_K(\chi'_c), \chi'_c]. \quad (2.94)$$

Since this is a displacement angle relative to a fiducial ray, we should rely on its derivatives to consider significant measurable effects, so that the choice of the peculiar ray will not be influential. So Eq.(2.94) is not uniquely determined. In a flat Universe $f_K(\chi_c) = \chi_c$ and $K = 0$, and so

$$\begin{aligned} \boldsymbol{\alpha}(\boldsymbol{\theta}, \chi_c) &= \frac{2}{c^2} \int_0^{\chi_c} d\chi'_c \left(1 - \frac{\chi'_c}{\chi_c}\right) \nabla_{\perp} \Phi(\boldsymbol{\theta}\chi'_c, \chi'_c) = \\ &= \frac{2\chi_c}{c^2} \int_0^1 dy (1 - y) \nabla_{\perp} \Phi(\boldsymbol{\theta}\chi'_c y, \chi'_c y) \end{aligned} \quad (2.95)$$

Let us write the deflection field Eq.(2.94) using the angular gradient

$$\boldsymbol{\alpha}(\boldsymbol{\theta}, \chi_c) = \frac{2}{c^2} \int_0^{\chi_c} d\chi'_c \frac{f_K(\chi_c - \chi'_c)}{f_K(\chi'_c)f_K(\chi_c)} \nabla_{\boldsymbol{\theta}} \Phi [\boldsymbol{\theta}f_K(\chi'_c), \chi'_c], \quad (2.96)$$

from which we can define the lensing potential

$$\phi(\boldsymbol{\theta}) = -\frac{2}{c^2} \int_0^{\chi_c} d\chi'_c \frac{f_K(\chi_c - \chi'_c)}{f_K(\chi'_c)f_K(\chi_c)} \Phi [\boldsymbol{\theta}f_K(\chi'_c), \chi'_c], \quad (2.97)$$

in order to write the deflection field as $\boldsymbol{\alpha}(\boldsymbol{\theta}) = -\nabla_{\boldsymbol{\theta}}\phi(\boldsymbol{\theta})$. This potential appears to be divergent near $\chi_c = 0$, but this singularity affects only the monopole potential, which does not contribute to the deflection field.

Effective Convergence

Analogously to what we have defined in Section 2.2.1, we can define an effective convergence for LSS lenses,

$$\kappa_e(\boldsymbol{\theta}, \chi_c) = \frac{1}{2} \nabla_{\boldsymbol{\theta}} \cdot \boldsymbol{\alpha}(\boldsymbol{\theta}, \chi_c) = \frac{1}{c^2} \int_0^{\chi_c} d\chi'_c \frac{f_K(\chi_c - \chi'_c)f_K(\chi'_c)}{f_K(\chi_c)} \nabla_{\perp}^2 \Phi [\boldsymbol{\theta}f_K(\chi'_c), \chi'_c]. \quad (2.98)$$

∇_{\perp}^2 is the two-dimensional Laplacian operator in comoving coordinates. For thin lens approximation we can replace it with the three-dimensional Laplacian operator, since the contribution along the LoS direction average to zero in the limit to which we are working. Explicitly,

$$\nabla_{\theta}^2 = \frac{\partial^2}{\partial \theta_1^2} + \frac{\partial^2}{\partial \theta_2^2} = f_K^2(\chi_c) \left(\frac{\partial^2}{\partial \xi_1^2} + \frac{\partial^2}{\partial \xi_2^2} \right) = f_K^2(\chi_c) \nabla_{\perp}^2 = f_K^2(\chi_c) \left(\nabla^2 - \frac{\partial^2}{\partial z^2} \right). \quad (2.99)$$

So, when we apply this operator to our lensing potential we have

$$\nabla^2 \Phi = \frac{1}{f_K^2(\chi_c)} \nabla_{\theta}^2 \Phi + \frac{\partial^2 \Phi}{\partial z^2}. \quad (2.100)$$

We can see how inserting this latter expression in Eq.(2.98) the LoS derivative is not influential, because the lens is gravitationally bound and $\partial \Phi / \partial z = 0$ at the boundaries. Thus

$$\kappa_e(\boldsymbol{\theta}, \chi_c) = \frac{1}{c^2} \int_0^{\chi_c} d\chi'_c \frac{f_K(\chi_c - \chi'_c) f_K(\chi'_c)}{f_K(\chi_c)} \nabla^2 \Phi [\boldsymbol{\theta} f_K(\chi'_c), \chi'_c] \quad (2.101)$$

and we are formally allowed to use Poisson's equation Eq.(2.37), obtaining

$$\kappa_e(\boldsymbol{\theta}, \chi_c) = \frac{3\Omega_m H_0^2}{2c^2} \int_0^{\chi_c} d\chi'_c \frac{f_K(\chi_c - \chi'_c) f_K(\chi'_c)}{f_K(\chi_c)} \frac{\delta [\boldsymbol{\theta} f_K(\chi'_c), \chi'_c]}{a}. \quad (2.102)$$

The Born approximation ($\boldsymbol{x} = \boldsymbol{\theta} f_K(\chi_c)$) allows us to neglect higher-order terms in potential derivatives in $\nabla_{\perp}^2 \Phi$ argument. These terms are usually determined by products of the density perturbations averaged along the LoS. Even if individual δ may be large, their average is really small. So $\kappa_e \ll 1$ and such terms are accordingly negligible.

If the sources are distributed in redshift or, equivalently, in coordinate distance $f_K(\chi_c)$, the mean effective convergence has to be averaged over the normalised source-distance distribution $G[f_K(\chi_c)]$, namely

$$\bar{\kappa}_e(\boldsymbol{\theta}) = \int_0^{\chi_c(z_H)} d\chi_c G[f_K(\chi_c)] \kappa_e(\boldsymbol{\theta}, \chi_c), \quad (2.103)$$

where $G[f_K(\chi_c)] d\chi_c = p_z(z) dz$ is the probability to find a source within $d\chi_c$. z_H is the horizon distance, defined as the comoving distance obtained for infinite redshift. So the source-weighted effective convergence is

$$\bar{\kappa}_e(\boldsymbol{\theta}) = \frac{3\Omega_m H_0^2}{2c^2} \int_0^{\chi_c(z_H)} d\chi'_c W(\chi_c) f_K(\chi_c) \frac{\delta [\boldsymbol{\theta} f_K(\chi_c), \chi_c]}{a}, \quad (2.104)$$

where the effective weight function is

$$W(\chi_c) = \int_{\chi_c}^{\chi_c(z_H)} d\chi'_c G(\chi'_c) \frac{f_K(\chi'_c - \chi_c)}{f_K(\chi'_c)}. \quad (2.105)$$

It is possible to see that Eq.(2.104) could be also recovered from generalising the definition of the surface mass density to a three-dimensional matter distribution. For example, in redshift coordinates we have

$$\kappa_e = \frac{4\pi G}{c^2} \int_0^{z_s} dz \frac{\mathcal{D}_L \mathcal{D}_{L_s}}{\mathcal{D}_s} \frac{dD_p}{dz} \Delta\rho, \quad (2.106)$$

which is analogous to a generalised three-dimensional version of Eq.(2.42), with the function W being $\langle \mathcal{D}_{L_s} / \mathcal{D}_s \rangle$ averaged over source distances at fixed lens distance.

Recovering Differential Effects

The Jacobian matrix, which describes the mapping between the source plane and the image plane is defined in analogy with the one already defined in the general lensing theory Eq.(2.48), namely

$$\mathcal{J}(\boldsymbol{\theta}, \chi_c) = I - \frac{\partial \boldsymbol{\alpha}(\boldsymbol{\theta}, \chi_c)}{\partial \boldsymbol{\theta}} = \frac{1}{f_K(\chi_c)} \frac{\partial \mathbf{x}(\boldsymbol{\theta}, \chi_c)}{\partial \boldsymbol{\theta}}. \quad (2.107)$$

This is explicitly

$$\mathcal{J}_{ij}(\boldsymbol{\theta}, \chi_c) = \delta_{ij} - \frac{2}{c^2} \int_0^{\chi_c} d\chi'_c \frac{f_K(\chi_c - \chi'_c) f_K(\chi'_c)}{f_K(\chi_c)} \Phi_{,ik} [\mathbf{x}(\boldsymbol{\theta}, \chi'_c), \chi'_c] \mathcal{J}_{kj}(\boldsymbol{\theta}, \chi'_c), \quad (2.108)$$

where i and j refer to the two perpendicular components on the transverse plan, and $\Phi_{,ik}$ is the second-order derivative of the potential with respect to the transverse coordinates. Here the integral is made along the actual perturbed photon path. Since we are dealing with weak gravitational fields, \mathcal{J} can be expanded into powers of Φ . This means that, to first order in perturbations, we have

$$\mathcal{J}_{ij}(\boldsymbol{\theta}, \chi_c) \simeq \delta_{ij} - \frac{2}{c^2} \int_0^{\chi_c} d\chi'_c \frac{f_K(\chi_c - \chi'_c) f_K(\chi'_c)}{f_K(\chi_c)} \Phi_{,ij} [\boldsymbol{\theta}, \chi'_c, \chi'_c], \quad (2.109)$$

(Schäfer et al., 2012). Here we have applied Born approximation, and, because of it, the integral is now computed along the unperturbed path. Moreover, the Jacobian matrix remains symmetric even in the case of weak lensing². From the Jacobian Eq.(2.109), we can define all the analogous quantities already defined in Section 2.2.2. For example the magnification is

$$\mu(\boldsymbol{\theta}, \chi_c) = \frac{1}{|\mathcal{J}(\boldsymbol{\theta}, \chi_c)|} \approx 1 + \nabla_{\boldsymbol{\theta}} \cdot \boldsymbol{\alpha} = 1 + 2\kappa_e, \quad (2.110)$$

so the magnification is determined mainly by the convergence, at lowest order in weak lensing approximation. As in the single lens-plane situation, the anisotropic deformation, *i.e.* the shear, is determined by the trace-free part of the matrix \mathcal{J} .

²This is not true for multiple lens-plane simulations (Schneider et al., 1992).

2.4 Convergence Power Spectrum from Limber's Equation

The approximation made in Eq.(2.101) allows us to use Limber's Equation (Limber, 1954) derived in Appendix A.5 to link the angular two-dimensional convergence power spectrum to the LoS integral of full three-dimensional matter power spectrum. This important relation, first derived in Fourier space by (Kaiser, 1992), defines the main observable used in this work. So, if we identify

$$q_1(\chi_c) = q_2(\chi_c) = \frac{3H_0^2\Omega_m}{2c^2} \frac{f_K(\chi_c)}{a(\chi_c)} W(\chi_c), \quad (2.111)$$

the result we get by applying Eq.(A.37) is

$$C_L^{KK} = \frac{9H_0^4\Omega_m^2}{4c^4} \int_0^{\chi_c(z_H)} d\chi_c \frac{W^2(\chi_c)}{a^2(\chi_c)} P_\delta\left(\frac{L}{f_K(\chi_c)}, \chi_c\right), \quad (2.112)$$

where $L = |L|$. The convergence power spectrum is sensitive to cosmological parameters and to non-linear evolution of matter power spectrum.

At large scales ≤ 30 arcmins we can assert that $\delta \propto ag(a)$. Therefore, the convergence power spectrum becomes

$$C_L^{KK} = \frac{9H_0^4\Omega_m^2}{4c^4} \int_0^{\chi_c(z_H)} d\chi_c W^2(\chi_c) g^2[a(\chi_c)] P_\delta^{\text{lin}}\left(\frac{L}{f_K(\chi_c)}, \chi_c\right), \quad (2.113)$$

where P_δ^{lin} is the linear matter contrast power spectrum extrapolated to the present epoch. For an Einstein-de Sitter matter dominated model, $P_\delta \propto a^2$, and $g(a) = 1$. So we have

$$C_L^{KK} = \frac{9H_0^4\Omega_m^2}{4c^4} \int_0^{\chi_c(z_H)} d\chi_c W^2(\chi_c) P_\delta^{\text{lin}}\left(\frac{L}{f_K(\chi_c)}, \chi_c\right), \quad (2.114)$$

where the weight function is

$$W(\chi_c) = \int_{\chi_c}^{\chi_c(z_H)} d\chi'_c G(\chi'_c) \left(1 - \frac{\chi_c}{\chi'_c}\right). \quad (2.115)$$

If the redshift distribution of the sources can be approximated by a delta function, *i.e.* if the angular size distance within the observed band is small compared to the angular size distance from the observer (at $z = 0$) to the center of the band, we can safely assume that the matter within the emission region is not contributing significantly to the lensing. So $G(\chi'_c) = \delta^D(\chi'_c - \chi'_c(z_s))$. This is applied to the CMB case, for example, and we will make use of this approximation in our work for the 21 cm as well. This means that W becomes an Heaviside function, namely

$$W(\chi_c) = \left[1 - \frac{\chi_c}{\chi_c(z_s)}\right] H[\chi_c(z_s) - \chi_c] = \begin{cases} 0, & \text{for } \chi_c > \chi_c(z_s) \\ \left[1 - \frac{\chi_c}{\chi_c(z_s)}\right] & \text{for } \chi_c < \chi_c(z_s) \end{cases} \quad (2.116)$$

Therefore, the convergence power spectrum reads

$$C_L^{\kappa\kappa} = \frac{9H_0^4\Omega_m^2}{4c^4} \int_0^{\chi_c(z_s)} d\chi_c \left[1 - \frac{\chi_c}{\chi_c(z_s)} \right]^2 \chi_c(z_s) P_\delta^{\text{lin}} \left(\frac{L}{f_K(\chi_c)}, \chi_c \right), \quad (2.117)$$

We can write this integral in redshift variables, namely

$$C_L^{\kappa\kappa} = \frac{9H_0^3\Omega_m^2}{4c^3} \int_0^{\chi_c(z_s)} dz \frac{W^2(z)}{a^2(z)E(z)} P_\delta^{\text{lin}} \left[\frac{L}{\chi(z)}, z \right], \quad (2.118)$$

where $E(z) = H(z)/H_0$ is defined in Eq.(2.13), and $W(z)$ is defined in Eq.(2.116).

We can write the analogous expression for the deflection field or lensing potential power spectra by considering the relations among these quantities in Fourier space. If $L = 2\pi/\theta$ is the Fourier dual of the angular coordinate θ , we get

$$\kappa(\theta) = \frac{\nabla_\theta^2 \phi(\theta)}{2} \Rightarrow \tilde{\kappa}(L) = -\frac{L^2}{2} \tilde{\phi}(L), \quad (2.119)$$

$$\alpha(\theta) = \nabla_\theta \phi(\theta) \Rightarrow \tilde{\alpha}(L) = iL \tilde{\phi}(L). \quad (2.120)$$

Therefore, the relation among the three power spectra is

$$C_L^{\kappa\kappa} = \frac{L^2}{4} C_L^{\alpha\alpha} = \frac{L^4}{4} C_L^{\phi\phi}, \quad (2.121)$$

and of course $C_L^{\alpha\alpha} = L^2 C_L^{\phi\phi}$. Because of this, we can write the deflection field power spectrum as

$$C_L^{\alpha\alpha} = \frac{9H_0^3\Omega_m^2}{L(L+1)c^3} \int_0^{z_s} dz \frac{W^2(z)}{a^2(z)E(z)} P_\delta^{\text{lin}} \left[\frac{L}{\chi(z)}, z \right]. \quad (2.122)$$

From Eq.(2.118) all density perturbation modes whose wavenumber are larger than $k_{\text{min}} = l/\chi(z_s)$, corresponde to a scale $\chi(z_s)\theta_{\text{Max}}$, contributing to $C_L^{\kappa\kappa}$. Consider that density perturbations on scales smaller than a few Mpc become non-linear at moderate z , breaking down the linear evolution assumption.

In order to see how much power is added at small scales by non linearities we can compare Eq. (2.122), computing using the computational linear power spectrum by Eisenstein & Hu (1998), with its analogous computed using the non-linear Λ CDM matter power spectrum as implemented by a code using the computation of Peacock & Dodds (1996), both for a Λ CDM universe. We can see in Figure 2.7 the convergence power spectrum at various redshifts, from $z = 0.5$ to $z = 100$, using a linear (dashed line) and a non-linear (solid line) evolving matter power spectrum. Depending on redshift, non-linear effects are important on scales of about 10 arcmins, where power spectrum amplitude is increased by more than an order of magnitude. Density fluctuations on angular scales smaller than 10 arcmins contribute more strongly to weak lensing by LSS.

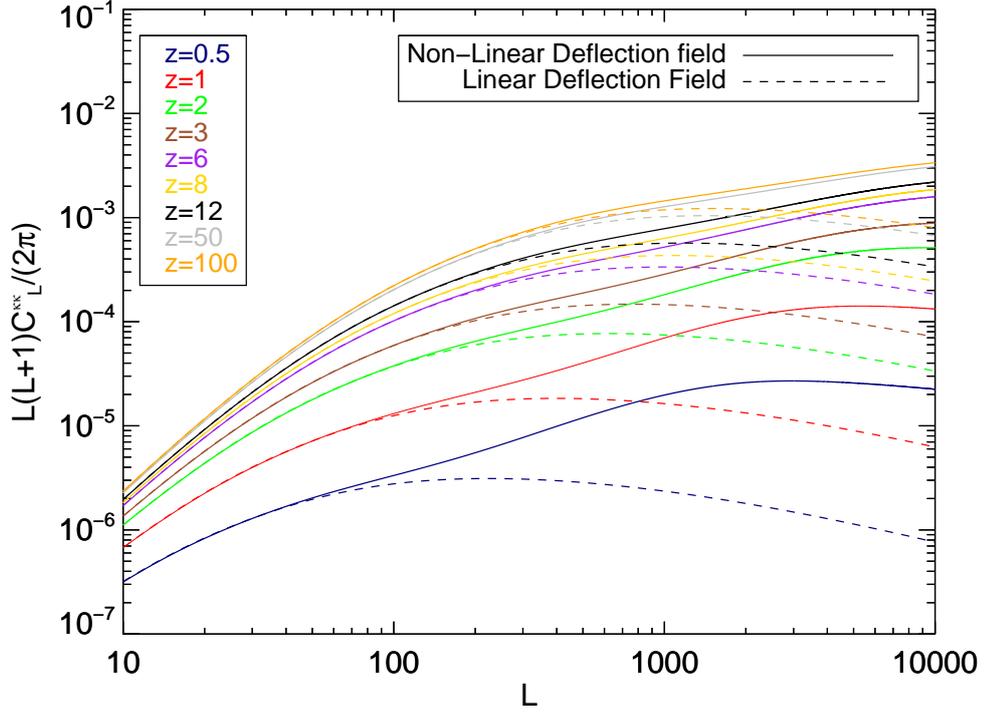


Figure 2.7: The convergence power spectrum as a function of $L = kf_K(\chi_c)$ computed for several redshifts, from $z = 0.5$ to $z = 100$. The dashed lines have been computed using a linear matter power spectrum, while solid lines have been computed using a non-linear matter power spectrum.

2.4.1 Shear Power Spectrum

It is also possible to find expressions for two-point statistics of other lensing observables, like shear. We recall that the shear is defined as

$$\gamma = \gamma_1 + i\gamma_2 = \frac{1}{2} \left(\frac{\partial^2 \phi}{\partial \theta_1^2} - \frac{\partial^2 \phi}{\partial \theta_2^2} \right) + i \frac{\partial^2 \phi}{\partial \theta_1 \partial \theta_2}, \quad (2.123)$$

so its Fourier transform is

$$\gamma(L) = -\frac{L_1^2 - L_2^2 + 2iL_1L_2}{2} \phi(L). \quad (2.124)$$

Now, if we apply Eq.(2.119), we get

$$\gamma(L) = -\frac{L_1^2 - L_2^2 + 2iL_1L_2}{L^2} \kappa(L) = e^{2i\varphi} \kappa(L). \quad (2.125)$$

We note that $e^{2i\varphi}$ has unitary modulus. This means that

$$\langle \gamma(L) \gamma^*(L') \rangle = \langle \kappa(L) \kappa^*(L') \rangle = (2\pi)^2 \delta^D(L - L') C_L^{\kappa\kappa}, \quad (2.126)$$

and so, in flat sky approximation, shear and convergence power spectra are the same quantity. Moreover we can only measure the reduced shear $g = \gamma/(1 - \kappa)$. The relation among observables and the underlying convergence field is more

complicated, but simulations can provide the cosmological corrections for it (Hilbert et al., 2009).

As already puntualised in Hoekstra (2013), shear cannot be measured reliably at every angular position, since the actual survey geometry could be really complicated, because of reflections in the optics or saturation due to brightest stars. Usually to avoid the problem one uses the ellipticity correlation function, which is a linear combination of shear's tangential and cross components respect to the line connecting a pair of galaxies in polar coordinates with polar angle α . These latter are defined as

$$\gamma_t = \gamma (\cos^2 \alpha - \sin^2 \alpha) = \gamma \cos 2\alpha \quad \gamma_\times = \gamma \sin 2\alpha. \quad (2.127)$$

Fourier transforming these definitions, one finds

$$\langle \gamma_t(\theta) \gamma'_t(\theta) \rangle = \frac{1}{2} \int \frac{LdL}{2\pi} C_L^{\kappa\kappa} [J_0(L\theta) + J_4(L\theta)], \quad (2.128)$$

$$\langle \gamma_\times(\theta) \gamma'_\times(\theta) \rangle = \frac{1}{2} \int \frac{LdL}{2\pi} C_L^{\kappa\kappa} [J_0(L\theta) - J_4(L\theta)]. \quad (2.129)$$

If we define $\varepsilon_\pm(\theta) = \langle \gamma_t(\theta) \gamma'_t(\theta) \rangle \pm \langle \gamma_\times(\theta) \gamma'_\times(\theta) \rangle$, we hence have

$$\varepsilon_{+/-}(\theta) = \int \frac{LdL}{2\pi} C_L^{\kappa\kappa} J_{0/4}(L\theta), \quad (2.130)$$

where J_i is the i -th order Bessel's function of the first kind. Since $\langle \varepsilon \rangle = 0$, any deviation from this result is a clue for systematic errors. For further discussions in real measurements of cosmic shear we advice the reader to consult Kilbinger et al. (2013); Schneider et al. (2002).

Chapter 3

21 cm Radiation Cosmology

The diffuse 21 cm radiation produced by neutral hydrogen atom in the intergalactic medium (IGM) can unveil a great wealth of cosmological and astrophysical information across a large redshift range, as shown in Section 1.1. It can potentially allow us to investigate epochs such as the Dark Ages, the Cosmic Dawn, the Epoch of Reionization (EoR), as well as the late Universe, mapping most of our horizon volume and providing accurate cosmological information (Mao et al., 2008).

The 21 cm line is produced by hyperfine splitting of the 1S ground state caused by the interaction between electron and proton magnetic moments. This splitting separates two energy states by $\Delta E = 5.9 \times 10^{-6}$ eV, corresponding to a wavelength of $\lambda_{21} = 21.1061$ cm or to a frequency of $\nu_{21} = 1420.4057$ MHz. Since almost the totality of baryonic matter in the Universe is made by neutral hydrogen, this is a tracer of IGM and of local galaxies properties. Historically, the 21 cm line was predicted by van de Hulst in 1942, and it has been first detected in emission by Ewen and Purcell in 1951 (Ewen & Purcell, 1951). Given the accuracy of this detection, the 21 cm line has been used in determining the velocity distribution of gas within our galaxy and others in the local Universe. These measurements are often used to trace galactic dynamics. The 21 cm line can also be seen in absorption against radio-loud background from individual sources at low redshifts.

Moreover, 21 cm can provide us information about fundamental physics and cosmology: part of the signal traces the density field giving information about neutrino masses and initial conditions from early epoch of inflation through the power spectrum. 21 cm surveys can use tomographic information to probe the large scale structure in three dimensions (across the sky and along time) and, assuming the various issues with foregrounds are dealt with, 21 cm radiation can be a powerful cosmological probe, complementing and competing with state-of-the-art CMB and optical galaxy surveys. As seen in Section 1.1, very sensitive and large radio telescopes like the SKA and its pathfinders can perform HI clustering and lensing measurements with very good precision and give us the first data from unexplored cosmological epochs.

In this chapter we will review the basic atomic physics of the 21 cm line, highlighting how it changes from epoch to epoch, and describing its fluctuations, which provide an observable power spectrum. We will also briefly dis-

cuss how 21 cm observations can be used to constrain cosmological parameters.

There is a number of extensive reviews on this subject. We will mainly follow Furlanetto et al. (2006), Pritchard & Loeb (2012), and Morales & Wyithe (2010) through this whole chapter.

3.1 Physics of the 21 cm Line of Atomic Hydrogen

The absorption α_ν and the emission e_ν of the hydrogen atom are described, in absence of scattering, by the radiative transfer equation through gas along the line of sight

$$\frac{dI_\nu}{dx} = -\alpha_\nu I_\nu + e_\nu, \quad (3.1)$$

which determines the change in brightness I_ν along a proper path x at a frequency ν . The brightness expresses the energy carried by rays per unit area, frequency, solid angle, and time. Sometimes this is expressed as angle-averaged $J_\nu = \int I_\nu d\Omega$. The flux produced by an individual gas cloud which subtends a solid angle $d\Omega$ is hence $S = \int I_\nu \cos\theta d\Omega d\nu$. This is measured in Janskys¹. Since solid angle elements are usually small with small apparent angular diameter, we can write $S_\nu \approx I_\nu \Delta\Omega = 2k_B T_b \nu^2 \Delta\Omega / c^2$, where all these quantities are measured in the observer's frame. Since $d\nu \propto d_L^{-2}$, the flux scales as $S_\nu \propto (1+z)d_L^{-2}$.

I_ν is quantified by the equivalent brightness temperature $T_b(\nu)$ defined as the temperature of a blackbody radiator such that its spectrum is $B_\nu = I_\nu(T_b)$. The 21 cm physics is well approximated by the low-frequency regime defined by the Rayleigh-Jeans limit: the relation between brightness and temperature is

$$T_b(\nu) \approx \frac{I_\nu c^2}{2k_B \nu^2}, \quad (3.2)$$

with k_B the Boltmann's constant. The 21 cm will be observed at an apparent frequency $T_b = T_b^0(\nu_{21})/(1+z)$ due to Doppler effect, since $\nu_o = \nu_{21}/(1+z)$. Similarly, the brightness temperature of the CMB in a comoving frame at redshift z scales from the observed value of $T_{\text{CMB}}(0) = 2.73$ K to $T(z) = 2.73(1+z)$ K.

We can rewrite Eq.(3.1) to give the radiative transfer from light from a background radio source of brightness temperature T_R along the LoS through a cloud of optical depth $\tau_\nu = \int dx \alpha_\nu(x)$ and uniform excitation temperature T_{ex} so that the observed temperature T_b^{obs} at a frequency ν is

$$T_b^{\text{obs}}(\nu) = T_{\text{ex}} (1 - e^{-\tau_\nu}) + T_R(\nu) e^{-\tau_\nu}. \quad (3.3)$$

Usually, the excitation temperature T_{ex} coincides with the spin temperature T_S , which quantifies the relative number densities n_i of atoms in the two hyperfine levels of the electronic ground state². If the singlet level is labeled by

¹ 1 Jy = 10^{-23} erg s⁻¹ cm⁻² Hz⁻¹.

²The assumption of a single T_S applying to the entire Eq.(3.3) is not strictly correct. Rigorously one should solve a Boltmann equation coupling spin and velocity distributions. For long collision time, this introduces percent level changes to $T_b(\nu)$

the subscript 0 and the triplet one is labeled by 1, we can write

$$\frac{n_1}{n_0} = \frac{g_1}{g_0} e^{-T_\star/T_S}, \quad (3.4)$$

with $g_1/g_0 = 3$ the ratio of statistical degeneracy factors of the two levels, and $T_\star = hc/\lambda_{21}k_B = E_{10}/k_B = 0.068$ K the equivalent temperature to the energy splitting $E_{10} = \Delta E$. For example if $T_S \gg T_\star$, three of four atoms are in the excited state.

Assuming stimulated emission, the optical depth of a cloud of hydrogen is then

$$\tau_\nu = \int dx \alpha_\nu(x) = \int dx \left[1 - e^{-E_{10}/k_B T_S}\right] \sigma_0 \phi(\nu) n_0, \quad (3.5)$$

where $n_0 = n_{\text{HI}}/4$ with n_{HI} the neutral hydrogen column density, and the 21 cm cross-section is defined as

$$\sigma(\nu) = \sigma_0 \phi(\nu) = \frac{3c^2 A_{10}}{8\pi\nu^2} \phi(\nu), \quad (3.6)$$

where $A_{10} = 2.85 \times 10^{-15} \text{s}^{-1}$ is the spontaneous decay rate of spin-flip transition, and the line profile is normalized so that $\int \phi(\nu) d\nu = 1$. To evaluate Eq.(3.5) we need to find the column length as a function of frequency $x(\nu)$, in order to determine $\phi(\nu)$. This can be done in two ways as already pictured by Pritchard & Loeb (2012):

- the IGM expands uniformly with the Hubble flow, and so

$$dx = -\frac{cH(z)}{1+z} dz, \quad (3.7)$$

where the redshifting of light is related to observed and emitted frequencies;

- assuming the Sobolev approximation, *i.e.* a linear velocity profile

$$v = \frac{dv}{dx} dx, \quad (3.8)$$

and using the Doppler law $v_{\text{obs}} = v_{\text{em}} (1 - v/c)$ self-consistently to $O(v/c)$. This approach describes the Hubble law in the absence of peculiar velocities, so we need to include this contribution.

Both approaches lead to the same result. We can write $\phi(\nu) = c/[\Delta v \nu] \approx c/[xH(z)\nu]$, where $\Delta v \approx xH(z)$ is the velocity broadening of a region of linear dimension x . The column density along x will depend on the neutral fraction x_{HI} of hydrogen, so $n_{\text{HI}} = x_{\text{HI}} n_{\text{HI}}(z)x$. Assuming uniformity throughout the cloud and small optical depth, we can write Eq.(3.5) as

$$\begin{aligned} \tau_\nu &\approx \frac{E_{10}}{k_B T_S} \frac{n_{\text{HI}}}{4} \sigma_0 \phi(\nu) \\ &= \frac{3}{32\pi} \frac{hc^3 A_{10}}{k_B T_S \nu_{21}^2} \frac{x_{\text{HI}} n_{\text{HI}}(z)}{(1+z) (dv_{\parallel}/dr_{\parallel})} \\ &\approx 9.2 \times 10^{-3} (1 + \delta_b) (1+z)^{3/2} \frac{x_{\text{HI}}}{T_S} \left[\frac{H(z)/(1+z)}{dv_{\parallel}/dr_{\parallel}} \right]. \end{aligned} \quad (3.9)$$

Here δ_b is the fractional baryon overdensity and $dv_{\parallel}/dr_{\parallel}$ is the gradient of the proper velocity along the LoS, including both Hubble expansion and peculiar velocities.

We can use Eq.(3.9) to compute the contrast between high-redshifted HI clouds and a background radiation. This contrast is defined as brightness temperature fluctuation, namely

$$\begin{aligned}\delta T_b &= \frac{T_S - T_R(z)}{1+z} (1 - e^{-\tau_\nu}) \\ &\approx \frac{T_S - T_R(z)}{1+z} \tau_\nu \\ &\approx 9 x_{\text{HI}} (1 + \delta_b) (1+z)^{1/2} \left[\frac{T_S - T_R(z)}{T_S} \right] \left[\frac{H(z)(1+z)}{dv_{\parallel}/dr_{\parallel}} \right] \text{ mK.}\end{aligned}\quad (3.10)$$

If the background is composed by CMB photons, then $T_R = T_{\text{CMB}}$. Brightness temperature fluctuations measurements are the observative key quantity related to cosmological signatures. 21 cm features are seen as a spectral distortion to the CMB blackbody at appropriate radio frequencies (since CMB temperature fluctuations are small and CMB is effectively a source of uniform brightness). So the distortion forms a diffuse background that can be studied across the whole sky in a similar way to CMB anisotropies. We can collect 3D map of 21 cm temperature brightness fluctuations, because observations at different z provide different shells of the observable Universe. It is clear from Eq.(3.10) that the detectability of this signal depends on the spin temperature, since the radiation will show up only if T_S deviates from T_{CMB} . T_S can be either above or below T_{CMB} , dictating whether the 21 cm signal will appear in emission, absorption, or not at all.

Apart from contrast with CMB photons, 21 cm line provides another important astrophysical probe. Suppose that the background radiation is made by a radio-loud point source, like a quasar at high redshift. In this case $T_R \gg T_S$, since the non-thermal source is always brighter than the diffuse HI emission ($T_{\text{src}} \approx 10^6 - 10^{10}$ K). Analogously to Ly α forest we see a 21 cm forest, made by absorption lines coming from regions of HI at different distances against the source. The high brightness of background sources allows for high frequency resolutions, making superb probes of cloud structure in neutral or partially reionized IGM.

3.2 Determining the Spin Temperature

From Eq.(3.10) we note that the generation of a 21 cm signal depends on the spin temperature value. The spin temperature is determined by three processes:

1. Absorption and/or emission of 21 cm photons from/to the radio background, mainly composed by CMB photons;
2. Collisions with other HI atoms, protons, and with free electrons;

3. Scattering of UV Ly α photons.

Excitation and de-excitation rates per atom must be balanced for thermal equilibrium, because all the relevant timescales are much shorter than the expansion time. If C_{10} and P_{10} are the de-excitation rates from collisions and UV scattering respectively, and if C_{01} and P_{01} are the corresponding excitation rates, we can write

$$n_1 (C_{10} + P_{10} + A_{10} + B_{10}I_{\text{CMB}}) = n_0 (C_{01} + P_{01} + B_{01}I_{\text{CMB}}), \quad (3.11)$$

where B_{01} and B_{10} are the Einstein coefficients for emission and absorption, and I_{CMB} is the energy flux of CMB photons. Thus, we know that, for detailed balance,

$$\frac{C_{01}}{C_{10}} = \frac{g_1}{g_0} e^{-T_\star/T_K} \approx 3 \left(1 - \frac{T_\star}{T_K}\right), \quad (3.12)$$

and

$$\frac{P_{01}}{P_{10}} = 3 \left(1 - \frac{T_\star}{T_\alpha}\right). \quad (3.13)$$

In these last two equations T_K is the gas kinetic temperature, while T_α is defined as the effective color temperature of the Ly α UV radiation field. Using them and Rayleigh-Jeans approximation, we can write Eq.(3.11) as

$$T_S^{-1} = \frac{T_{\text{CMB}}^{-1} + x_\alpha T_\alpha^{-1} + x_c T_K^{-1}}{1 + x_c + x_\alpha}, \quad (3.14)$$

where x_c and x_α are the coupling coefficients for atomic collisions and Ly α photons scattering, respectively. Here we have set $T_\gamma = T_{\text{CMB}}$ since the photon bath is mainly composed by CMB photons. In most astrophysical situations $T_\alpha \rightarrow T_K$, because of recoils after repeated scattering. In this limit Eq.(3.14) can be written as

$$\left(1 - \frac{T_{\text{CMB}}}{T_S}\right) = \frac{x_c + x_\alpha}{1 + x_c + x_\alpha} \left(1 - \frac{T_{\text{CMB}}}{T_K}\right). \quad (3.15)$$

So the spin temperature is strongly coupled to gas temperature if $x_c + x_\alpha \geq 1$ and relaxes to T_{CMB} if $x_c + x_\alpha \ll 1$. Now we are going to explicit these contributions.

3.2.1 Collisional Coupling

Spin-flip transitions in a hydrogen atom could be generated by collisions among different particles, causing excitation and de-excitation of the hyperfine levels. Such a contribution is important in the Early Universe, where the gas density is high. We can write down the generic coupling coefficient for a contribution produced by the i -th kind of particle as

$$x_c^i = \frac{C_{10}^i T_\star}{A_{10} T_\gamma} = \frac{n_i \kappa_{10}^i T_\star}{A_{10} T_\gamma}, \quad (3.16)$$

where C_{10} is the collisional excitation rate, T_γ is the temperature of the surrounding bath photons, and κ_{10}^i is the specific rate coefficient for spin de-excitation by collisions with species i , in units of cm^3s^{-1} . There are three main available channels for this process:

- collisions between two hydrogen atoms;
- collisions between a hydrogen atom and an electron;
- collisions between a hydrogen atom and a proton.

Other subdominant channels are collisions with deuterium atoms, helium atoms or ions. The relevant contributions to the total collisional coupling are the sum of the i -th coefficients, namely

$$x_c = x_c^{\text{HH}} + x_c^{\text{eH}} + x_c^{\text{pH}} = \frac{T_\star}{A_{10}T_\gamma} \left[\kappa_{10}^{\text{HH}}(T_K)n_{\text{H}} + \kappa_{10}^{\text{eH}}(T_K)n_{\text{e}} + \kappa_{10}^{\text{pH}}(T_K)n_{\text{p}} \right]. \quad (3.17)$$

In order to compute these rates, one would need to perform quantum mechanical calculations, which we will not treat but only see in a more qualitative way. For details we refer the reader to consult (Furlanetto et al., 2006) and the references there indicated.

HH Collisions

The HH collision can be described as a typical scattering event between two identical particles that form an intermediate (virtual) hydrogen molecule, before separating again. Let us label by a the $F = 0$ ground hyperfine state, while b, c, d are the $F = 1$ ones with $m = -1, 0, 1$, where m is the quantum number related to spin state. Remembering that an electron exchange must conserve spin number, we can have interactions producing $\Delta F = 2, 1, 0$. An example for these three cases are transitions like $cc \rightarrow aa$, $bd \rightarrow ac$, and $bd \rightarrow cc$, respectively. We can write the Schroedinger equation for the total wavefunction $\Psi(\mathbf{r}, \mathbf{R})$, where \mathbf{r} is the two electrons' positions, and \mathbf{R} is the vector joining the two nuclei. For slow collisions, the general solution for this equation is a superposition of the lowest energy eigenstates, *i.e.* the singlet χ_s and the triplet χ_t states of hydrogen molecule, so

$$\Psi(\mathbf{r}, \mathbf{R}) = F_s(\mathbf{R})\chi_s(\mathbf{r}, \mathbf{R}) + F_t(\mathbf{R})\chi_t(\mathbf{r}, \mathbf{R}). \quad (3.18)$$

In the elastic scattering limit, the asymptotic solutions for $F_{s,t}$ can be written in spherical coordinates, whose angular parts $f_{s,t}$ are expanded in Legendre polynomials. In the end the problem is reduced to an infinite set of radial equations, indexed by the order l of the associated Legendre polynomial. The phase shifts $\delta_l^{s,t}$ quantify the coherence of the scattering amplitudes over the different waves. These phases are determined by the H_2 energy potential curves in singlet and triplet states. So, the total cross section averages over the spins of the particles and must reflect the antisymmetry of Ψ for two identical particles with respect to the interchange of two nuclei. It can be showed that the resulting cross sections are

$$\sigma^\pm = \frac{\pi}{4k^2} \sum_{l=0}^{\infty} (2l+1) \sin^2(\delta_l^s - \delta_l^t) \left[1 - (-1)^{l+1/2 \pm 1/2} \right], \quad (3.19)$$

where k is the relative momentum module, and a proportionality dependance on $|f_i - f_s|^2$ is present, reflecting the fact that this is a coherent sum. If we had distinguishable particles, the square bracket factor would be 1, and hence there would be only one cross section. Usually, the relevant cross section is the smoothed thermally averaged cross section:

$$\sigma^\pm = \frac{1}{(k_B T_K)^2} \int E dE \sigma^\pm(E) e^{-E/k_B T_K}. \quad (3.20)$$

From detailed balance we can write the excitation rates as

$$k_x^\pm = \sqrt{\frac{8k_B T_K}{\pi M}} \sigma^\pm e^{-\omega^\pm} = k^\pm e^{-\omega^\pm}, \quad (3.21)$$

where $\omega^- = \omega = E_{10}/k_B T_K$ and $\omega^+ = 2\omega$. Having these cross sections, we can compute how the level populations evolve, assuming they are independent of atomic velocities. Assuming statistical and thermodynamic equilibrium among these sublevels, a linearised balance between excitations and de-excitations can be written down, namely

$$\dot{n}_1 = n_0 \kappa_{01}^{\text{HH}} n_{\text{HI}} - n_1 \kappa_{10}^{\text{HH}} n_{\text{HI}}, \quad (3.22)$$

with $n_0 = n_a$, $n_1 = n_b + n_c + n_d$, and

$$\kappa_{10}^{\text{HH}} = \frac{k^+ + k^-}{2} = \kappa_{01}^{\text{HH}} \frac{e^\omega}{3}. \quad (3.23)$$

A useful fitting function exists in the range $T_K = [10, 1000]$ K (Kuhlen et al., 2006)

$$\kappa_{10}^{\text{HH}}(T_K) \approx 3.1 \times 10^{-11} T_K^{0.357} e^{-32/T_K} \text{ cm}^3 \text{ s}^{-1}. \quad (3.24)$$

The linearisation is sufficiently accurate throughout the regime of interest. By the way, when collisions dominate the assumption of atomic velocity independence is not always good. The actual velocity dependence leads to a non-thermal distribution for the hyperfine occupation (Hirata & Sigurdson, 2007). This effect causes a 5% suppression of 21 cm signal level, which could be important for precision cosmology studies during Dark Ages.

eH Collisions

Free electrons can induce spin exchange when they collide with HI. This scattering problem is solved exactly by following the same steps as the one overviewed for HH collisions for distinguishable particles. The cross section will have the form of Eq.(3.19), without any factor involving nuclear symmetry. The fitting function for this collision rate is (Liszt, 2001)

$$\log \kappa_{10}^{\text{eH}}(T_K) = \begin{cases} -9.607 + 0.5 \log T_K e^{-(\log T_K)^{4.5}/1800} & T_K \leq 10^4 \text{ K} \\ \log \kappa_{10}^{\text{eH}}(10^4 \text{ K}) & T_K > 10^4 \text{ K} \end{cases}. \quad (3.25)$$

As can be seen in Figure 3.1, $\kappa_{10}^{\text{eH}} \gg \kappa_{10}^{\text{HH}}$ because, at a fixed temperature, free electrons have much larger velocities than HI atoms. Unlike for HH collisions, the eH cross section does not cut off for small T_K .

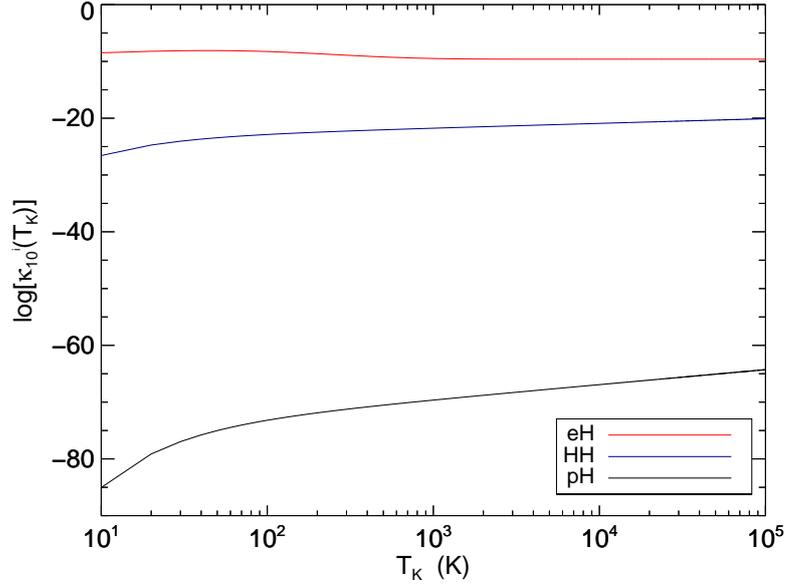


Figure 3.1: The spin de-excitation rate eH, HH, pH collisions, in red, blue, and black respectively.

pH and Other Collisions

As shown in Figure 3.1, proton collisions are generally subdominant, since $\kappa_{10}^{\text{pH}} \approx 3.2\kappa_{10}^{\text{HH}}$ for relatively high temperatures, and are thus less efficient than free electrons (Furlanetto et al., 2006).

There could be collisions with other species like neutral He, but Pauli exclusion principle prevents electron exchange in the ground state and spin-flip transitions, unless He is found excited to the triplet state. This possibility is quite unlikely since it requires significantly more energy than the one provided usually by cold neutral IGM. Ionized He could be significant in partially ionized gas, though free electrons will still dominate because of their high velocities.

Deuterium collisions are possible but, although they produce bigger cross sections at small T_K respect to HH ones, are a rare event. They will not hence produce a significant effect on the spin temperature.

So, HH collisions dominate in the early Universe. When the Universe is highly heated and ionized, collisions with electrons become a strong coupling mechanism, and its high-temperature behaviour is important.

3.2.2 The Wouthuysen-Field Effect

As realised at the end of previous section, coupling processes are inefficient for most of the redshifts that are likely to be observationally probed in next years. In fact this contribution is important only during Dark Ages, namely for $z > 40$. Once first stars begin to form, resonant scattering of Ly α photons, also known as Wouthuysen-Field effect (Field, 1958) can cause a spin-flip transition in HI atoms. We illustrate this effect in Figure 3.2 considering the hyperfine splittings of the 1S and 2P levels of HI. Consider an HI atom which finds

itself in the hyperfine singlet state. For the dipole selection rules, excitations due to Ly α photon absorption can only produce transitions with total angular momentum difference $\Delta F = 0, 1$, excluding the transition $F = 0 \rightarrow 0$. So the atom will be in one of the central $2P$ hyperfine states pictured in left panel of Figure 3.2. The de-excitation can happen with a spontaneous emission of a Ly α , making the HI atom decay to either of the two ground state hyperfine levels. If the HI atom decays into the ground level triplet state, then a spin-flip occurred. In reality this could happen with the emission of any Lyman-series photon. In fact, Ly α photons can be produced by atomic cascades from photons redshifting into these higher n series levels, as shown in the right panel of Figure 3.2. The Ly α flux is substantially amplified, since the photon distribution will be one-sided. For large n , the conversion rate into Ly α is about 30%.

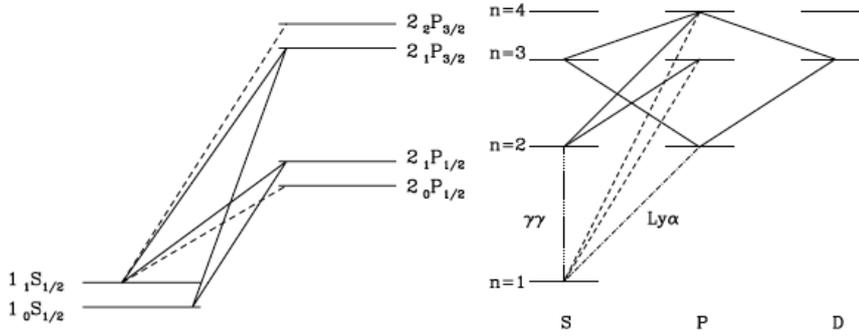


Figure 3.2: *Left panel:* Hyperfine structure of the hydrogen atom and relevant transitions for the Wouthuysen-Field effect. Solid lines allow spin-flip transitions, while dashed lines represent allowed transitions but not contributing to spin-flips. *Right panel:* the conversion process of Ly n into Ly α via atomic cascades. Taken from Pritchard & Loeb (2012).

The Wouthuysen-Field effect will produce a Ly α photons scattering rate within the IGM

$$P_\alpha = 4\pi \int d\nu J_\nu(\nu) \sigma_\nu, = 4\pi \chi_\alpha \int d\nu J_\nu(\nu) \phi_\alpha(\nu), \quad (3.26)$$

where σ_ν is the local absorption cross-section, $\chi_\alpha = (\pi e^2/m_e c) f_\alpha$ is the oscillation strength of the Ly α transition, $\phi_\alpha(\nu)$ is the Ly α absorption profile and $J_\nu(\nu)$ is the angle averaged specific intensity of the background photon field. The excitation rate P_{10} can be computed considering all the possible emitting and absorbing processes allowed by spin vector rules. Assuming a constant radiation field across the hyperfine lines, one has $P_{10} = P_\alpha(4/27)$. Hence, the coupling coefficient is

$$x_\alpha = \frac{4P_\alpha}{27A_{10}} \frac{T_\star}{T_\gamma} = \frac{16\pi^2 T_\star e^2 f_\alpha}{27A_{10} T_\gamma m_e c} S_\alpha J_\alpha = S_\alpha \frac{J_\alpha}{J_\alpha^C}, \quad (3.27)$$

where we have used Eq.(3.26), J_α is the specific flux computed at Ly α frequency, and

$$S_\alpha = \int dx \phi_\alpha(x) \frac{J_\nu(x)}{J_\infty} \quad (3.28)$$

is a correction factor of order unity introduced to describe variations of photon distribution near the Ly α resonance. J_∞ is the flux away from the absorption line, while $J_\alpha^C = 1.165 \times 10^{10} [(1+z)/20] \text{ cm}^{-2}\text{s}^{-1}\text{Hz}^{-1}\text{sr}^{-1}$. We get a critical value of the coupling coefficient for $x_\alpha = S_\alpha$. This condition is easily reached when first stars begin to form, and this is explicitly noticeable expressing this critical flux in terms of Ly α photons per HI atom $J_\alpha^C/n_{\text{HI}} = 0.0767 [(1+z)/20]^{-2}$.

We have seen in Eq.(3.14) that the Ly α coupling depends on the color temperature T_α other than on the coupling coefficient. The color temperature is a measure of the shape of the radiation field near the Ly α line. Considering the photon occupation number $n_\nu = c^2 J_\nu / 2\nu^2$, this is (Rybicki, 2006)

$$\frac{h}{k_B T_\alpha} = -\frac{d \log n_\nu}{d\nu}. \quad (3.29)$$

As long as the medium is optically thick, there is a big number of scattering events. This makes the Ly α profile equal to a blackbody of temperature T_K near the central frequency, where there is local equilibrium. Because of cosmic expansion, the photons will flow in frequency, and, once they reach the Ly α frequency, they can scatter to larger or smaller frequencies. Hence, the net flow rate is preserved if the cross section is symmetric. This will not be true, since at every scattering event a Ly α will lose energy due to the recoil of the atom. The result is a net deficit of photons near the central frequency and a flow towards lower energies: the photon distribution is asymmetric and tilted to the red. Because of this, the local equilibrium is restored with $T_\alpha \approx T_K$. This asymmetry determines S_α , and, since recoil produces an absorption feature, $S_\alpha \leq 1$. The Wouthuysen-Field effect will be more suppressed for low IGM temperatures. This picture is complicated by spin excitation, which contributes to the fraction of lost energy through recoil with less than 10%.

In principle we need to compute the photon spectrum near the Ly α frequency more accurately, and Monte Carlo methods are used to satisfy the resulting radiative transfer equation in Fokker-Planck limit, *i.e.* when this equation is writable as a steady-state diffusion equation. A useful approximation for the profile is the Voigt profile $S_\alpha \approx e^{-1.79\alpha}$, with $\alpha = \eta (3\Gamma/8\pi^2\gamma\Delta\nu_D)^{1/3}$, in which Γ is the inverse lifetime of the upper 21 cm level, $\Delta\nu_D = \nu_0 (2k_B T_K/mc^2)^{1/2}$ is the Doppler parameter, ν_0 is the line center frequency, $\gamma = \tau_{\text{GP}}^{-1}$ is the inverse of the Gunn-Peterson optical depth, and $\eta = h\nu_0^2/mc^2\Delta\nu_D$ is the mean frequency drift per scattering due to recoil. This expression is accurate to $\sim 5\%$ for $T_K \geq 1$ and for high τ_{GP} . At smaller temperatures, we would need a numerical solution including a full Voigt profile and the spin exchange.

All these effects can modify the 21 cm signal at the $\sim 10\%$ level, which is substantial for 21 cm fluctuations measurements.

3.3 21 cm Radiation Global Evolution

21 cm temperature brightness fluctuation history can be written as a function of four parameters. These are the IGM kinetic temperature T_K , the volume-

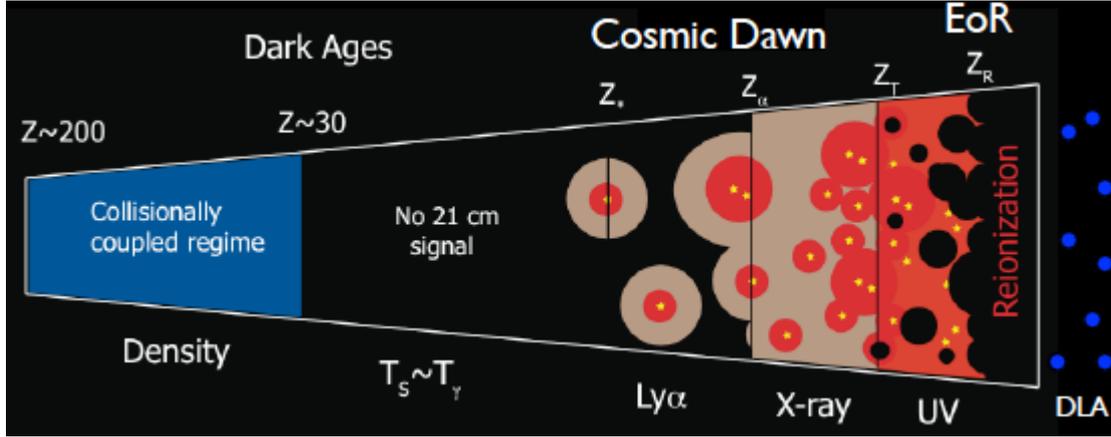


Figure 3.3: Illustration of the global history of 21 cm radiation across a redshift range going from Dark Ages to today. The first transition is between an early phase of collisional coupling to a later phase of Ly α through a short period in which the signal is very low. Then fluctuations will be dominated by spatial variation in Ly α , X-rays, and ionizing UV background. After EoR, the only residual signal is found within galaxies. Taken from Pritchard & Loeb (2012).

averaged ionized fraction of HI x_i , the flux computed at Ly α frequency J_α , and the neutral fraction of HI n_{HI} . There are cosmic epochs in which one of these variables contribution dominates the signal over the others. Following Figure 3.3, we have:

- $200 \lesssim z \lesssim 1100$: here we do not expect to see any 21 cm signal, since $T_S = T_\gamma$. In fact, the high-density IGM will be collisionally coupled to CMB because of Compton scattering due to the residual free electron fraction left after recombination, and so $T_K = T_\gamma$.
- $40 \lesssim z \lesssim 200$: during this period the gas cools adiabatically. This means $T_K < T_\gamma$, and $T_S < T_\gamma$ because of collisional coupling. This causes the 21 cm signal to be potentially detected in absorption against the CMB field. The temperature brightness fluctuations are sourced by the mass density fluctuations at these redshifts, allowing us to probe initial conditions (Loeb & Zaldarriaga, 2004). A tomographic detection of 21 cm absorption will map the HI density, and its predictions will be straightforward and robust, because the 21 cm absorption physics is well understood.
- $z_* \lesssim z \lesssim 40$: Universe expansion makes the HI gas less dense, making collisional coupling ineffective. The IGM will be coupled to CMB, setting $T_S = T_\gamma$. So there is no 21 cm signal produced.
- $z_\alpha \lesssim z \lesssim z_*$: during the Cosmic Dawn, first stars will switch on, emitting Ly α and X-rays photons. The Ly α is considerably smaller than that required for heating the IGM over the CMB. So here we expect to find the spin temperature coupled to gas, namely $T_S \sim T_K < T_\gamma$, resulting in an absorption signal. After a certain amount of time near z_α , the Ly α coupling will saturate and $x_\alpha \gg 1$.

- $z_h \lesssim z \lesssim z_\alpha$: because of Ly α saturation, fluctuation in its flux do not determine 21 cm signal anymore. So, heating is significant and brightness temperature fluctuations are originated by IGM. As T_K approaches to T_γ , the signal will be progressively seen in emission from hotter regions. At z_h the gas will be uniformly heated and $T_K = T_\gamma$.
- $z_T \lesssim z \lesssim z_h$: at this point $T_K > T_\gamma$ and the signal is seen in emission. Brightness temperature fluctuations are sourced by fluctuations in ionization, density and gas temperature. Here the ionization fraction becomes important and 21 cm brightness temperature goes towards saturation at z_T .
- $z_r \lesssim z \lesssim z_T$: heating will keep increasing T_K , making fluctuations unimportant. At this point $T_S \sim T_K \gg T_\gamma$, and in Eq.(3.10) $1 - (T_\gamma/T_S) = 1$. Here the filling fraction of HII regions becomes significant and ionization fluctuations begin to dominate 21 cm signal.
- $z \lesssim z_r$: after EoR, any 21 cm radiation will be originated from isolated HI systems (like galaxies).

Consider that there is no certainty on the correct definition of these epochs: most of them could be overlapped and not happen in this order. For example, if X-ray preheating is more important if it allows collisional coupling, and $z_h > z_\alpha$. There could also be very early weak shocks in IGM, if heating is much more efficient due to dark matter annihilations or other exotic particle mechanisms.

3.3.1 Evolution of the IGM

At the beginning, the IGM is made of HI with temperature T_K and a small fraction of electrons x_e . When first stars turned on, energetic UV photons ionize HII regions, whose boundaries are quite sharp because of the short mean path of the photons. These ionized HII bubbles will evolve, and become fully ionized with fixed uniform temperature $T_{\text{HII}} = 10^4$ K, which determines the collisional recombination rate inside them, and a volume filling fraction x_i . Since the photons redshifted into the Ly α resonance have initial long mean free paths, the Ly α flux, J_α , is the same during these phases. So the evolution of the 21 cm is parametrised by four quantities: x_i , x_e , T_K , and J_α .

For the temperature we have

$$\frac{dT_K}{dt} = \frac{2T_K}{3n} \frac{dn}{dt} + \frac{2}{3k_B} \sum_j \frac{\epsilon_j}{n}, \quad (3.30)$$

where the first term express adiabatic cooling of IGM due to Universe expansion, and the second term indicates other sources of heating or cooling having ϵ_j rate per unit volume for the process j . For volume filling fraction and electron fraction, we have

$$\frac{dx_i}{dt} = (1 - x_e) \Lambda_i - \alpha_A C x_i^2 n_{\text{H}}, \quad (3.31)$$

$$\frac{dx_e}{dt} = (1 - x_e) \Lambda_e - \alpha_B(T) x_e^2 n_{\text{H}}, \quad (3.32)$$

where Λ_i is the rate of production of ionizing photons per unit time per baryon applied to HII regions, Λ_e is the same but in IGM bulk, $\alpha_A = 4.2 \times 10^{-13} \text{ cm}^3 \text{ s}^{-1}$ is the recombination coefficient at $T = 10^4 \text{ K}$, $\alpha_B(T)$ is the recombination rate for the electron fraction evolution, and $C = \langle n_e^2 \rangle / \langle n_e \rangle^2$ is the clumping factor. Both equations are a balance between ionizations and recombinations in dense clumps which are thick respect to ionizing radiation. The clumping factor C takes into account matter inhomogeneities due to this local recombination rate in these fully ionized bubbles. The secondary photons produced by recombinations on bubbles edges are likely to be absorbed inside the clumps rather than in the IGM. In the second case B, recombinations in the IGM bulk are absorbed by the IGM itself.

The balance between recombination and ionization determines the growth of ionized HII regions. Let us define the ionization rate per H atom as

$$\Lambda_i = A_{\text{He}} f_{\text{esc}} N_{\text{ion}} \dot{\rho}_*(z), \quad (3.33)$$

with N_{ion} the number of ionizing photons per baryons produced in stars, A_{He} a correction factor which takes into account the presence of Helium, and f_{esc} the fraction of ionizing photons that escape the halo. $\dot{\rho}_*(z)$ is the star formation rate density, which is usually modeled by tracking the collapsed matter, namely

$$\dot{\rho}_*(z) = \bar{\rho}_b^0 f_* \frac{df_{\text{coll}}(z)}{dt}, \quad (3.34)$$

with $f_{\text{coll}}(z)$ the fraction of gas inside collapsed objects at z , $\bar{\rho}_b^0$ the cosmic mean baryon density today, and f_* the fraction of baryons converted into stars. This formalism is valid for $z \gtrsim 10$, since on late times merging gives another contribution to star formation. So, the ionization rate is

$$\Lambda_i = A_{\text{He}} f_* f_{\text{esc}} N_{\text{ion}} \frac{df_{\text{coll}}(z)}{dt} = \varepsilon(z) \frac{df_{\text{coll}}(z)}{dt}, \quad (3.35)$$

where we have defined the ionization efficiency parameter. To compute f_{coll} one uses a mass function dn/dm and determines a minimum mass m_{min} for having collapse by setting the virial temperature $T_{\text{vir}} \geq 10^4 \text{ K}$.

It is important to set the properties of first galaxies, since they are the main sources of ionizing photons, but these are poorly constrained. Difficulties are posed by small sky coverage and limited frequency coverage. Moreover, remember that the UV photons are seen in the optical range, and this does not necessarily correspond to the UV origin range. The mass distribution is uncertain, so the number of ionizing photons N_{ion} is not well determined. The clumping factor can depend on the composition of the bubbles, feedback mechanisms, and hydrodynamical effects. Usually one assumes that the Universe is fully ionized up to some critical overdensity, and, provided an expression for the probability distribution of the gas, the clumping factor is the second moment of this distribution. This critical overdensity depends on the effective photon path, which is determined by bubbles' dimensions and so depends on the patchy nature of EoR.

3.3.2 Sources for Heating and Ionization

The heating rate is redshift dependent, since for every epoch we need to know which mechanisms are relevant to solve Eq.(3.30). At high z the most relevant effect is Compton heating of the IGM caused by CMB photons scattering from a small residual free electron fraction. This is important for $z \gtrsim 150$ and couples T_K to T_γ , setting initial conditions for star formation. At lower z other sources of heat, like shocks caused by IGM turnarounds undergone before collapsing and separated from the Hubble flow, become important because of non-linear structures growth. We need to take into account the scattering of Ly α photons against H atoms, but it requires large Ly α fluxes to heat the gas to the CMB temperature.

The most important source of energy injection into the IGM at these redshifts is likely X-ray heating. As soon as compact objects are formed, X-ray photons can be produced in large number and have a long mean free path. This quantity is a function of energy (Furlanetto et al., 2006), namely

$$\lambda_X \approx 4.9 \bar{x}_H^{-1/3} \left(\frac{1+z}{15} \right)^{-2} \left(\frac{E}{300 \text{ eV}} \right)^3 \text{ Mpc}. \quad (3.36)$$

From this relation we can see that the Universe will be optically thick over a Hubble length to all photons with energy lower than $E \approx 2[(1+z)/15]^{1/2} \bar{x}_H^{1/3}$ KeV. The E^{-3} dependence of cross-section means that heating is dominated by soft X-rays, with an additive uniform component caused by harder X-rays. X-rays will photo-ionize the IGM, exciting HI and HeI: this generates photo-electrons, which will de-excite by heating, secondary ionizations, and atomic excitation.

The total rate of energy deposition per unit volume ϵ_X is summed over the i species involved, and must contain the fraction of energy converted into form i at a specific frequency $f_i(\nu, x_e)$, in order to divide the energy for every heating, ionization, and excitation. The X-ray number flux is

$$J_X(z) = \int_{\nu_{\text{th}}}^{\infty} d\nu J_X(\nu, z), = \int_{\nu_{\text{th}}}^{\infty} d\nu \int_z^{z^*} dz' \frac{(1+z)^2}{4\pi} \frac{c}{H(z)} \hat{\epsilon}_X(\nu', z') e^{-\tau(\nu, z, z')}, \quad (3.37)$$

where $\hat{\epsilon}_X(\nu, z)$ is the comoving photon emissivity, ν_{th} is the threshold energy for ionization, $\nu' = \nu(1+z)/(1+z')$, and the optical depth depends on the cross-section of each process involved,

$$\tau(\nu, z, z') = \int_z^{z'} \frac{dl}{dz''} dz'' [n_{\text{HI}}\sigma_{\text{HI}}(\nu'') + n_{\text{HeI}}\sigma_{\text{HeI}}(\nu'') + n_{\text{HeII}}\sigma_{\text{HeII}}(\nu'')]. \quad (3.38)$$

The main astrophysical processes producing X-rays are supernova remnants (SNR), starburst galaxies with high rates of star formation, and mini-quasars. Starburst galaxies are likely producers of low-mass and high-mass X-ray binaries. These latter are expected to be dominant at EoR redshifts in the contribution of L_X tracking the star formation rate (SFR). However they form a contribution which is quite uncertain (Mirabel et al., 2011; Dijkstra et al., 2012).

SNR produce X-rays via inverse Compton scattering, but the estimates on the luminosity are still uncertain, because of the dependence on SFR. Mini-quasars produce X-rays because of accretion onto black holes, but here uncertainties are due to seed source and merger history of black holes. They will track SFR too, but the evolution is more complex.

As stated by Dijkstra et al. (2004), complete X-rays re-ionisation is ruled out because of the upper limit imposed by the present day unresolved X-ray background (SXRb). On the other hand, heating requires less energy than re-ionization, and it could have an high importance. Constraining the the spectral distribution function for X-ray emissivity will allow a better understanding of the thermal history of IGM and the population of X-ray sources at high redshifts.

3.3.3 Evolution of Ly α Flux

The Ly α flux is determined mainly by stellar emission $J_{\alpha,\star}$ and X-ray excitation of HI $J_{\alpha,X}$. Photons emitted by stars between Ly α and the Lyman limit are redshifted until they enter a Lyman series resonance. So the total contribution comes from a sum over the Ly n levels

$$J_{\alpha,\star}(z) = \sum_{n=2}^{n_{max}} J_{\alpha}^{(n)}(z) = \sum_{n=2}^{n_{max}} f_{recyc}(n) \int_z^{z_{max}(n)} dz' \frac{(1+z)^2}{4\pi} \frac{c}{H(z')} \hat{\epsilon}_{\star}(\nu'_n, z'), \quad (3.39)$$

where $n_{max} \sim 23$ is set by the size of the HII region of a typical isolated galaxy, $z_{max}(n)$ is the maximum redshift from which emitted photons are redshifted into the resonance, ν'_n is the emission frequency at z' in which the photon is absorbed and f_{recyc} is the probability of producing a Ly α photon by cascade from level n . $\hat{\epsilon}_X(\nu, z)$ is the comoving photon number emissivity for stellar sources and is linked to SFR similarly to X-rays. Stellar sources typically have a spectrum falling quickly above the Ly β level.

Photoionization of HI by X-rays can produce Ly α photons because of relaxation. The rate is similar to the one due to X-ray heating, but with a different fraction of total X-ray energy that goes into excitations rather than heating. This is not a trivial issue and has to be addressed by Monte-Carlo simulation. Pritchard & Furlanetto (2007) found that a 70% of the total energy going into excitation produces Ly α photons.

3.3.4 Heating by Exotic Mechanisms

Because of the direct connection between 21 cm brightness temperature and the IGM temperature, it is possible to model the physics beyond the Standard Model (SM) and make concrete predictions on exotic heating mechanisms like Dark Matter annihilation. This should contribute to the measured value of Ω_m . This contribution could be important both at early times where DM has yet to be diluted by cosmic expansion, and when a significant number of DM haloes is formed. Other mechanisms include decaying into SM particles or photons.

Other heating mechanisms could be produced because of Hawking radiation generated by black holes evaporation and moving cosmic strings.

In general incorporating heating requires a knowledge of the energy spectrum of photons produced by the source. This then has to be processed to determine how much radiative energy goes into the IGM.

3.4 21 cm Brightness Fluctuation Power Spectrum

Brightness temperature fluctuations express deviation from the 21 cm global signal averaged over large angular scales, which can be seen as a zeroth order approximation of the full signal. This latter is highly inhomogeneous, because of the inhomogeneity of involved radiation fields and of IGM features.

We will treat these perturbations in an analogous way to what already studied for the CMB case (*e.g.* Hu & White (1997); Hu & Dodelson (2002)) but in the flat sky approximation. Hence, the statistical properties of the 21 cm signal will be expressed in terms of 21 cm brightness temperature fluctuation power spectrum. We can expand Eq.(3.10) up to linear order, so that

$$\delta T_b = \beta_b \delta_b + \beta_x \delta_x + \beta_\alpha \delta_\alpha + \beta_T \delta_T - \delta_{\partial v}, \quad (3.40)$$

where each δ_i describes the fractional variation for the i -th quantity, which are baryon density, neutral HI fraction, Ly α coupling coefficient, IGM temperature, and LoS peculiar velocity gradient, respectively. The β_i expansion coefficients are

$$\begin{aligned} \beta_b &= 1 + \frac{x_c}{x_{\text{tot}}(1 + x_{\text{tot}})}, \\ \beta_x &= 1 + \frac{x_c^{\text{HH}} - x_c^{\text{eH}}}{x_{\text{tot}}(1 + x_{\text{tot}})}, \\ \beta_\alpha &= \frac{x_\alpha}{x_{\text{tot}}(1 + x_{\text{tot}})}, \\ \beta_T &= \frac{T_\gamma}{T_K - T_\gamma} + \frac{1}{x_{\text{tot}}(1 + x_{\text{tot}})} \left(x_c^{\text{eH}} \frac{d \log \kappa_{10}^{\text{eH}}}{d \log T_K} + x_c^{\text{HH}} \frac{d \log \kappa_{10}^{\text{HH}}}{d \log T_K} \right), \end{aligned} \quad (3.41)$$

where $x_{\text{tot}} = x_c + x_\alpha$ and the collisional term is split into the dominant eH and HH components. We have assumed $T_c = T_K$: if this is not valid the expressions would be more complicated. By the way, these perturbation fields have to be properly written for specific physical regimes. In general the baryonic density field δ_b is equivalent to the total matter density δ , because the 21 cm background is directly related to δ_b , or to the HI density field. On small scales the finite pressure of the baryons introduces a cutoff absent from the dark matter (Naoz & Barkana, 2005), and galaxy formation processes and feedback can also work on the two separately. For $z < 10$, β_b and β_x tend to be ~ 1 , while β_α and β_T get close to 0.

The power spectrum is defined as

$$\langle \tilde{\delta T}(\mathbf{k}) \tilde{\delta T}^*(\mathbf{k}') \rangle = (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{k}') P_{\delta T}(k), \quad (3.42)$$

and can contain all possible terms of the form $P_{\delta_i \delta_j}$ which have to be written for a specific physical situation at a given redshift. 21 cm observations hope to measure individually these quantities, even if some of them are correlated. For example the collision rate is directly linked to the ionized fraction, or δ_α is related on neutral fraction and temperature because of the Wouthuysen-Field effect.

Now we are going to examine each of the contributions written in Eq.(3.40).

3.4.1 Redshift Space Distortions

The power spectrum Eq. (3.42), should be spherically symmetric in Fourier space and should depend on $|k| = k$, but the term $\delta_{\partial v}$ breaks this symmetry since redshift space distortions induce a preference over the LoS direction, preserving cylindrical symmetry. This effect, studied first by Kaiser (1987), is caused by the flowing of the bulk on large scales, and in particular by infalling onto massive structures, which causes compression in redshift space. On small scales, random motions in virialized regions create elongation in redshift space (the so-called Finger of God effect) when the peculiar velocity is high.

Peculiar velocities can have a significant effect on the 21 cm signal because our observations are made in frequency space, while the theory predicts observables in coordinate space. The conversion between the two is affected by the local bulk velocity of the gas. If we consider Eq.(2.18), we can express the Hubble parameter in comoving space using the conformal time, such that $H = (1/a)da(\eta)/d\eta$. The coordinate distance will be modified by the peculiar velocity effect by

$$\chi = \chi_z - \mathbf{v}(\mathbf{x}) \cdot \frac{\hat{\mathbf{n}}}{H} \Big|_z. \quad (3.43)$$

If $\mathbf{x} = \chi \hat{\mathbf{n}}$ is the real space coordinate and $\mathbf{s} = \chi_z \hat{\mathbf{n}}$ is the redshift space coordinate, the mapping between the two is

$$\mathbf{s} = \mathbf{x} + \left[\mathbf{v}(\mathbf{x}) \cdot \frac{\hat{\mathbf{n}}}{H} \right] \hat{\mathbf{n}}, \quad (3.44)$$

where the term $\mathbf{s} - \mathbf{x}$ generates the redshift space distortions. As demonstrated by Kaiser (1987), in linear theory we can write the radial peculiar velocity gradient can be written as

$$\nabla \cdot \mathbf{v}(\mathbf{x}) = -Hf(\Omega_m) \delta(\mathbf{x}), \quad (3.45)$$

whose Fourier transform is proportional to the Fourier transform of the density field, namely

$$\tilde{\nabla}_r \mathbf{v} = -\mu^2 f(\Omega_m) \tilde{\delta}, \quad (3.46)$$

where $\mu = k_{\parallel}/k$ is the cosine of the angle between the wavevector \mathbf{k} and the its component parallel to the LoS direction, and

$$f(\Omega_m) = \Omega_m^{0.6}(z) = \frac{d \ln D}{d \ln a} \quad (3.47)$$

tracks the evolution of the linear growth factor. At high redshift the Universe is matter dominated, so $f \approx 1$. Hence, we can write Eq.(3.40) as formed by an isotropic and a non-isotropic part in Fourier space

$$\delta\tilde{T}_b = \mu^2\tilde{\delta} + \tilde{\delta}_i. \quad (3.48)$$

Neglecting second order terms, we have

$$P_{\delta T}(\mathbf{k}) = \mu^4 P_{\delta\delta} + 2\mu^2 P_{\delta_i\delta} + P_{\delta_i\delta_i}. \quad (3.49)$$

If we neglect temperature or ionization fluctuations, the velocity term boosts the power spectrum by a factor $\langle(1 + \mu^2)^2\rangle = 1.87$: redshift space distortions are non-negligible. This is true for the angular power spectrum as well. If we explicit the power spectrum contributions, we get

$$\begin{aligned} P_{\delta T}(\mathbf{k}) = & P_{bb} + P_{xx} + P_{\alpha\alpha} + P_{TT} + 2P_{bx} + 2P_{\alpha b} + 2P_{bT} + 2P_{\alpha x} + 2P_{xT} + 2P_{\alpha T} \\ & + P_{x\delta x\delta} + \text{other quartic terms} \\ & + 2\mu^2 (P_{\delta b} + P_{\delta x} + P_{\delta\alpha} + P_{\delta T}) \\ & + \mu^4 P_{\delta\delta} \\ & + 2P_{\delta x\delta_{\partial v}x} + P_{\delta_{\partial v}\delta_{\partial v}xx} + \text{other quartic terms with } \delta_{\partial v}. \end{aligned} \quad (3.50)$$

It is clear that all quartic terms must be quadratic in x_H , and their separation depends on whether they contain powers of $\delta_{\partial v}$ or not. Those that contain powers of $\delta_{\partial v}$ are anisotropic and cause the angular dependence of $P_{\delta T}$. So, in principle it is possible to isolate the contribution from $P_{\delta\delta}$, as already showed by Barkana & Loeb (2005); Lidz et al. (2007), and all the quartic terms with anomalous μ dependency can be grouped into a term $P_{f(k,\mu)}$, so

$$P_{\delta T}(k, \mu) = P_{\mu^0}(k) + \mu^2 P_{\mu^2}(k) + \mu^4 P_{\mu^4}(k) + P_{f(k,\mu)}(k, \mu). \quad (3.51)$$

Assuming independent measurements of the matter power spectrum, we can extrapolate from the velocity field the factor depending on $x_H(1 - T_\gamma/T_S)$, and hence extract $x_H(z)$ at late times when $T_S \gg T_\gamma$. High precision measurements of 3D power spectrum can allow for the angular separation of $P_{\delta T}(k, \mu)$, but this goal depends on $P_{f(k,\mu)}$ form, which is important during the end of EoR. The feasibility of this measurement is still uncertain. As written by Pritchard & Loeb (2012), the first generation of 21 cm could measure the angle averaged quantity

$$\bar{P}_{\delta T}(k) = P_{\mu^0}(k) + P_{\mu^2}(k)/3 + P_{\mu^4}(k)/5. \quad (3.52)$$

Consider that it is required a high signal-to-noise to measure 21 cm power spectrum. But this is challenging, since the noise is anisotropic: radio foregrounds are likely to have more power across the sky than on the LoS direction.

Moreover, second-order terms can be non-negligible. Remember that fluctuations in x_H can be of order unity, so terms in higher order of δ_x which can be misleadingly thought to be small, can still contribute at a significant level to the power spectrum, leading to non-trivial four-point terms in the power spectrum (McQuinn et al., 2006) with undefined μ -angular dependence. These

terms make attempts to separate the μ^n powers during reionization more difficult. There could be an improvement only before δ_x becomes important.

At linear order, redshift space distortion contribution is well understood (Bharadwaj & Ali, 2005; Wang & Hu, 2006), since the $\delta_{\partial v}$ is related to the total density field. But this latter evolves, and non-linear corrections to the velocity field become important, especially at redshifts where the EoR is not uniform. Mao et al. (2012) studied this effect in detail studying three different models: the first one is a model without redshift space distortion contribution, the second one is a quasi-linear μ -decomposition scheme derived slightly modifying Eq.(3.51), while the third one is a numerical scheme that finds the fully non-linear redshift-space distorted 21 cm brightness temperature signal as a function of position and frequency. The peculiar velocity distorts the 21 cm mapping not only by shifting the apparent location in redshift space, but also by modifying the brightness temperature in real space. The relation between these two effects is broken down for $T_S \lesssim T_{\text{CMB}}$ or when $\tau \gtrsim 1$. In fact, high peculiar velocities can be larger than the Hubble flow in virialized haloes, but the optically thin 21 cm approximation is valid in most of the cases, making the observed power spectrum in redshift space remain finite. In the high spin temperature regime, the linear theory derived by Barkana & Loeb (2005) is 30% level non-accurated during EoR ($\sim 50\%$ of the ionized epoch). Their quasi-linear scheme incorporates relevant higher order correlations of ionization and density fluctuations, and this differs from the numerical scheme by 10%. Non-linearities may introduce larger deviations when the 3D power spectrum is decomposed to extract only the $P_{\mu^A}(k)$ for cosmology, and further investigations are needed to study the nature of these contributions. Moreover their accurate scheme avoids divergences that appear in real space evaluation when peculiar velocities gradients are large. These gradients appear because of non-linear structure formation on small scales, but more studies are needed to understand the dependence of the errors from grid dimensions, redshift and ionization fraction.

3.4.2 The Alcock-Paczynski Effect

The Alcock-Paczynski (AP) effect is another contribution that generates an isotropic term on the power spectrum. This depends on the underlying background cosmology, and it is due to the general difference between transverse and LoS distances scale. Assuming a wrong cosmology would create apparent errors in the scaling of angular sizes (which depends on angular diameter distance) compared to LoS sizes (which depends on the Hubble parameter), which break the isotropy. 21 cm surveys could provide a definitive detection of the AP effect (Ali et al., 2005; Barkana, 2006).

Generally, anisotropic terms are modeled following Eq.(3.51), where the AP effect distorts the shape and the normalization of the 21 cm power spectrum. This effect is generally undistinguishable from the one caused by redshift space distortions. The only way to disentangle the two effects is a measurement of the term $\mu^6 P_{\mu^6}(k)$ which is uniquely due to AP effect. This allows

for a measurement of

$$(1 + \alpha) = \frac{H\mathcal{D}(\text{Assumed Cosmology})}{H\mathcal{D}(\text{True Cosmology})}. \quad (3.53)$$

k and z dependence of the AP effect are well known and aid in separating it from noise and foregrounds. The AP effect is sensitive to background cosmology even out to high redshifts, where peculiar velocities are not sensitive to it (because the Universe is close to be Einstein-de Sitter, and $f = 1$). This has to be detected in the early stage of the EoR, when the effect of ionized bubbles is negligible (they would add non-trivial μ dependencies).

Its detection is difficult, because we need an accuracy of better than $\sim 5\%$ in $\log H$ and \mathcal{D} to make a better measurement respect to CMB (which is sensitive to $\Omega_\Lambda(z)$ contribution in \mathcal{D} integral too). At the moment, real-world challenges prevent any planned experiment, SKA included, from reaching this level.

3.4.3 Ionization Fluctuations

These source of fluctuations are originated by a balance between ionizing photons produced by highly clustered galaxies and recombinations in dense matter regions. We can assert that a region of gas is ionized if this region contains a sufficient number of galaxies able to ionize it. This means

$$m_{\text{ion}} \geq \varepsilon m_{\text{g}}, \quad (3.54)$$

where ε is the ionizing efficiency and m_{g} is the total galaxy mass in that region which produces enough ionizing photons. Translating this in collapse fraction, we get the equivalent expression

$$f_{\text{coll}} \geq f_{\text{x}} = \frac{1}{\varepsilon}, \quad (3.55)$$

where

$$f_{\text{coll}} = \int_{m_{\text{min}}}^{\infty} dm m n(m) = \text{erfc} \left[\frac{\delta_{\text{c}}(z)}{\sqrt{2}} \sigma(m_{\text{min}}) \right]. \quad (3.56)$$

Through the Press-Schechter mass function (Press & Schechter, 1974) $n(m)$, we can express this condition in mass overdensity terms,

$$\delta \geq \delta_{\text{x}}(m, z) = \delta_{\text{c}}(z) - K(\varepsilon) \sqrt{2(\sigma_{\text{min}}^2 - \sigma^2(m))}, \quad (3.57)$$

where

$$K(\varepsilon) = \text{erfc}^{-1} \left(1 - \frac{1}{\varepsilon} \right), \quad (3.58)$$

and

$$\sigma^2(m) = \int_0^{\infty} \frac{dk}{2\pi^2} k^2 P_{\text{lin}}(k) \left[\frac{3j_1(kR)}{kR} \right]^2 \quad (3.59)$$

is the variance of the density field smoothed in top-hat spheres of mass m . This condition to self-ionization allows for computing the probability distribution

of ionized regions, or bubble sizes $n_{\text{bub}}(m)$. This analytic distribution of bubble sizes is a good match to numerical EoR simulations. This single-point statistics is non-trivially connected to the power spectrum for ionization fluctuations, giving complicated expression for it, and some simplifying assumption is generally needed. For example Furlanetto et al. (2004) incorporates some simple ansatzes for P_{xx} and $P_{\text{x}\delta}$ based on the expected clustering properties of the bubbles.

3.4.4 Ly α Coupling Fluctuations

The temperature dependence of S_α is usually neglected, and this kind of fluctuations are sourced only by fluctuations in the flux, so that $\delta_\alpha = \delta_{J_\alpha}$. These are proportional to the matter overdensity through

$$\delta_\alpha = \delta_{J_\alpha} = W_\alpha(k)\delta, \quad (3.60)$$

with window function $W_{\alpha,\star}(k)$ defined for a gas element at z for stellar sources

$$W_{\alpha,\star}(k) = \frac{1}{J_{\alpha,\star}} \sum_{n=2}^{n_{\text{max}}} \int_z^{z_{\text{max}}(n)} dz' \frac{dJ_\alpha^{(n)}}{dz} \frac{D(z')}{D(z)} \left\{ [1 + b(z')] j_0(kr) - \frac{2}{3} j_2(kr) \right\}, \quad (3.61)$$

where $D(z)$ is the linear growth function, and $r = r(z, z')$ is the distance to the source. Each resonance contributes a differential comoving $dJ_\alpha^{(n)} dz'$. We can see from Eq.(3.61) that these density perturbations are sourced by three effects. First, the number of galaxies tracing matter is biased by a factor $[1 + b(z')\delta]$, so an overdense region emits more strongly. Second, photon trajectories are modified by gravitational lensing, and the effective area is increased by a factor $(1 + 2\delta/3)$. Finally, the peculiar velocity of overdense region of the gas will distort the observed frequency. This analytic prescription fits well with simulations, but more accurate treatments need to account for the full Ly α radiative transfer equation.

Here we assumed that UV photons redshift until they reach the line center of a Lyman series resonance, and only then they scatter. In reality the Universe is optically thick for all but the highest n transitions, so the photons will scatter in the wings of the line, and only a few mill make it to line center. This means that the former will travel a significantly reduced distance from the source before they scatter, reducing the size of the coupled region and steeping the flux profile around a source. This contribution will modify the $dJ_\alpha^{(n)}/dz'$ term.

Moreover, the photons in this approach propagate through the IGM with mean properties. There are density inhomogeneities and velocity flows that may produce more scattering and production of UV photons, leading to extra fluctuations.

3.4.5 Temperature and Ionization Fluctuations from X-rays

The effective value of W_α Eq.(3.60) has to consider effective fluctuations in J_X , so that

$$W_\alpha = \sum_i W_{\alpha,i} \frac{J_{\alpha,i}}{J_\alpha}. \quad (3.62)$$

In general X-ray photoionization depends on photon energy $\sigma \sim E^{-3}$, making the IGM optically thick for soft X-rays ($E \sim 20$ eV), while is optically thin for hard X-rays ($E \gtrsim 1$ KeV). Differently from Ly α flux, X-ray heating is a continuous process and depends on its past history. Usually X-ray fluctuations are described as produced by clustering of X-ray sources, since it is not clear yet how to calculate Poisson fluctuations in these sources.

With these assumption, a form for W_X can be computed by perturbing equations (3.31) for the evolution of δ_T and δ_e (the fractional fluctuation in x_e which is related to the neutral fraction by $\delta_x = -x_e/(1 - x_e)\delta_e$). For details consult Pritchard & Loeb (2012). These equations are completed using the Compton scattering contribution and introducing the ionization or heating rate per baryon $\Lambda = \epsilon/n$. For X-rays the rate of heating, ionization, and production of Ly α photons differ only by constant multiplicative factors (neglecting small fluctuations in x_e and focusing on X-ray energies $E \gtrsim 100$ eV). This means

$$\delta_{\Lambda_{\text{ion}}} = \delta_{\Lambda_{\text{heat}}} = \delta_{\Lambda_\alpha} = \delta_{\Lambda_X} = W_X(k)\delta, \quad (3.63)$$

and one finds that

$$W_X(k) = \frac{1}{\bar{\Lambda}_X} \int_{E_{\text{th}}}^{\infty} dE \int_z^{z^*} dz' \frac{d\Lambda_X(E)}{dz} \frac{D(z')}{D(z)} \left\{ [1 + b(z')] j_0(kr) - \frac{2}{3} j_2(kr) \right\}, \quad (3.64)$$

where the contribution to energy deposition rate by X-rays of energy E emitted with energy E' from between redshifts z' and $z' + dz'$ is

$$\frac{d\Lambda_X(E)}{dz} = \frac{4\pi}{h} \sigma_{\nu}(E) \frac{dJ_X(E, z)}{dz'} (E - E_{\text{th}}), \quad (3.65)$$

where $\sigma_{\nu}(E)$ is the cross-section for photoionization, E_{th} is the minimum energy threshold needed for photoionization, and $\bar{\Lambda}_X$ is a normalization obtained by performing the energy and redshift integrals. Respect to Eq.(3.61), we are not summing over discrete levels, but we integrate over X-ray energies. $W_X(k)$ acts as a mask: on scales smaller than 100 Mpc its shape depends on the details of X-ray source spectrum and the heating cross-section, while on larger scales these details are unresolved, so it traces density fluctuations. Moreover, there should be a dependency on X-rays coming from more distant sources than the ones at the redshifts considered in the integration.

The solution to perturbed equations are written as $\delta_T = g_T(k, z)\delta$, $\delta_e = g_e(k, z)\delta$, $\delta_\alpha = W_\alpha(k, z)\delta$, and $\delta_{\Lambda_X} = W_X(k, z)\delta$. Since these quantities are not independent of scale, the equations need to be solved for each mode. Note that we neglected the scale dependence induced by coupling to CMB. In MDE, one finds equations for the evolution of $g_T(z)$ and $g_e(z)$. The first term is usually linked to the adiabatic index γ_a of a gas through $g_T = \gamma_a - 1$. Adiabatic expansion and cooling make $g_T \rightarrow 2/3$, but at high z Compton heating is effective and leads the gas to isothermality $g_T \rightarrow 0$. At lower z , X-ray heating is insignificant and temperature fluctuations are dominated by spatial variation in the heating rate $g_T \rightarrow W_X$, linking temperature anisotropies to clustered source X-ray emission. When the heating rate is uniform $W_X \approx 0$, and $g_T \rightarrow 0$.

On the other hand $g_x \rightarrow -1$ at high redshift, because of highly dense IGM and the ionization fraction is dominated by the recombination rate. As IGM density gets low, $g_x \rightarrow 0$. When ionization becomes important, $g_x \rightarrow W_X$.

3.4.6 Full Power Spectrum Evolution and other source of fluctuations

The contributions of every single term we explored in the last sections are summed to produce the full 21 cm three-dimensional power spectrum, which is displayed on Figure 3.4 as a function of z at fixed k -values (Pritchard & Loeb, 2008).

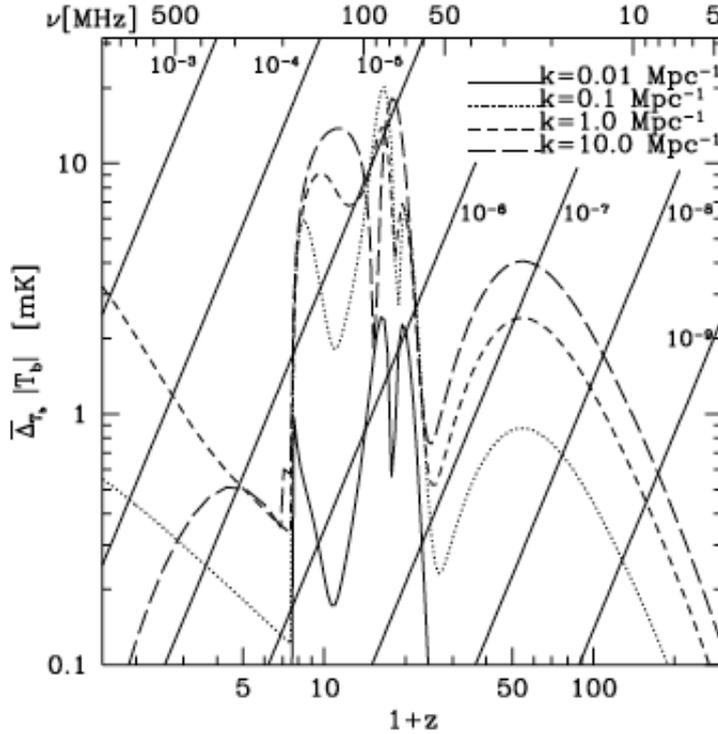


Figure 3.4: Evolution of the 21 cm brightness temperature fluctuations as a function of z and for fixed $k = 0.01, 0.1, 1, 10 \text{ Mpc}^{-1}$. The diagonal lines show the foreground temperature $T_{\text{fg}}(z)$ scaled by a factor ϵ which goes from 10^{-9} to 10^{-3} , indicating the required level of foreground needed to detect the 21 cm signal. Taken from Pritchard & Loeb (2008).

We can clearly see 4 different epochs. At early times $z \gtrsim 30$ before star formation, the power spectrum rises a peak around $z \approx 50$ and then drops off, as the 21 cm power spectrum tracks the density field modulated by the mean brightness fluctuation temperature. Once first stars are formed, coupling and temperature fluctuations become important. Then ionization fluctuations become important, leading to a decay at the end of EoR. After that, a weaker signal arises from the remaining HI in dense clumps found within collapsed structures. The foreground lines are showed to give an idea of the difficulty of removing them and of the required instrumental sensitivity for a detection.

In this discussion we have neglected other source of fluctuations coming from non-linear growth of structures, like the ones that can be originated from dense HI clumps which are collapsed but have not enough mass to give birth to stars. Such minihaloes should be abundant in early Universe and should contribute in the collision term, due to high density.

Another contribution could be given by the infalling of baryons into dark matter overdensities. Their relative velocity exceeds the local speed of sound generating supersonic flows. These can suppress the formation of the first IGM clouds by preventing the baryons from collapsing into dark matter haloes with low escape velocities and might be important for earliest phases of EoR, for example delaying the onset of Ly α coupling.

In the local Universe, shocks are known to be an important IGM heating mechanism. If there are magnetic fields, these shocks can accelerate charged particles, generating photons through inverse Compton scattering of CMB photons. This energy range is extended from radio to X-rays and large scale shocks are significant only at $z \lesssim 20$ and could have an important role in ionizing the IGM.

Finally, the diffuse radio background generated by bright sources might contaminate 21 cm flux, albeit the ambient 21 cm radiation field is mainly dominated by CMB $T_\gamma = T_{\text{CMB}}$. Fluctuations in this radio background would come from clustering of radio sources, and are important only when Ly α and collisional coupling are not important, so that the spin temperature relaxes to the background radio temperature T_γ .

3.4.7 Simulations: State of the Art

Apart from the analytic treatment presented here, numerical and semi-numerical techniques are needed to give predictions at a more detailed level, as stated in Trac & Gnedin (2011).

The spectrum of ionized fluctuations depends primarily on a single parameter, the ionized fraction x_{H} . Once this is fixed, the ionization pattern can be computed by filtering the density field on progressively smaller scales, and asking if a region can self-ionize itself. These regions will contribute to the photon counting, and are the main basis of codes such 21cmFAST³ (Mesinger et al., 2011), or SIMFAST21⁴ (Santos et al., 2010). These codes give a 21 cm signal with reasonable accuracy. Adding fluctuations in Ly α and temperature requires FFT convolution techniques. BEARS (Thomas et al., 2009) is another semi-numerical code based on painting spherically symmetric ionization, heating or coupling profiles from a library of 1D radiative transfer equations.

Fully numerical simulations are the best option for 21 cm studies, but require a simulation volume and dynamic range which are difficult to reach. These can take into account hydrodynamics or not. Those that do not take it into account are essentially dark matter N-body simulations which add a baryon component in a subsequent step. Then they apply radiative transfer to calculate the evolution of ionized bubbles. Some code can be found in McQuinn et al. (2007); Partl et al. (2011). Hydrodynamical simulations like the one performed by Trac et al. (2008) require a large amount of simulated volume and are often cut to an unrepresentative cosmic volume. On the other hand these codes allow for a proper study of dark matter, baryons, and bubbles evolution.

³<http://homepage.sns.it/mesinger/Download.html>

⁴<http://www.simfast21.org/>

The assumption $T_S \gg T_{\text{CMB}}$ is largely used in the literature, but analytic T_S variations have to be verified numerically. Such kind of codes (Baek et al., 2009, 2010) have to keep track of the radiative transfer in frequency bins and are numerically expensive.

3.4.8 Higher Order Statistics

The 21 cm brightness temperature distribution is highly non-Gaussian, because of the presence of ionized bubbles. So, high order statistics should contain information about sizes and topology of these bubbles. The challenge is to develop statistics matched to that form of non-Gaussianity, which is an unresolved problem. A source of skewness is the 1 point probability distribution function (Furlanetto & Loeb, 2004), which is important as reionization leads to many pixels with zero signal. The number of haloes in a connected surface might be determined by the Euler characteristic (Friedrich et al., 2011). These non-Gaussianities might also modify the shape of the power spectrum (Joudaki et al., 2011), giving birth to non-trivial connected terms. Of course there are other possibilities including bispectrum, wavelets and threshold statistics (Lee & Spergel, 2011) to explore.

Just like the CMB case, we could also hope to measure primordial non-Gaussianities from the 21 cm power spectrum. We can characterise the primordial non-Gaussianity with the parameter f_{NL} , defined by assuming a quadratic correction to the Gaussian inflaton potential, *i.e.* $\phi = \phi_G + f_{\text{NL}}\phi_G^2$. This sources a non-null bispectrum signal, the Fourier transform of the 3 point correlation function. A measurement of f_{NL} would effectively distinguish among different inflation models, but from CMB measurements (Planck Collaboration et al., 2015), f_{NL} is expected to be small, namely $f_{\text{NL}} = 0.8 \pm 5.0$ for the local bispectrum. The only hope to detect it with 21 cm observation is to measure 21 cm signal from $z > 30$, since these surveys can probe very large volumes (Cooray, 2006; Pillepich et al., 2007) and the signal is less contaminated from other sources of three-point correlation functions.

3.5 Constraining Cosmology with 21 cm Radiation

As already stated in Section 3.3 and can be seen in Figure 1.1, 21 cm survey can probe an unpreviously explored cosmological volume. CMB probes only a thin shell at $z \sim 1100$, and current surveys, such as the Sloan Digital Sky Survey (SDSS), can map only a small volume nearby. HI is potentially capable of mapping most of our horizon volume. To have a stronger impression about this issue, we can look at the comoving volume in function of redshift in Figure 3.5. With detecting the 21 cm signal at just a redshift range $z \sim 1 - 3$, the comoving volume explored is two orders of magnitude bigger than the one explored by SDSS. This is of crucial importance for cosmological parameters estimation: the constraints depend on the number of independent Fourier mode available from observations, and, hence, from the observed volume as $V^{-1/2}$ (Loeb & Wyithe, 2008).

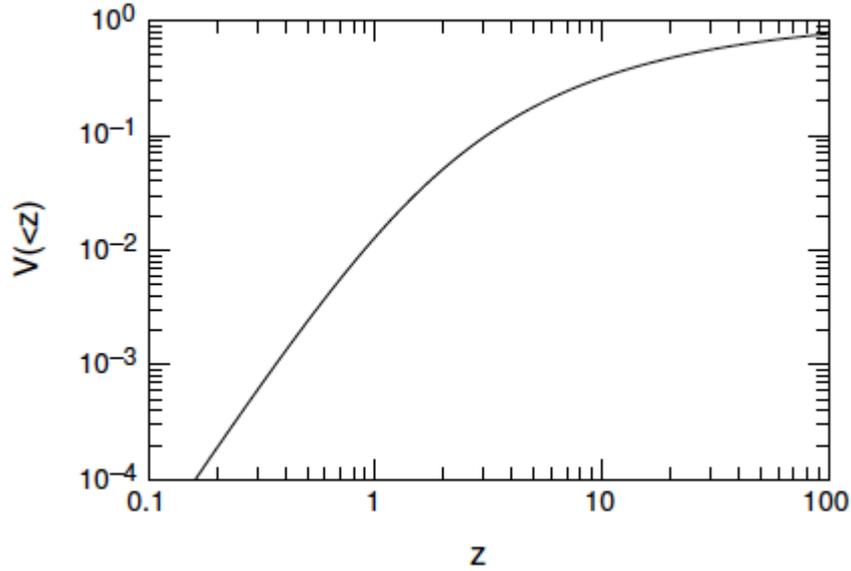


Figure 3.5: The available comoving volume in function of redshift z . Taken from Loeb & Wyithe (2008).

Mao et al. (2008) studied how accurately cosmological parameters can be measured, depending on factors such as ionization power modeling, reionization history, redshift range, experimental noise and configuration, and astrophysical foregrounds. Varying the assumptions on these, for example assuming various collecting areas or knowledge on the various 21 cm power spectrum Eq. (3.51) contributions, they constructed three models, a Pessimistic, a Middle, and an Optimistic case. They found that the constraints on Ω_k and Δm_ν for example, varying the ionization power modeling from the Pessimistic to the Optimistic case, improve by a factor 6 – 15. In particular, the assumptions can be ordered by importance as ionization power spectrum modeling \gg foregrounds \sim redshift ranges \sim array layout $>$ collecting area \sim system temperature \sim observational time \sim small-scales cutoff for non-linearities \sim non-Gaussianity. On the other hand, the pessimistic approach generated by marginalizing over our ignorance over the reionization history destroys too much cosmological information. In practice the truth will be in the middle case, but EoR models need to be better constrained. In principle an experiment like the SKA could improve the sensitivity of the Planck satellite by two orders of magnitude in best cases, and detect at 4σ the running of the spectral index to test inflation models.

Ideally a 21 cm tomography of the dark ages would provide huge improvements, since the physics in this phase is linear and well understood, and so the astrophysics can be easily distinguished and separated from the cosmology, for example using redshift space distortions as seen in Section 3.4.1. In reality low-frequency foregrounds pose a serious barrier, overwhelming the cosmological signal by several orders of magnitude. The required sensitivity can be reached with very large interferometer. For this reason the optimal experimental environment could be provided by the nearly-absent lunar atmo-

sphere (Jester & Falcke, 2009). Such an experiment would be even able to test the scaling of the fine-structure constant.

Gravitational lensing of 21 cm radiation could also improve the current estimates on cosmological parameters, as shown by Metcalf & White (2009). This topic will be investigated more accurately at the end of the next chapter.

Chapter 4

21 cm Radiation Weak Lensing Reconstruction

The 21 cm brightness temperature fluctuation radiation field can be seen as a diffuse background radiation which undergoes the same secondary fluctuations studied in the CMB. One of these effects, which can transfer power into the 21 cm spectrum, is the weak gravitational lensing.

The 21 cm emission/absorption has two major advantages over the CMB as a background source for lensing studies. Since lensing conserves surface brightness, it can only redistribute structure that already exists in the source. The CMB has very little structure on the angular scales where lensing is significant (< 1 arcmin) so that lensing effects are very weak, due to Silk damping. The second advantage is that the CMB provides only one temperature field on the sky while the 21 cm emission/absorption provides many, all of which are lensed by the same foreground mass distribution. Although the CMB comes from a higher redshift, this is a relatively minor advantage since most of the structure detected by lensing is at much smaller redshift than either source.

Zahn & Zaldarriaga (2006); Metcalf & White (2009) had shown that if the EoR is at redshift $z \sim 8$ or later, a large radio array such as SKA could provide a 3D tomography for the matter density, measure the lensing convergence power spectrum, and constrain the standard cosmological parameters. The authors extended the Fourier-space quadratic estimator technique, which was first developed by (Hu & Okamoto, 2002) for CMB lensing observations to three dimensional observables, *i.e.* the 21 cm intensity field.

Gravitational lensing moves temperature points in 21 cm emission and absorption maps, conserving the surface brightness. If the background radiation field is smooth, then lensing has no effect on it. The scales on which 21 cm power spectrum changes slope are the ones useful to probe lensing by taking advantage of the uncorrelation among different frequency maps of the fluctuation source field, while the lensing source is assumed to be always the same. Then the maps can be combined to average out the intrinsic temperature fluctuations, because lensing will induce non-null correlations between temperature field gradients and mass distribution.

Moreover, if the observed fraction of the sky is big enough, 21 cm lensing allows for competitive estimates of cosmological parameters (Metcalf &

White, 2009; McQuinn et al., 2006) through convergence auto-correlations or correlations with a foreground source. Using the gravitational lensing of galaxies in different source redshifts as a background and correlating them to a given lens population as a foreground provides an excellent way to test the Λ CDM paradigm (Pourtsidou et al., 2015). In fact, if the foreground and background distributions are narrow enough in redshift, the cross-correlation between the foreground galaxy density and the background lensing convergence from 21 cm can be measured from different bins. This is a purely geometric quantity which depends on the ratio of comoving radial distances and, hence, depends on cosmology and in particular on Dark Energy equation of state.

In the next section we will describe the simulation we set to study these issues, exploring various SKA configurations and observation frequencies. These simulations will include a discrete grid formalism which takes into account effects such as imperfect foreground subtraction or incomplete visibility coverage which can not be treated analytically. The aim of this work is to provide plausible simulations for high-fidelity images of the underlying mass distribution, as predicted by Metcalf & White (2007), and this will depend on the nature of 21 cm source signal, foreground subtraction technique, and telescope design. This would provide the opportunity of correlating visible objects with mass.

These studies can be extended, as we will see in Chapter 6, for 21 cm observations from redshifts after reionization, when the average HI density in the universe is much smaller. Pourtsidou & Metcalf (2014) extended the aforementioned studies to redshifts after reionization, but before those probed by galaxy surveys in the visible bands and showed that lensing can be measured using the HI intensity mapping technique.

In this chapter we will make use of the relations found in Chapters 2 and 3 to derive all the tools we adopt to get our results in Chapter 5.

4.1 Lensing Preliminaries

In this section we will resume what we need to describe 21 cm radiation weak lensing by LSS. The 21 cm radiation emitted from sources at a redshift z_s is lensed by the matter distribution lying between us and the emission region. Gravitational lensing will shift the observed position of a point on the sky without changing the surface brightness. If the lensing is weak compared to structure in the source, the observed temperature can be expressed as a Taylor expansion of the unlensed temperature:

$$\tilde{T}(\boldsymbol{\theta}, \nu) = T(\boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}), \nu) \simeq T(\boldsymbol{\theta}, \nu) - \boldsymbol{\alpha}(\boldsymbol{\theta}) \cdot \nabla_{\boldsymbol{\theta}} T(\boldsymbol{\theta}, \nu) + \dots \quad (4.1)$$

(Lewis & Challinor, 2006), where $\boldsymbol{\alpha}(\boldsymbol{\theta})$ is the deflection caused by lensing (with $\boldsymbol{\theta}$ the true position of the source) and dots denote higher-order terms in the expansion. Expansion Eq. (4.1) is trivially valid in CMB case because of the smallness of temperature gradients on medium scales and Silk damping on smaller scales. This is true also in 21 cm case, where temperature gradients can be large, but the deflections (or deflection gradients) are small compared

to them on all scales of interest. The deflection field $\alpha(\theta)$ is related to the 2D projected lensing potential via $\nabla\Phi = -\alpha(\theta)$, in the weak lensing limit. The lensing potential comes from the integration over the redshift direction of the full 3D gravitational potential Eq. (2.97), namely

$$\Phi = \frac{2}{c^2} \int_0^{z_s} dz \frac{\mathcal{D}(z)\mathcal{D}(z_s - z)}{\mathcal{D}(z_s)} \phi[\mathcal{D}(z)\theta(z), z], \quad (4.2)$$

where $\mathcal{D}(z)$ is the comoving angular diameter distance at redshift z . Taking the observed lensed position to be θ and the unlensed one to be ξ , the shear $\gamma_{1,2}$ and the convergence κ are related to the gravitational potential by the Jacobian matrix Eq. (2.51),

$$\begin{aligned} \mathcal{J}(\theta, z_s) = \frac{\partial \xi}{\partial \theta} &= \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} \\ &= \begin{pmatrix} 1 - \Phi_{,11} & \Phi_{,12} \\ \Phi_{,12} & 1 - \Phi_{,22} \end{pmatrix}, \end{aligned} \quad (4.3)$$

where we have neglected any rotational variable in off-diagonal elements and the subscripts 1 and 2 stand for the derivative operation with respect to the two transverse coordinates of the lensing potential.

The convergence field is related to the potential - or, equivalently, the deflection field - via the Poisson equation $\kappa = -\nabla^2\Phi/2 = -\nabla \cdot \alpha/2$. We have seen in Section 2.4 that using the Limber approximation for small angles, it is possible to define the power spectrum of the deflection (or convergence) field. This will be related to the 3D density fluctuations power spectrum through Eq. (2.122)

$$C_L^{\alpha\alpha} = \frac{9\Omega_m^2 H_0^3}{L(L+1)c^3} \int_0^{z_s} dz \frac{\mathcal{W}^2(z)}{a^2(z)E(z)} P_\delta\left(k = \frac{L}{\mathcal{D}(z)}, z\right), \quad (4.4)$$

where $E(z) = H(z)/H_0$ and $\mathcal{W}(z) = 1 - [\mathcal{D}(z)/\mathcal{D}(z_s)]$. H_0 is the Hubble parameter today and Ω_m is the density of the matter in the Universe relative to the critical density. Throughout this work we adopt a standard Λ CDM cosmology with the Planck parameters set (Planck Collaboration et al., 2015), namely $\Omega_m = 0.3183$, $\Omega_\Lambda = 0.6817$, $\Omega_b = 0.049$, $h = H_0/100 = 0.67$ Km/s/Mpc, $\sigma_8 = 0.8347$, and $A = 1.2 \times 10^{-4}$.

4.2 21 cm Brightness Temperature Fluctuation Field

The brightness temperature for the 21 cm line is given by expliciting and slightly modifying Eq. (3.51), namely

$$\bar{T}(z) \simeq 26(1 + \delta_b)x_H \left(1 - \frac{T_{\text{CMB}}}{T_S}\right) \left(\frac{\Omega_b h^2}{0.022}\right) \left[\left(\frac{0.15}{\Omega_m h^2}\right) \left(\frac{1+z}{10}\right)\right]^{1/2} \text{ mK}, \quad (4.5)$$

(Furlanetto et al., 2006; Zahn & Zaldarriaga, 2006) where x_H is neutral hydrogen fraction, T_S is the 21 cm spin transition temperature, $T_{\text{CMB}} = 2.73(1+z)$ K

is the CMB temperature at redshift z , $\delta_b = (\rho_b - \bar{\rho}_b)/\bar{\rho}_b$ is the baryon density contrast measured in redshift space, and Ω_b is the average density of baryons today relative to the critical density. In the regimes of interest here, i.e. $z < 15$, $T_s \gg T_{CMB}$ so that there is no dependence on the CMB temperature. The ionization fraction and the density of HI will depend on the considered epoch, the ionization history and structure formation history, as seen in Chapter 3.

The brightness temperature will be represented in the simulation within a rectangular volume centred at a redshift z . The comoving, radial length of this volume is $\mathcal{L}(z, \Delta\nu)$ which depends on the bandwidth, $\Delta\nu$. We will make the approximation that angular distance to the simulation box $\mathcal{D}(z)$ is very large compared to $\mathcal{L}(z, \Delta\nu)$ so that the angular size of the front of the box is the same as the back of the box. With this and the flat-sky approximation for small patches of the sky, the 3D temperature field is represented in Fourier space by defining the wave vectors $\mathbf{k}_\perp = \mathbf{l}/\mathcal{D}(z)$ and $k_\parallel = 2\pi k_p/\mathcal{L}(z, \Delta\nu)$, where \mathbf{l} is the multipole vector, the Fourier space dual of the angle coordinate, and k_p is an integer which discretises the k_\parallel direction. The frequency band is broken up into many channels which can be interpreted as tangential slices. The Fourier dual of the radial distance is then the discrete values of k_\parallel or k_p . Homogeneity dictates that there will be no correlations between modes with different k_p .

We will take the simulation box to be square in the angular dimensions with the obvious extension to rectangular geometry. The angular area of the survey and box will be Ω_s . The number of grid points in each dimension on the sky will be N_\perp so that the total number of grid positions in each frequency channel is $N_s = N_\perp^2$. The angular resolution is $\Delta\theta$ and (m, n) are the pixel indices.

The conversion between radial distance and frequency is given by

$$dr = \frac{c}{H_0} \frac{dz}{\sqrt{\Omega_m(1+z)^3 + \Omega_K(1+z)^2 + \Omega_\Lambda}} \simeq \frac{c(1+z)^2}{v_{21}H_0\Omega_m^{1/2}} d\nu, \quad (4.6)$$

where Ω_K is the energy density parameter for curvature, Ω_Λ is the one for a cosmological constant, and the approximation made in the last passage holds at high redshifts when the Universe is matter dominated. We will assume this is true throughout this work. The rest frame frequency is $\nu_{21} = 1420.4$ MHz. With this the total depth of the box can be calculated,

$$\mathcal{L}(z, \Delta\nu) \simeq \frac{c(1+z)^2}{v_{21}H_0\Omega_m^{1/2}} \Delta\nu, \quad (4.7)$$

and the frequency of each channel, ν , can be converted into radial distances r_ν within the box.

Following the notation introduced in Appendix D, the Discrete Fourier Transform (DFT) of the temperature intensity field is then

$$T_{\mathbf{l}, k_p} = \frac{\Omega_s}{N_s N_\nu} \sum_{m, n, r} \exp \left[-2\pi i \left\{ \frac{1}{N_s} \mathbf{l} \cdot (m, n) + \frac{1}{\mathcal{L}} r_\nu k_p \right\} \right] T_{m, n, \nu} \quad (4.8)$$

where N_ν is the number of channels within the band that is used.

From Eq. (4.8), we can derive the angular power spectrum of the 21 cm temperature field, C_{l,k_p} , defined by

$$\langle T_{l,k_p} T_{l',k'_p}^* \rangle = \Omega_s C_{l,k_p} \delta_{l,l'}^K \delta_{k_{\parallel},k'_{\parallel}}^K \quad (4.9)$$

(see Appendix D for a detailed derivation of this relation). Throughout this work the averaging operation denoted as $\langle \dots \rangle$ is performed over 21 cm intensity field realisations. The angular power spectrum is related to the discrete temperature field power spectrum P_k via

$$P_k = \frac{P_{\Delta T}(k)}{V_s} = \frac{P_{\Delta T} \left[\sqrt{(l/D)^2 + (2\pi k_p/\mathcal{L})^2} \right]}{\Omega_s \mathcal{D}^2 \mathcal{L}} = \frac{C_{l,k_p}}{\Omega_s}. \quad (4.10)$$

(Zahn & Zaldarriaga, 2006).

For our first set of simulations we will adopt a simple model for the brightness temperature distribution which has been used before and can be compared to analytic results. First we will consider a time before ionization when hydrogen is completely neutral ($x_H = 1$ in expression (4.5)). The brightness temperature is then only dependent on the density distribution of hydrogen. To model this we will make the assumption that the baryons are not yet significantly biased with respect to the mass so that their power spectrum in redshift space is given by

$$P_{\Delta T}(k) = \bar{T}^2(z) (1 + f\mu_k^2)^2 P_{\delta}(k), \quad (4.11)$$

where $P_{\delta}(k)$ is the dark matter power spectrum. Lu & Pen (2008) remark that this simplification is not valid on small scales. In this work we are mainly interested in recovering a high scale signal while the small scale one will be dominated by noise, so this approximation holds for our purposes.

We have included the redshift space distortion term in which we applied Eq. (3.49), and $f = d \ln D / d \ln a \simeq \Omega_m(z)^{0.55}$ with D the linear growth rate. The cosine of the angle formed by the parallel component of the wavevector \mathbf{k} and the wavevector itself is denoted $\mu_k = k_{\parallel}/k$. We will also assume that these fluctuations can be modeled with a Gaussian random field. We will use a fiducial source redshift of $z_s = 8$ for single band results and bandwidths of $\Delta\nu = 5$ MHz ($\Delta z = 0.286$) and $\Delta\nu = 8$ MHz ($\Delta z = 0.456$). The explored range when different redshift bands are combined is $z \approx 7 - 11.6$. In Figure 4.1 the power spectra (4.11) are shown for different k_p . Depending on the noise model (which will be specified in the next sections), modes beyond some k_p^{\max} are dominated by noise and thus not useful for detecting lensing.

At the high redshifts and the resolution considered in this study, redshift space distortions can be modeled assuming $f \approx 1$. As discussed in Mao et al. (2012), redshift distortions have non negligible effects on 21 cm power spectrum.

We also assume that no reionization has occurred. The actual temperature distribution is likely to be considerably more complicated because of segregation between baryons and dark matter, and non-uniform ionization. In fact,

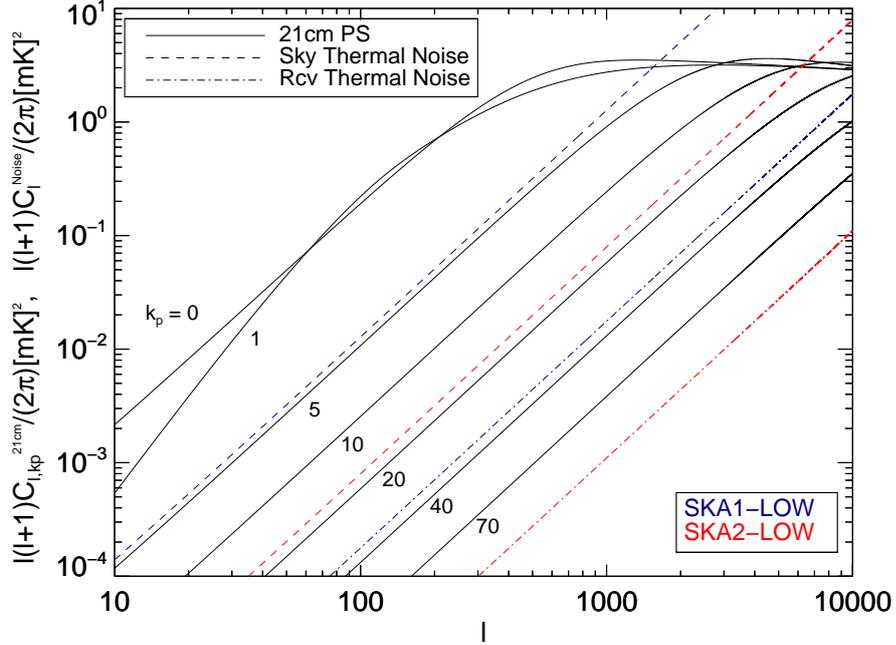


Figure 4.1: The 21 cm angular power spectra C_{ℓ,k_p} for several values of k_p at $z_s = 8$. The thick dashed lines are the sky noise power spectra, while the dashed-point lines are the receiver noise power spectra. The latter are produced assuming the SKA1-Low (blue) and SKA2-Low (red) settings described in Section 5.4.

EoR could be happened for a considerable redshift range, so bubbles of HI regions are likely to be formed. These grow, evolve, and merge until a uniform reionized region is filled up and reionization can be considered complete. Albeit this will improve 21 cm detection because of higher temperature brightness contrast, the higher moments terms in the temperature brightness fluctuation can be non-negligible. The form of our input power spectrum Eq. (4.11) should be modified including these contributions (Metcalf & White, 2009).

These are cases our code is designed to handle, but will be investigated in future works when it is combined with more detailed reionization simulations.

4.3 Lensing Reconstruction in Fourier Space

If the bandwidth is small and the redshift is relatively moderately high, to a good approximation the matter within the band does not contributing significantly to the lensing of that band, i.e. there is no self-lensing. In this case, as shown in Appendix D.3 identifying the generic intensity field I_{l,k_p} with the 21 cm brightness temperature field T_{l,k_p} , the correlation between brightness temperature modes can be written derived from equation (4.1),

$$\left\langle \tilde{T}_{l,k_p} \tilde{T}_{l-L,k'_p}^* \right\rangle = \mathbf{L} \cdot \left[l C_{l,k_p} + (L-l) C_{|l-L|,k'_p} \right] \Phi(\mathbf{L}) \delta_{k_p k'_p}^K. \quad (4.12)$$

(for $L \neq 0$) (Hu & Okamoto, 2002). We can then define a discrete quadratic estimator for the gravitational potential of the form

$$\hat{\Phi}_L = \sum_{l, k_p} f(l, L, k_p) \tilde{T}_{l, k_p} \tilde{T}_{l-L, k_p}^*, \quad (4.13)$$

in which the form of the filter $f(l, L, k_p)$ depends on the kind of source we are analysing and on its statistics. The important effect of lensing is to induce correlations between different modes that would otherwise be uncorrelated because of homogeneity. In the case of a Gaussian temperature field, an optimal kernel can be derived by requiring $\langle \hat{\Phi}(L) \rangle = \Phi(L)$, and minimising its variance. The resulting estimator is

$$\hat{\Phi}_L = \frac{N_L^{\hat{\Phi}}}{2\Omega_s} \sum_{l, k_p} \left[\frac{\mathbf{L} \cdot \mathbf{l} C_{l, k_p} + \mathbf{L} \cdot (\mathbf{L} - \mathbf{l}) C_{l-L, k_p}}{C_{l, k_p}^T C_{l-L, k_p}^T} \right] \tilde{T}_{l, k_p} \tilde{T}_{l-L, k_p}^* \quad (4.14)$$

(Zahn & Zaldarriaga, 2006). The variance of this estimator is

$$\langle \hat{\Phi}_L \hat{\Phi}_L^* \rangle = \Omega_s (N_L^{\hat{\Phi}} + C_L^{\Phi\Phi}), \quad (4.15)$$

with $N_L^{\hat{\Phi}}$ being the reconstruction noise. For the optimal estimator with a Gaussian source field this is

$$N_L^{\hat{\Phi}} = \left\{ \frac{1}{2\Omega_s} \sum_{l, k_p} \frac{[\mathbf{L} \cdot \mathbf{l} C_{l, k_p} + \mathbf{L} \cdot (\mathbf{L} - \mathbf{l}) C_{l-L, k_p}]^2}{C_{l, k_p}^T C_{l-L, k_p}^T} \right\}^{-1}, \quad (4.16)$$

where $C_{l, k_p}^T = C_{l, k_p} + N_l^{\text{Sky}} + N_l^{\text{Rcv}}$ is the total power spectrum that includes the sky and receiver noises. Note that the presence of the factor 2 ensures the validity of the invariance $\mathbf{l} \rightarrow \mathbf{l} - \mathbf{L}$ for every independent k_p slice.

One can note in Appendix D.3 that, in deriving this expression, and the optimal form of the kernel, the fourth order correlations of the field are required. These are easily found for a Gaussian field, but for a more complicated and realistic source field the noise will need to be found numerically through simulations like the ones described in this work.

Expressions for estimator and noise, for both the deflection and convergence fields, are trivially found by remembering that in Fourier space $\hat{\alpha}_L = i\mathbf{L}\hat{\Phi}_L$ and $\hat{\kappa}_L = -(L^2/2)\hat{\Phi}_L$. Moreover, $N_L^{\hat{\kappa}} = (L^4/4)N_L^{\hat{\Phi}} = (L^2/4)N_L^{\hat{\alpha}}$. These results can be linked to the continuous result by making the substitution $\Omega_s \rightarrow (2\pi)^2\delta(0)$.

Note that equation (4.16) is of the form $N_L^{\hat{\Phi}} = 1 / \sum_{k_p} [N_{L, k_p}^{\hat{\Phi}}]^{-1}$, a result of the different k_p modes being uncorrelated. Adding more k_p modes reduces the total noise, but, as pointed out in Zahn & Zaldarriaga (2006), only the first 20 k_p modes contribute to lensing reconstruction. This is caused by the monotonically decreasing behaviour of C_{l, k_p} on all scales of interest as shown in Figure 4.1. For high values of k_p the signal is well below the thermal noise level so these modes do not contribute to the estimator. Hence, the estimator noise

saturates at $k_p^{\max} \sim 20 - 25$ for $z_s = 8$ and $\Delta\nu = 5$ MHz in this case, and this effect will be clearly demonstrated in Section 5.5 for our particular model.

A non uniform EoR process in the considered redshift range can cause a bias in the estimator since the Gaussianity assumption is broken, especially at small scales. Zahn & Zaldarriaga (2006) already pointed out that the noise will increase at small scales because of correlation induced by connected four-points function and because of 21 cm fluctuation level decreasing. Further investigations on more realistic reionization models are needed, and we plan to accomplish these issues in future works, when more detailed reionization models will be available. For the moment we will assume that these features will show for higher resolutions and redshifts than the one considered in this work, and, as already explored by Kovetz & Kamionkowski (2013), there is no reason to include higher moments in our estimator.

4.3.1 Faster Lensing Estimator

Estimator (4.14) is computationally slow to calculate. As shown in Anderes (2013), Lewis & Challinor (2006) and Carvalho & Moodley (2010) for the analogous 2D CMB case, the estimator can be interpreted as a convolution in Fourier space which is equivalent to a real space product. Doing the product in real space allows one to take advantage of Fourier Transforms (FFTs) methods¹ such as FFTW² to do the sums. Extending their derivation for k_p images, we have:

$$\hat{\Phi}_L = -\frac{N_L^{\hat{\Phi}}}{\Omega_s} (iL) \cdot \sum_{k_p} \left[\sum_{\theta} e^{-iL \cdot \theta} F_{\theta} \nabla G_{\theta} \right]_{k_p} = -\frac{N_L^{\hat{\Phi}}}{\Omega_s} (iL) \cdot \sum_{k_p} H_{L,k_p}, \quad (4.17)$$

where H is defined here and F and G are 2D angular space maps of the input 21 cm intensity temperature field, defined by applying the following high-pass filters in Fourier space

$$F_{l,k_p} = \frac{\tilde{T}_{l,k_p}}{C_{l,k_p}^T}, \quad G_{l,k_p} = \frac{C_{l,k_p} \tilde{T}_{l,k_p}}{C_{l,k_p}^T}. \quad (4.18)$$

In this way every k_p contribution to the estimator is computed individually, by filtering the input fields and multiplying their inverse DFTs in real space.

As pointed out by Lewis & Challinor (2006), seen from this point of view, the estimator measures the correlations in the product of two Wiener filtered fields, the temperature gradient field, $\nabla G(\theta)$, and the small-scale weighted field $F(\theta)$.

¹This will be true in full-sky representation too, since the azimuthal integrals can be treated similarly.

²<http://www.fftw.org/>. Notice that our DFT convention agrees with normalized FFTW one except for an Ω_s factor.

4.4 Including the Beam

In order to simulate more accurately the observational effects we included a beam in our simulation and estimator. The beam smooths the signal coming from scales that are small with respect to the beam resolution scale. This scale corresponds to a cutoff multipole L_{cut} that plays the same role as L_{max} in the non-beamed estimator, and determines the total level of estimator noise. The estimator will be insensitive to all modes $L > L_{\text{cut}}$, which will not be reconstructable.

The observed point will have a sky noise contribution n_x^{Sky} and a receiver noise contribution n_x^{Rcv} , so that

$$\tilde{\mathcal{T}}_x = \sum_{x'} W_{xx'} (\tilde{\mathcal{T}}_{x'} + n_{x'}^{\text{Sky}}) + n_x^{\text{Rcv}}, \quad (4.19)$$

with Fourier transform

$$\tilde{\mathcal{T}}_{l,k_p} = W_l (\tilde{\mathcal{T}}_{l,k_p} + n_l^{\text{Sky}}) + n_l^{\text{Rcv}}. \quad (4.20)$$

The beaming function W_l does not depend on frequency and it is assumed not to generate any spurious correlation among k_p modes. From this definition of the lensed and beamed temperature field, the discrete quadratic estimator can be found following the procedure outlined in Appendix D.4. We find

$$\hat{\phi}_L = \frac{\mathcal{N}_L^{\hat{\phi}}}{2\Omega_s} \sum_{l,k_p} \left\{ \frac{W_l W_{|l-L}^* [L \cdot l C_{l,k_p} + L \cdot (L-l) C_{|l-L|,k_p}]}{C_{l,k_p}^T C_{|l-L|,k_p}^T} \right\} \tilde{\mathcal{T}}_{l,k_p} \tilde{\mathcal{T}}_{l-L,k_p}^* \quad (4.21)$$

with $C_{l,k_p}^T = |W_l|^2 (C_{l,k_p} + N_{l,k_p}^{\text{Sky}}) + N_{l,k_p}^{\text{Rcv}}$. The estimator noise will consequently be modified into

$$\mathcal{N}_L^{\hat{\phi}} = \left\{ \frac{1}{2\Omega_s} \sum_{l,k_p} \frac{|W_l|^2 |W_{|l-L}|^2 [L \cdot l C_{l,k_p} + L \cdot (L-l) C_{|l-L|,k_p}]^2}{C_{l,k_p}^T C_{|l-L|,k_p}^T} \right\}^{-1}. \quad (4.22)$$

This beamed discrete estimator noise is easily computable in a reasonable amount of time by parallelizing the innermost sums in the latter equation. So, if we re-define our filters Eqs.(4.18) as

$$\mathcal{F}_{l,k_p} = \frac{W_l \tilde{\mathcal{T}}_{l,k_p}}{C_{l,k_p}^T}, \quad \mathcal{G}_{l,k_p} = \frac{W_l C_{l,k_p} \tilde{\mathcal{T}}_{l,k_p}}{C_{l,k_p}^T}, \quad (4.23)$$

we can find the beamed version of Eq. (4.17), namely

$$\hat{\phi}_L = -\frac{\mathcal{N}_L^{\hat{\phi}}}{\Omega_s} (iL) \cdot \sum_{k_p} \mathcal{H}_{L,k_p}, \quad (4.24)$$

where \mathcal{H}_{L,k_p} is again the Fourier transformed vectorial field formed by multiplying the inverse transformed Fourier \mathcal{F}_{l,k_p} with the inverse Fourier transformed gradient of \mathcal{G}_{l,k_p} . A detailed derivation is presented in Appendix D.5.

For multipoles $L > L_{\text{cut}}$, the estimator noise diverges because of the dominance of thermal noise at those scales and the smoothing of structure by the beam. Explicitly incorporating the beam allows us to avoid aliasing and pixelization effects. The value of L_{cut} will be specified in Section 5.3, and will depend on the observed redshift and on telescope design. A low L cutoff, L_{min} reflecting the finite field-of-view can also be incorporated into the beam. We choose here to allow the boundaries of the simulated maps implicitly impose this cutoff at $L_{\text{min}} = 2\pi/\theta^{\text{max}}$.

The overall estimator noise level will depend only on the chosen beam resolution and so uniquely on L_{cut} value. This level will be higher for stronger beam suppressions. Just like its analogous Eq. 4.16, the estimator noise level will be weakly sensitive to the resolution in Fourier space, namely L_{min} . A computation of the estimator reconstruction noise including the beam is shown in Figure 4.2 for $\Omega_s 5^\circ \times 5^\circ$, $z = 8$, $L_{\text{cut}} = 13237$ and noise model specified on Section 5.4, together with the single k_p contributions. We will often consider a $k_p^{\text{min}} \neq 0$, in order to take into account the foreground subtraction, as will be better explained on Section 5.5.4. Note that the signal-to-noise is not too much affected by removing the first k_p modes.

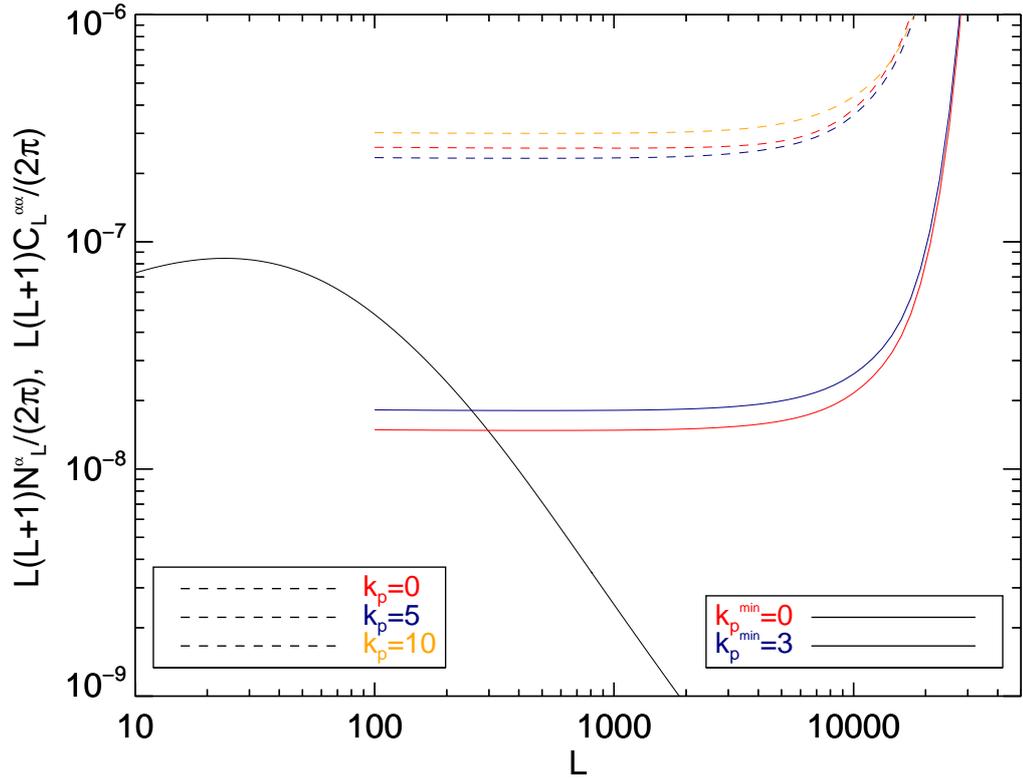


Figure 4.2: The estimator reconstruction noise including the beam function. The dashed lines correspond to single $k_p = 0, 5, 10$ modes contributions in red, blue, and orange respectively. The solid lines are for $k_p^{\text{min}} = 0, 3$ in red and blue, respectively. The assumed noise model is described in Section 5.4. We assumed $\Omega_s = 3.6^\circ \times 3.6^\circ$ at $z = 8$, with $L_{\text{cut}} = 13237$.

4.5 Measuring 21 cm Lensing Power Spectrum

A 21 cm lensing survey covering a large fraction of the sky would be able to measure the two-point statistics of the underlying lensing field averaged over a given observation redshift z . The statistical error in the deflection field power spectrum given by Eq. (4.4) is

$$\Delta C_L^{\alpha\alpha} = \sqrt{\frac{2}{(2L+1)f_{\text{sky}}\Delta L}} (C_L^{\alpha\alpha} + \mathcal{N}_L^{\hat{\alpha}}), \quad (4.25)$$

(Metcalf & White, 2009), where ΔL is the multipole binning, f_{sky} is the observed fraction of the sky, and $\mathcal{N}_L^{\hat{\alpha}}$ is the discrete estimator reconstruction noise related to Eq. (4.22) via $\mathcal{N}_L^{\hat{\alpha}} = L^2 \mathcal{N}_L^{\hat{\Phi}}$.

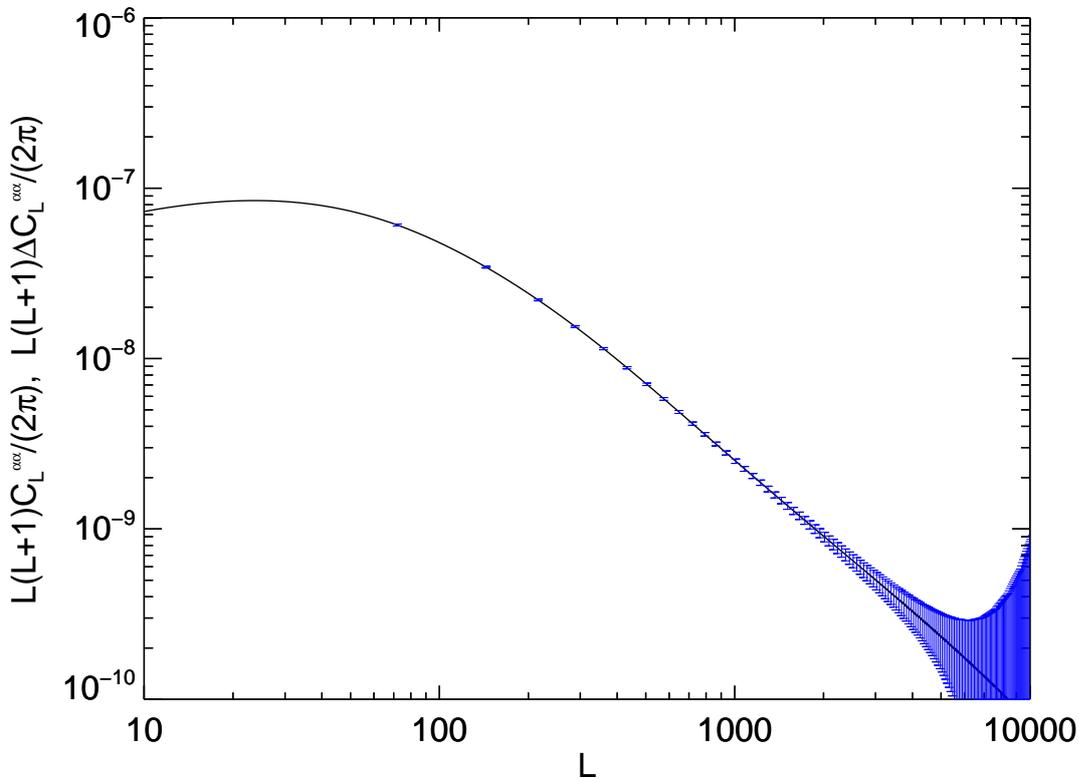


Figure 4.3: Deflection field power spectrum for $z = 8$ and corresponding measurement errors computed using SKA1-Low settings described in Sections 5.6 and 5.5.6.

In Figure 4.3 we show the deflection field power spectrum with measurement errors for a single bandwidth of $\Delta\nu = 5$ MHz centered around $z = 8$ and using the SKA1Low-R1 configuration which will be specified in Section 5.6.1. Note that we have adopted a large $f_{\text{sky}} = 0.6$ and we have used $\Delta L = 72$. The width ΔL defines the number of independent L -bands and is determined by the condition³ $\Delta L \geq 2\pi/\sqrt{\Omega_s}$ (Bowden et al., 2004). So if $f_{\text{sky}} = 0.6$, $\sqrt{\Omega_s} = \sqrt{f_{\text{sky}} \times 41253} \approx 160^\circ$, ΔL can be any number ≥ 2.25 .

³This holds for a square survey. If the survey is not square one needs to use the smallest angular dimension.

In reality the situation will be worse than the one pictured in Figure 4.3, since, as stated by Pourtsidou & Metcalf (2014), the fraction of the sky observed by an SKA-Low-like instrument will be very small ($\sim 10^{-4}$). More details on what SKA can measure at EoR redshifts will be given in Section 5.5.6, when we will show our reconstruction results and introduce our SKA-Low models to compute the discrete reconstruction noise.

4.6 Cosmological Parameter Estimation

In this work we are not interested in giving constraints on cosmological parameters to test the Λ CDM model. This can be done in future works. Here we introduce the possibility and the formalism to perform such estimates, since 21 cm lensing is in principle capable of providing competitive measurements without constructing density maps.

We assumed that the estimator reconstruction noise is, to a good approximation, statistically independent for different k_p modes (and also for different frequency bands) and for different L 's. In this case it is possible to define a data vector which is the difference of the estimated potential and the theoretical one (containing all the information about the global cosmological parameters). This is

$$D = \hat{\Phi} - \Phi, \quad (4.26)$$

with the components of Φ running over all measured positions and redshift components. The logarithm of the likelihood function is

$$\ln \mathcal{L} = -\frac{1}{2} D^\dagger N^{-1} D - \frac{1}{2} |N| - H, \quad (4.27)$$

where N is the noise covariance matrix, and H is an additional regularizing function which represents a prior distribution on the parameters. Usually the noise covariance matrix is not diagonal, in order to take into account possible correlations among frequency bands. The uncorrelation assumption here is translated in making the noise correlation matrix diagonal. Such assumption might be significantly broken for example by foreground subtraction or because of correlations between different L 's due to the finite beam of the telescope.

The maximum likelihood estimate for any parameter can be found by maximizing the likelihood function with respect to that parameter. The error of this estimation is evaluated through the Fisher matrix

$$F_{ij} = - \left\langle \frac{\partial^2 \mathcal{L}}{\partial p_i \partial p_j} \right\rangle, \quad (4.28)$$

so that the expected error in a given parameter p_a is $\sigma_a^2 \simeq (F^{-1})_{aa}$, *i.e.* when we marginalise over all the other parameters, or $(F_{aa})^{-1}$ when all other parameters are fixed and the error is unmarginalized. Eq. 4.28 can be non-diagonal if some parameter is degenerate, so correlated with each other. It is possible to find a

better error estimate considering the matrix that diagonalize F , $F = V^\dagger \lambda V$, which defines the linear combinations of the parameters $\hat{p} = Vp$ that are uncorrelated with variance λ_{aa}^{-1} . This can help in redefining the parameters space, in order to find the best one which describes the measured data.

Estimates of the cosmological parameters can be performed marginalizing the actual distribution of matter assuming a suitable prior. For example the components of Φ can be assumed to be normally distributed (Takada & Jain, 2004). In this case

$$H = \frac{1}{2} \Phi^\dagger C_{\Phi\Phi}^{-1} \Phi, \quad (4.29)$$

with $[C_{\Phi\Phi}]_{ij} = \langle \phi(\mathbf{L}, (k_p)_i) \phi^*(\mathbf{L}, (k_p)_j) \rangle$ and the likelihood function can be integrated over all components of \mathbf{K} . If k_p modes are uncorrelated, $C_{\Phi\Phi}$ is diagonal, namely $[C_{\Phi\Phi}]_{ij} = C_L^{\phi\phi} \delta_{ij}^K$. The likelihood function Eq. 4.27 can be written as

$$\ln \mathcal{L} = -\frac{1}{2} \hat{\Phi}^\dagger C^{-1} \hat{\Phi} - \frac{1}{2} |C|, \quad (4.30)$$

with

$$C = N + C_{\Phi\Phi}, \quad (4.31)$$

the sum of noise and potential covariance matrices. As pointed out by (Metcalf & White, 2009), Eq. 4.30 can be broken up into L -bins, each resolved region in Fourier space. This means that $\sim (2L + 1)f_{\text{sky}}$ modes are independently measured, with f_{sky} the observed fraction of the sky. The Fisher matrix is hence

$$F_{ab} = \frac{1}{2} \sum_{L=L_{\min}}^{L_{\max}} (2L + 1) f_{\text{sky}} \text{tr} \left[C^{-1} C_{,a} C^{-1} C_{,b} \right], \quad (4.32)$$

where commas indicate derivatives over the a, b parameter.

As written in Metcalf & White (2009), the same formalism can be used to do tomography, *i.e.* reconstruct the cosmic mass density distribution, by fixing the cosmological background and maximizing $\ln \mathcal{L}$ with respect to the pixelized foreground density. This is because $\Phi = G\delta$, where G is the matrix encoding all the cosmological information, while δ encloses the information about the angular and redshift position.

Metcalf & White (2009) found that an SKA-like experiment could constrain the dark matter and the shape of the matter power spectrum with high accuracy from $z \sim 0.5$ to $z \sim 7$. On the other hand, the dark energy parameters are less well constrained with respect to a galaxy survey, since Dark Energy is primarily felt at low redshifts. The best constraints come from combining a 21 cm survey with a galaxy survey, improving by more than an order of magnitude the constraints obtained considering galaxy surveys alone.

Chapter 5

Simulations for 21 cm Lensing Imaging at EoR redshifts

In this chapter we will describe the computational tools, the method, and the approximations used to perform lensing reconstruction on simulated 21 cm temperature maps for EoR redshifts ($z \sim 6.5 - 12$). We will use the lensing and the 21 cm formalism developed in the previous two chapters to use the quadratic estimator introduced in Section 4.3.1 and derived in Appendix D. We first describe how we generated the Gaussian random temperature fluctuation field, the lensing potential field and how we combine them to get the lensed temperature field. Then we explain how we have modelled the thermal noise components due to sky contamination and instrumentation. Finally we apply an explicit model for the beam.

Since the reconstruction signal-to-noise is dependent on telescope's specifications, the telescope design will be crucial for these observations. In particular we will discuss the results we get for two different SKA models: the first will be modeled as a uniform array, while the second is more accurate and will take into account the real visibility space distribution of the baseline. We will see that the results can be quite different and that adopting a more realistic telescope model will considerably improve them.

5.1 Discrete 21 cm Field Modeling

We have developed the properties and the statistics of Gaussian random fields (GRF) in Appendix A, in order to apply such notions to the 21 cm brightness temperature fluctuation field and to the underlying lensing mass distribution. A Gaussian random temperature field has been generated in our C++ code following the storing prescriptions developed in Appendix C.4. At each k_p contribution, we assigned a real and imaginary number from

$$\mathcal{R}(T_{l,k_p}) \propto G_1 \sqrt{\frac{C_{l,k_p}}{2}} \quad \mathcal{I}(T_{l,k_p}) \propto G_2 \sqrt{\frac{C_{l,k_p}}{2}}, \quad (5.1)$$

in order to create a GRF in Fourier space, where $G_{1,2}$ are two random Gaussian numbers with null mean and unitary standard deviation. The normalisa-

tion constant depends on desired discrete temperature fluctuation field units and on DFT definition. In particular, following our definition Eq. (D.25), we adopted a $\sqrt{\Omega_s}$ normalization. This is equivalent to a \sqrt{V} normalization if we translate multipoles into wave-vectors \mathbf{k} , with V being the physical volume of the box in Mpc^3 , since $P(k) = VC_{l,k_p}/\Omega_s$ and the discrete power spectrum is defined following Eq. (4.11). The two Gaussian numbers are extracted using a Box-Muller polar transform. The resulting real and imaginary parts of the field will be Gaussian distributed with uniform phases in the range $[0, 2\pi]$ and are mutually independent. If $\sigma_{l,k_p}^2 = C_{l,k_p}/2$, we can write their statistically homogeneous and isotropic probability distribution as:

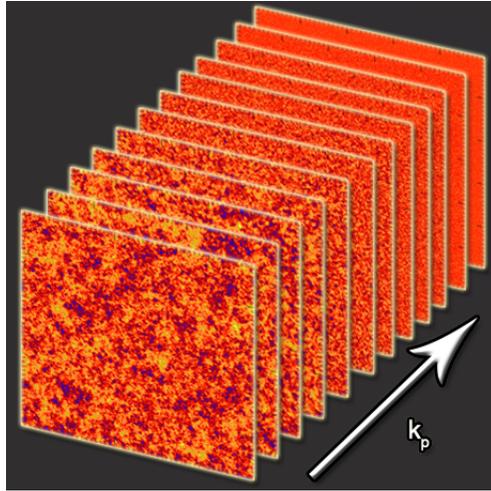
$$\mathcal{P}[\mathcal{R}(T_{l,k_p}), \mathcal{I}(T_{l,k_p})] = \frac{1}{\sqrt{2\pi\sigma_{l,k_p}^2}} e^{-[\mathcal{R}(T_{l,k_p}), \mathcal{I}(T_{l,k_p})]^2 / 2\sigma_{l,k_p}^2}, \quad (5.2)$$

so that the moduli of the Fourier amplitudes will have a Rayleigh distribution (Coles & Lucchin, 2002). As learned in Appendix C.2.1, the Nyquist sampling theorem ensures that modes aliasing is avoided if the Nyquist frequency, *i.e.* the maximum sampled frequency, is bigger than the maximum component frequency. So we generated a positive and negative frequency spectrum, satisfying for every k_p mode the Hermitian condition $T_{-l,k_p} = T_{l,k_p}^*$ for the field's reality¹ and following the prescription described in Appendix C.4.1.

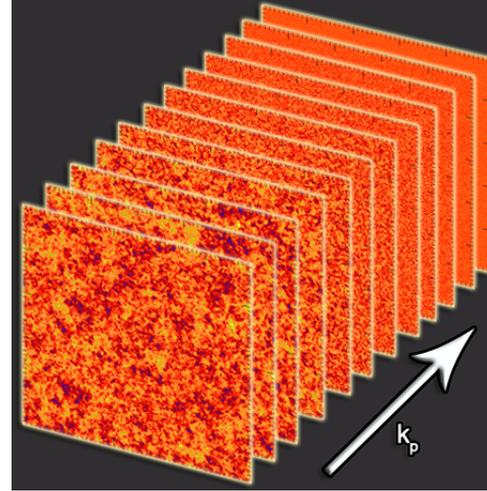
Although the 21 cm power spectrum depends substantially on reionization parameters such as gas density, mean temperature, Ly- α flux and neutral hydrogen fraction at emission time, we adopted the simplified model Eq. (4.11) in our simulation. As discussed in section 4.2, the approximation equation (4.11) for the brightness temperature power spectrum, ignores partial reionization, temperature fluctuations and HI bias. We approximate the non-linear matter power spectrum for structure formation using the Peacock & Dodds (1996) method although the lensing signal and noise are relatively insensitive to non-linear scales.

In Figure 5.1.1 we show a sample of our simulation boxes, where we can see the unlensed 21 cm brightness temperature fluctuation field produced for several k_p . Modes with a larger values of k_p have less power and the signal quickly decays below the thermal noise with increasing k_p , as shown in Figure 4.1. Because of this, we do not need to simulate a large number of k_p maps, allowing the code to be faster and optimized to produce reconstructed maps in reasonable time.

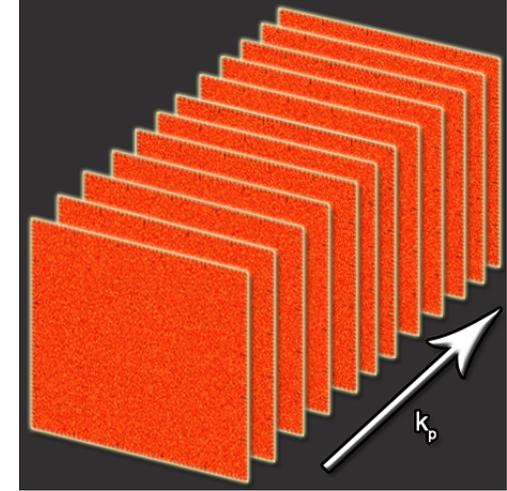
¹Notice that in case of full 3D treatment, the reality condition would be applied to k_p modes too: this would cause undesired correlations among redshift modes.



5.1.1 Unlensed 21 cm simulated box



5.1.2 Lensed 21 cm simulated box



5.1.3 Thermal noise simulated box

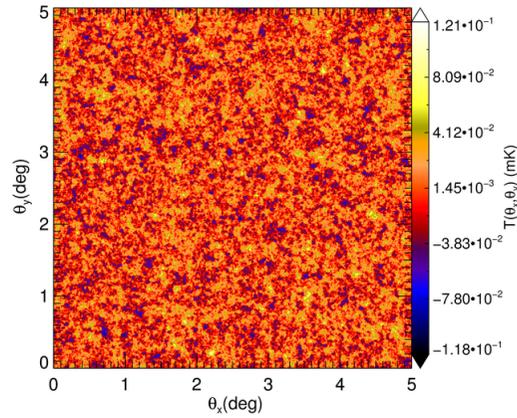
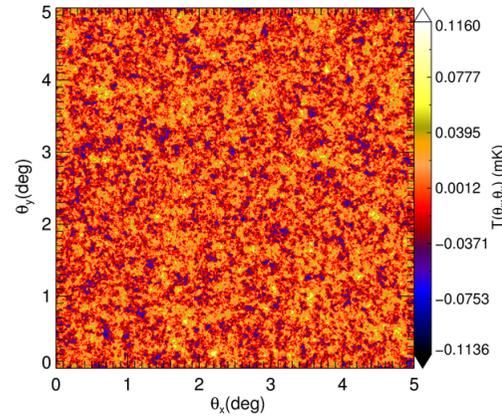
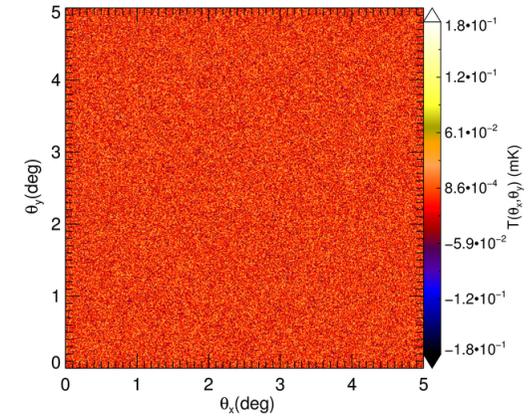
5.1.4 Unlensed 21 cm slice for $k_p = 3$ 5.1.5 Lensed 21 cm slice for $k_p = 3$ 5.1.6 Thermal noise slice for $k_p = 3$

Figure 5.1: *Top*: Sample realisations of the simulated box centred around $z_s = 8$ for every simulated component, made for individual k_{\parallel} modes. The angular area is $\Omega_s = 5^\circ \times 5^\circ$ and $N_{\text{side}} = 650$. *Bottom*: The $k_p = 3$ map sextacted from the cubes are shown in the bottom panels.

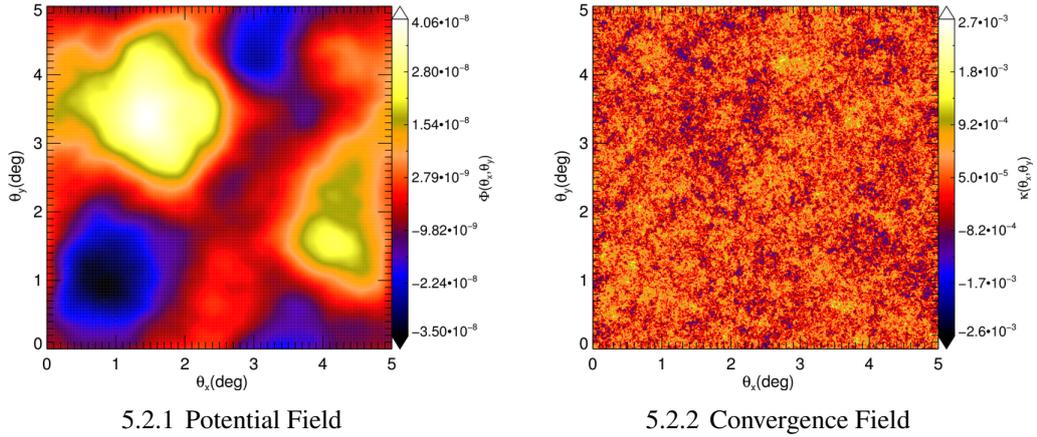


Figure 5.2: *Left*: A sample realisation of the gaussian random potential field that generated the deflection of 21 cm intensity points. *Right*: The corresponding convergence field.

5.2 Lensing Maps

Since we do not expect highly non-linear objects in the deflection potential to be detectable, we model the deflection field as a Gaussian random field in much the same way as we did the brightness temperature fluctuations. First potential field is generated analogously to what has been done in Section 5.1, but using the power spectrum (4.4) and $C_L^{\Phi\Phi} = C_L^{\alpha\alpha}/L^2$ into Eq. (5.1). One example of a potential and convergence field realisation is shown in Figure 5.2 for $N_{\text{side}} = 650$ and $\Omega_s = 5^\circ \times 5^\circ$. Then we produced the components of the deflection field in Fourier space, namely $\tilde{\alpha}_L = iL\tilde{\Phi}_L$. Lensed 21 cm temperature brightness maps are hence produced applying, for each k_p map, a realisation of our randomly generated x - and y -deflection field maps². In practice, if \tilde{x} is the lensed angular position, and x is the old one, we have

$$\tilde{x} = x - \alpha = \Delta x(l, m) - \alpha. \quad (5.3)$$

Then, the value of the lensed field is computed through bicubic interpolation of the values at the undeflected position. Bicubic interpolation is a method which smooths the interpolated surface on a grid more than the corresponding fields obtained by bilinear interpolation or nearest-point interpolation, showing less interpolation artifacts and more realistic results. It gives the same degree of smoothness as bicubic spline interpolation, but it has the advantage of being a local method (Press et al., 2002). In spite of considering only 4 pixel values as for bilinear interpolation, bicubic interpolation considers the contributions from 16 pixel values. In fact, for each nearest point found in a square grid cell, we need to specify not only the value of the field but its first and

²In order to have a gradient operation consistent with Eq. (4.8), we need to add an Ω_s factor when the deflection field is included.

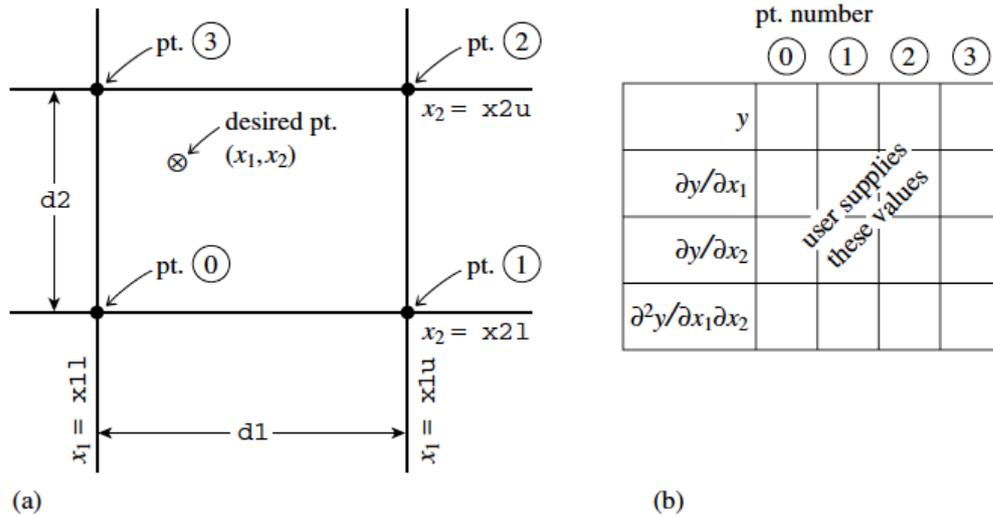


Figure 5.3: (a) Labeling of points used in the two-dimensional bicubic interpolation routines. (b) For each of the four points in (a), the user supplies one function value, two first derivatives, and one cross-derivative, a total of 16 numbers. Taken from Press et al. (2002).

second mixed derivatives as well, as shown in Figure 5.3. This will allow for finding a cubic interpolating function whose function and derivatives values are reproduced exactly on grid points and change continuously as the interpolating point crosses from one grid square to another. Interested readers are advised to consult Press et al. (2002) for a more detailed discussion. In general, the method's accuracy will depend on derivative's one, which needs to be provided separately.

When light rays are shot outside the simulated source boundaries, periodic boundary conditions are applied by mirroring the source plane grid points. So, if the new lensed pixels are $\tilde{l} = \tilde{x} / \Delta x = N - 1$ and $\tilde{m} \neq N - 1$, the $l + 1$ -th point is the first one on the grid, $l = 0$. If $\tilde{l} = \tilde{m} = N - 1$, then the $(l, m) + 1$ -th point is $(0, 0)$.

A sample box image for the resulting lensed temperature field is shown in Figure 5.1.2. It is hard, but not impossible, to see differences by eye between the unlensed and the lensed image, as one can notice from Figure 5.1.5.

As pointed out by (Lewis, 2005) another way to simulate CMB or 21 cm lensing is by downsampling a temperature field image produced at a much higher resolution (usually 3 or 4 times the desired resolution) in order to smooth undetectable sub-pixel structures when rays are shot. To re-map the field, one can always use the bicubic interpolation scheme, but that can be time consuming if not properly parallelized. In any case, moving the points and re-mapping them via bicubic interpolation without generating a map at a higher resolution is fast and safe enough to avoid any problem for our purposes³. In our case, bicubic interpolation is accurate enough to avoid the generation of pixelization artefacts in both lensing reconstruction and lensed maps. To this

³This issue becomes more important if one wishes to lens a polarization image, because lensed polarized modes will depend more considerably on high multipoles.

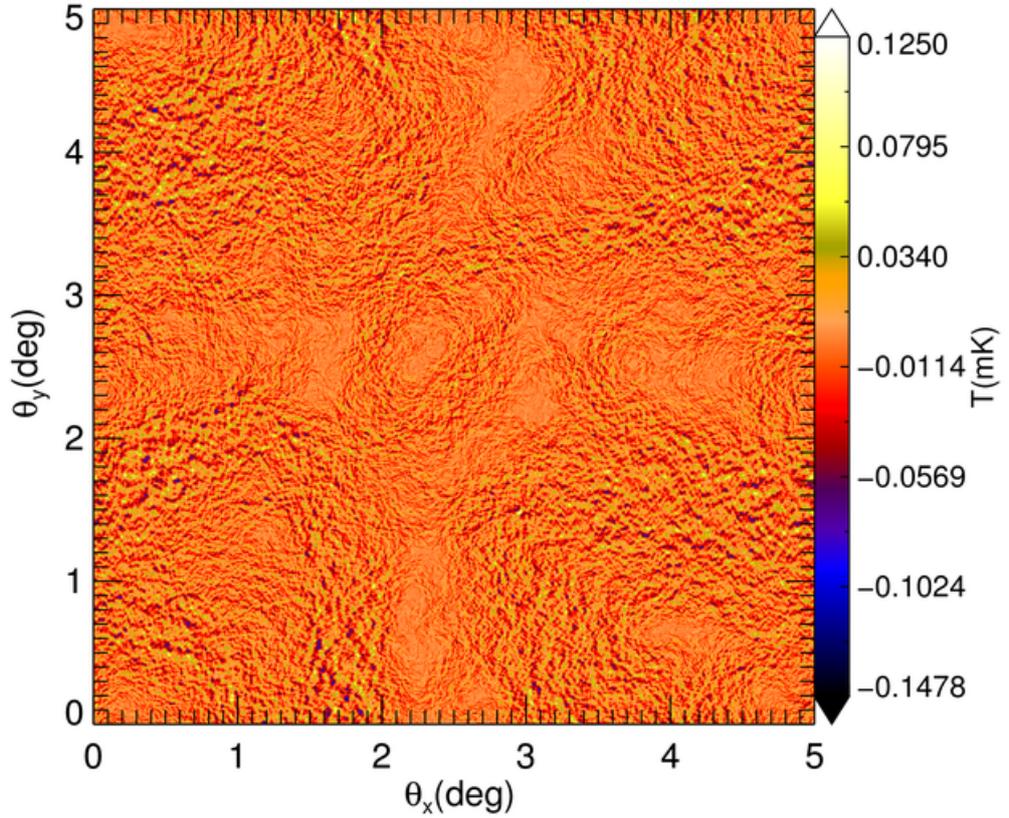


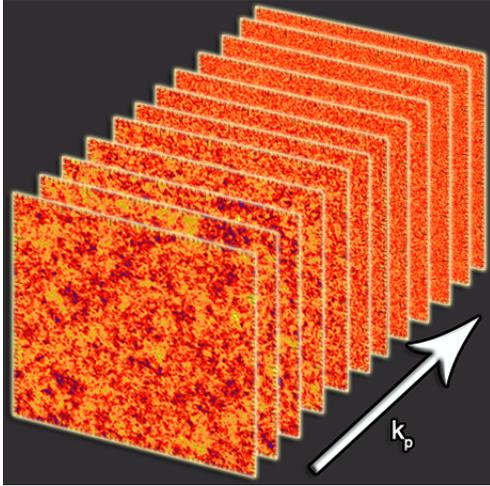
Figure 5.4: The residual image of the difference between Figure 5.1.5 and Figure 5.1.4, using as input potential field Figure 5.6.1. Here $\Omega_s = 5^\circ \times 5^\circ$ and $N_{\text{side}} = 650$.

purpose we show also Figure 5.4, the residual map of the lensed temperature map Figure 5.1.5 respect to the unlensed one Figure 5.1.4. The lensing potential field used to produce this map is pictured on Figure 5.6.1, and we see that this residual map traces the gradient magnitude of the potential point.

Both lensing potential and 21 cm temperature maps benefit from periodic boundary conditions induced by FFTs. This periodicity could be compromised by effects such as cut masked areas or incomplete coverage of our field of view. For simplicity we will not consider them in this work, avoiding the need to apply methods such as padding or apodization (consult (Pearson et al., 2014) or (Plaszczynski et al., 2012) for a CMB example) to treat these cases. Our code is designed to handle these methods, but they will be included in future works.

5.3 Modeling the Beam

As already stated in Section 4.4 the beaming function has the effect of smoothing the Fourier frequencies near to the characteristic beam frequency, which we call L_{cut} corresponding to a beam resolution σ . The form of the beam function does not depend on observed frequency ν . We use a simple Gaussian



5.5.1 Beamed 21 cm simulated lensed box.

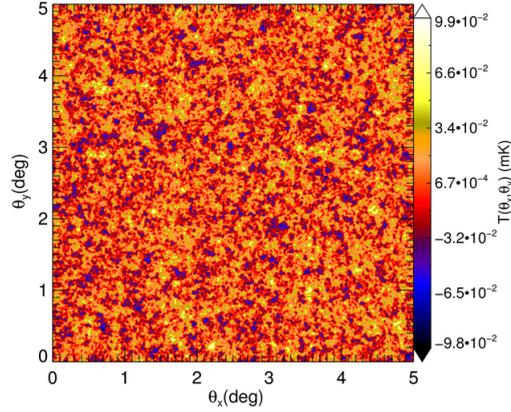
5.5.2 Beamed 21 cm simulated lensed map for $k_p = 3$.

Figure 5.5: *Left*: A sample realisation of the simulated beamed box centred around $z_s = 8$, for different k_p modes. The angular area is $\Omega_s = 5^\circ \times 5^\circ$ and $N_{\text{side}} = 650$. *Right*: The $k_p = 3$ map extracted from the above cube. The beam smoothing is 2.5 times the basic map resolution, namely $\Delta b = 1.15$ arcmins.

beam

$$W_l = e^{-l(l+1)\sigma^2/2}, \quad (5.4)$$

with $\sigma = b\Delta\theta / \sqrt{8 \ln 2}$, where b quantifies the beam size with respect to angular resolution of the simulated map $\Delta\theta$. Seen under another point of view b is the ratio between the Nyquist mode L_{cut} and the beam characteristic one L_{cut} . The left panel of Figure 5.5.1 is a realisation of our simulation box centered at $z = 8$ which includes the beam, whose $k_p = 3$ slice is shown in the right panel. It is hard to notice from this simple picture the suppression of smallest scale modes with respect to Figure 5.1.5.

Assuming the SKA-Low instrument, the typical beam resolution will depend on the maximum baseline length, which defines the smallest observable scale. Its definition is

$$L_{\text{cut}} = \frac{2\pi D_{\text{max}}}{\lambda(z)}, \quad (5.5)$$

and, as we will see in next section, assuming an observation made at $z = 8$ and a maximum baseline array of 4 Km, we have $\Delta b = b\Delta\theta = 1.15$ arcmins. This value is set by $\Delta b = \sqrt{2}\pi/L_{\text{cut}}$ for a square grid, where $L_{\text{cut}} = L_{\text{max}}$. In fact, $L_{\text{Nyq}} = \sqrt{L_x^2 + L_y^2} = \sqrt{2}\Delta L N_{\perp}/2$, with $\Delta L = 2\pi/\theta_{\text{max}}$ the resolution in Fourier space.

5.3.1 Aliasing and the Beam

When using the fast lensing estimator, (4.24), we find that when the beam cutoff, σ , is close to the resolution of the image or when no beam is taken into account, spurious aliasing effects occur causing the lensing signal and noise to disagree with the input signal and the analytically calculated noise.

Setting the beam resolution and its relation to the Nyquist multipole L_{Nyq} is a fundamental point in this study: the quadratic estimator is a convolution of filtered fields and there will be a visible aliasing effect if these two frequencies are too close in Fourier space. With the slow Fourier space summed estimator one can avoid the contamination of spurious frequencies by setting with the sums all the allowed L modes (*i.e.*, imposing that they are confined inside the map's boundaries). Aliasing is indeed generated by spurious bouncing of modes with $L > L_{\text{Nyq}}$ whose contributions have already taken into account when slow Fourier space sums are performed over map's boundaries. In this case the convolution in Fourier space is developed into Eq.(4.21) and the convolved modes L can be kept under control. For the fast estimator this is not possible because the filtered fields' convolution is performed with FFTs transforming periodic data into periodic data. As already pointed out by Bowman & Roberts (2010), the result is an aliasing effect generated by the difference between our FFT discrete periodic convolution operation and the linear one, analytically extended from $-\infty$ to $+\infty$, as performed in Eq. (4.21).

Incorporating a beam solves this problem for the fast estimator because it acts as a low-pass filter that reduces the aliased contamination coming from high frequency modes⁴. This problem is much less prominent for CMB lensing because in that case there is relatively little power in the high frequency modes.

To reduce memory usage and computational time it is advantageous to keep L_{Nyq} as small as possible while avoiding this aliasing problem. We found that the beam resolution has to be bigger than $2.5\Delta\theta$: this means that $L_{\text{Nyq}} \geq 2.5L_{\text{cut}}$ to avoid this problem. Tests of this limit are discussed in Section 5.5.2.

As a final caveat, let us point out that simulating a beamed estimator allows us to avoid any complications due to maps de-convolution in Fourier space, which could be another source for aliasing. This is just another aspect of our previous problem, because it depends on the way we sample frequencies in Fourier space, and hence on the relative distance between the Nyquist frequency and the cut frequency imposed by the beam.

5.4 Uniform Thermal Noise Component

A crucial point in this work is set by discussing how the high-fidelity quality of reconstructed mass depend drastically on the experimental design, foreground and noise. To this purpose we introduce two models for describing an SKA-Low telescope design and will discuss the results obtained within our simulation for both. The first one considers an uniform visibility space coverage, leading to a noise power spectrum which is constant in Fourier space and has been widely used in the SKA literature (Zaldarriaga et al., 2004; Metcalf & White, 2009; Pourtsidou & Metcalf, 2015). The second one will consider a more

⁴Another way to solve this problem is by padding the temperature field in Fourier space with a sufficient number of null arrays. This may be computationally expensive, especially for multi-dimensional arrays. Alternative FFTW efficient methods that do not involve padding in Fourier space have been developed. These eliminate the aliasing error from multi-dimensional convolutions with small computational costs. Interested readers can consult Bowman & Roberts (2010).

accurate description of the array density distribution in visibility space, taking into account the fact that SKA-Low will have a core region in which the antennae are densely distributed (until a radius of ~ 1 Km) and then we have a more sparse region which forms a characteristic spiral arms feature. To find an expression for the power spectrum in this case we will make use of the notation adopted by Santos et al. (2015); Bull et al. (2015); Pourtsidou et al. (2015), resulting in a lower thermal noise level respect to the uniform case for the first 2000 multipoles, *i.e.* where the lensing signal can be reconstructed with higher signal-to-noise. As we will see in Appendix E, this noise presents a steep increment at very high multipoles which forbids reconstruction at these scales.

For the moment we will introduce the model and the results obtained in the first case, postponing the discussion for the second case results in Section 5.6.

The lensing estimator and noise, (4.14) and (4.16), include a total power spectrum contribution C_{ℓ, k_p}^T which depends on on the power spectrum of the noise. The noise consists of a component coming from the sky and from the instrumentation. For the observation we are simulating, the noise from foreground removal will dominate. We will use a simple, but widely used model for the thermal noise in an SKA-like interferometer,

$$C_{\ell, \Delta\nu}^N = \frac{(2\pi)^3 T_{\text{Sky}}^2}{t_o \Delta\nu f_{\text{cov}}^2 L_{\text{max}}^2} \quad (5.6)$$

(Zaldarriaga et al., 2004; Furlanetto et al., 2006), assuming that the telescopes in the array are uniformly distributed on the ground. The observation bandwidth is $\Delta\nu$, t_o is the total observation time, and f_{cov} is the total collecting area of the telescope divided by $\pi(D_{\text{max}}/2)^2$, *i.e.* the aperture covering fraction, with D_{max} the diameter of the maximum array baseline considered. The highest multipole $L_{\text{max}} = 2\pi D_{\text{max}}/\lambda_{\text{obs}}(z)$ that the array is able to probe at the observed wavelength $\lambda_{\text{obs}}(z) = \lambda_{21}(1+z)$.

The planned SKA1-Low (Dewdney, 2013) and SKA2-Low instruments have a collecting area of $A_{\text{coll}} = 0.3 \text{ Km}^2$ and $A_{\text{coll}} = 1.2 \text{ Km}^2$, respectively. In both cases the maximum baseline is $D_{\text{max}} = 4 \text{ Km}$. Thus the telescope covering fraction is $f_{\text{cov}} \sim 0.024$ for SKA1-Low and $f_{\text{cov}} \sim 0.095$ for SKA2-Low. The maximum probed multipole at $z = 8$ is $\ell_{\text{max}} \simeq 13230$.

We will consider, for comparisons and discussions, the survey strategies summarised in Table 5.1.

	$\Delta\nu$ [MHz]	t_o [hrs]
R0	8	1000
R1	5	1000
R2	5	2000

Table 5.1: The considered SKA simulation settings for this study at $z = 8$. For every case we considered a telescope diameter of $D = 4 \text{ Km}$. The total collecting area of the telescope is $A_{\text{coll}} = 0.3 \text{ Km}^2$ for SKA1-Low and $A_{\text{coll}} = 1.2 \text{ Km}^2$ for SKA2-Low. For every case we assume a total frequency range of 50 – 350 MHz.

The frequency range explored by SKA phases is 50 – 350 MHz. We can break

it up in several smaller frequency bands, each one centered on a given observation redshift. For R1 and R2 strategies we have used $\Delta\nu = 5$ MHz. For our single-band results this band has been centered around a source redshift of $z_s = 8$. This bandwidth $\Delta\nu$ is sufficiently thin to have good resolution over a certain redshift range, which is good for exploring EoR epoch, but thick enough so that correlations between bands can be ignored. The band $\Delta\nu$ is divided into channels which will detect individual k_p modes in visibility space. There is a maximum number of detectable k_p modes which will depend on the ratio between the bandwidth and the frequency resolution in a single channel (Parsons et al., 2012). The frequency channel resolution is $\delta\nu = 100$ KHz. Notice that Pritchard et al. (2015) have assumed $B = 8$ MHz and $t_o = 1000$ hours, which corresponds to our R0 survey strategy. R2 survey strategy has been introduced to keep a comparable thermal noise level to R0 and have the possibility to stack more frequency bands to fit a given redshift range.

At such high redshifts the most important source of astrophysical noise is galactic synchrotron emission which produces a representative sky temperature of

$$T_{\text{sky}} = 1.1 \times 60 \left(\frac{\nu_{\text{obs}}}{300 \text{ MHz}} \right)^{-2.55} \text{ K} \quad (5.7)$$

(Dewdney, 2013) will be dominated by this contribution for observed regions that are enough far away from Galactic Plane. The receiver noise power spectrum is computed analogously by following Eq. (5.6) and setting T_{Rcv} in the place of T_{sky} . It is assumed that this contribution is uncorrelated with the signal and with the sky noise term, in order to add it after the inclusion of the beam. We assumed $T_{\text{Rcv}} = 40$ K (Dewdney, 2013) and it contributes significantly only for low-redshift experiments. In Figure 5.1.3 we show a sky noise component cube, where a different noise realisation is produced for each k_p channel. Sky and receiver noise power spectra for the R2 configuration are shown in Figure 4.1, where they are compared with the brightness temperature angular power spectrum for different k_p modes.

Another important instrumental parameter, especially for statistical detection of the signal and for cosmological purposes, is the observed fraction of the sky $f_{\text{sky}} = \Omega_s[\text{sr}]/4\pi[\text{sr}] = \Omega_s[\text{deg}]^2/41253[\text{deg}]^2$. In order to get a good measurement of the lensing power spectrum and a competitive estimate of cosmological parameters, a large f_{sky} is essential. We will see in next sections that the angular power spectrum error scales as $1/\sqrt{f_{\text{sky}}}$ (Pourtsidou & Metcalf, 2015; McQuinn et al., 2006). Current plans take into account a small observed fraction of the sky (6.06×10^{-4} for a $5^\circ \times 5^\circ$ survey), and detecting accurately the lensing power spectrum seems to be a difficult task to achieve if the reconstruction noise is high (see Eq. (4.25)).

5.5 Results for the Uniform Thermal Noise Model

In this section we present our results obtained using the uniform thermal noise model introduced in Section 5.4, focusing on the how we used the derived quadratic estimator to produce images of the underlying weak lensing

potential. We then discuss how the telescope and survey parameters influence the results.

5.5.1 Single-Band Reconstruction

We use a simulated beam cut off multipole that is close to the one permitted for SKA-Low, *i.e.* $L_{\text{cut}} = 13237$ at $z = 8$, or $\nu = 157.82$ MHz, corresponding to an angular resolution of $\Delta\theta \simeq 1.15$ arcmins, and study the performance of the reconstruction technique for different map sizes. Because of the aliasing effect we need to generate the field with a L_{Nyq} that is at least 2.5 times the beam cut off, L_{cut} . The pixel resolution is the same for every case, namely $\Delta\theta \simeq 27.7$ arcsecs and so the corresponding Nyquist multipole is $L_{\text{Nyq}} = 33092.6$. We have considered three observational strategies (a), (b), and (c), listed in Table 5.2. Here N_{side} is the number of pixels per map side, $\Delta l = 2\pi/\theta_{\text{side}}^{\text{max}}$ is the resolution in Fourier space, and f_{sky} is the observed fraction of the sky given the survey area Ω_s .

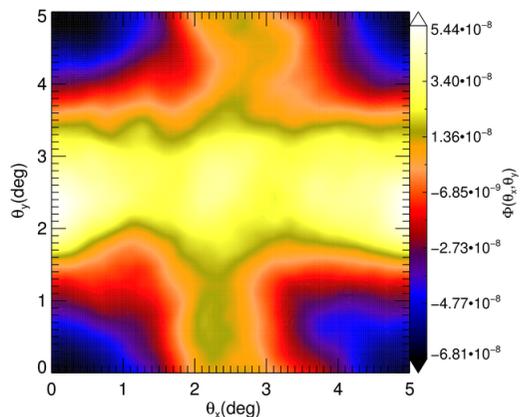
	Ω_s	N_{side}	Δl	f_{sky}
(a)	$5^\circ \times 5^\circ$	650	72	6.06×10^{-4}
(b)	$10^\circ \times 10^\circ$	1300	36	2.42×10^{-3}
(c)	$20^\circ \times 20^\circ$	2600	18	9.7×10^{-3}

Table 5.2: The three considered simulation settings for this study. For every case we have $L_{\text{cut}} = 13237$, $L_{\text{Nyq}} = 2.5 \times L_{\text{cut}}$, and $z_s = 8$.

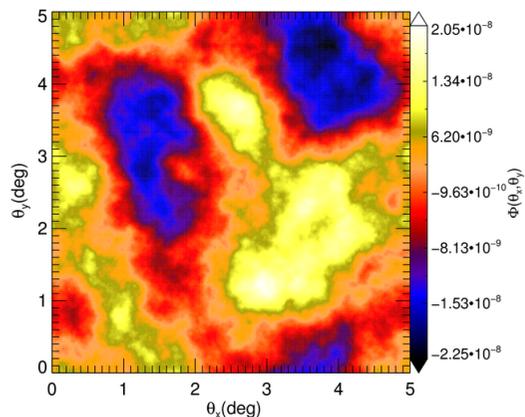
In the following few sections we consider the SKA2-Low instrument described in Section 5.4, with the configuration denoted as R2 in Table 5.1. This is because the reconstruction estimator noise level is too high for the single band detection adopting a SKA1-Low design. This will be more deeply discussed in Section 5.5.4.

We generate our 21 cm temperature brightness and lens each k_p mode separately using a single realisation of the lensing potential field. The band $\Delta\nu$ is centred around redshift $z_s = 8$ and the beam resolution is $\sigma = 1.15$ arcmins. In the final step we added a different receiver noise contribution for every k_p slice. Using the estimator described in Section 4.4 summing over the first 20 k_p modes we obtain maps of the estimated lensing potential.

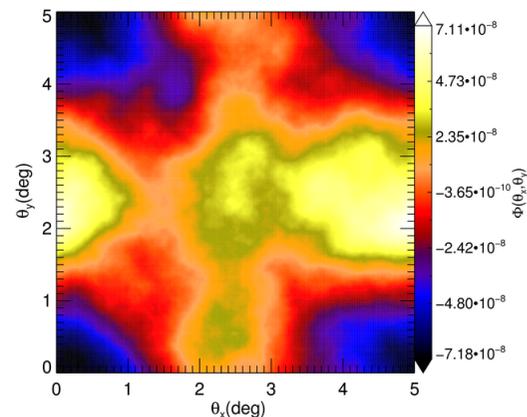
The results for the three cases listed in Table 5.2 are presented in Figures 5.6, 5.7, and 5.8, where the input potential field, the recovered pure estimator noise image and the recovered potential field are shown. For case (a) we are able to recover a noisy version of the input map shape; in this case a large number of available modes are under the noise level. As we increase the map dimensions in cases (b) and (c), we have more large scale structure available at low multipoles and a better image is recovered. Hence, if the resolution is fixed, the performance of the reconstruction technique increases with increasing map size, as more additional and becoming available modes are above the noise level.



5.6.1 Input potential field case (a)

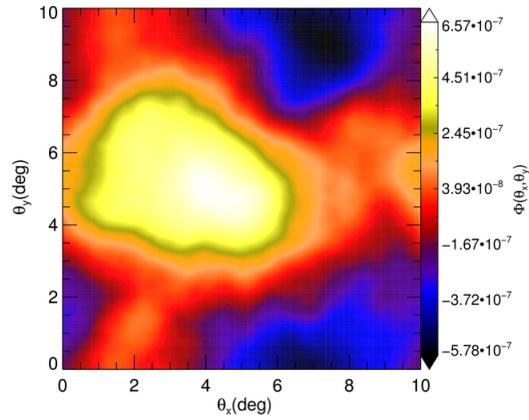


5.6.2 Estimator Noise map case (a)

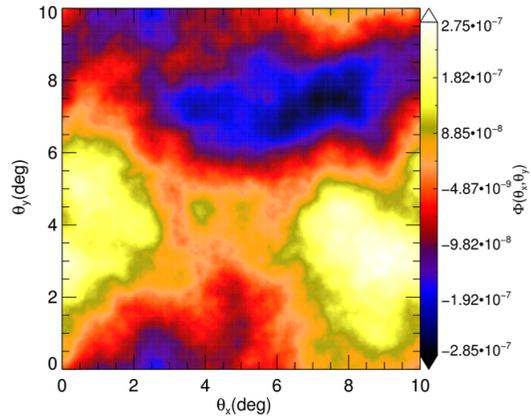


5.6.3 Estimated potential map case (a)

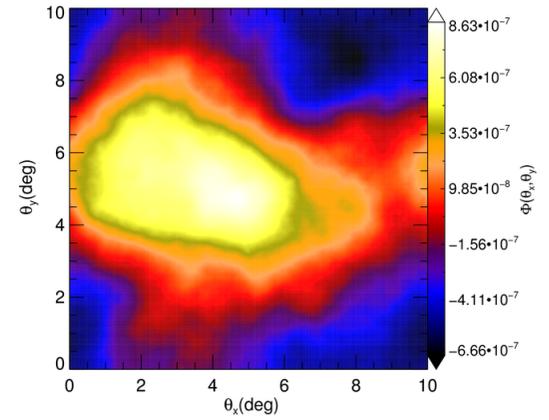
Figure 5.6: Reconstructed potential images from a realisation of the input 21 cm source box with beam cut off $L_{\text{cut}} = 13237$ at $z_s = 8$ for case (a) listed in Table 5.2. On the middle panel we see the estimator noise image produced by our estimator without any input lensing signal. On the right panel we show the reconstructed potential using the contribution of 20 k_p modes. For every case we used a SKA2-Low configuration with choices for observational time and bandwidth denoted as R2 in Table 5.1.



5.7.1 Input potential field case (b)

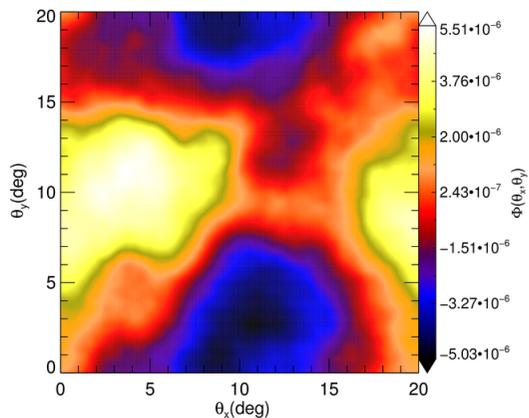


5.7.2 Estimator Noise map case (b)

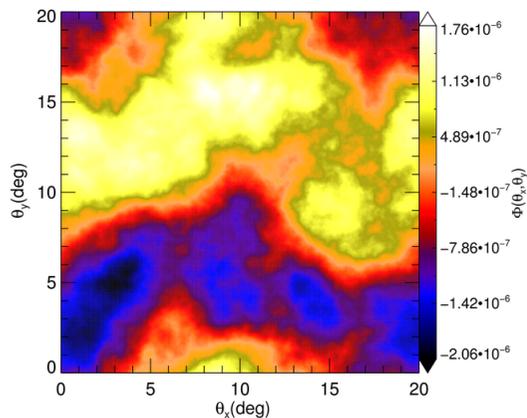


5.7.3 Estimated potential map case (b)

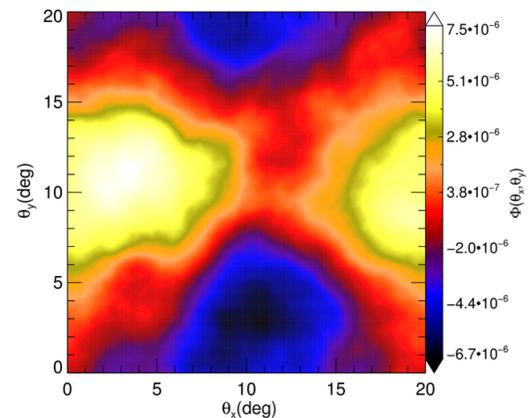
Figure 5.7: Reconstructed potential images from a realisation of the input 21 cm source box with beam cut off $L_{\text{cut}} = 13237$ at $z_s = 8$ for case (b) listed in Table 5.2. On the middle panel we see the estimator noise image produced by our estimator without any input lensing signal. On the right panel we show the reconstructed potential using the contribution of 20 k_p modes. For every case we used a SKA2-Low configuration with choices for observational time and bandwidth denoted as R2 in Table 5.1.



5.8.1 Input potential field case (c)

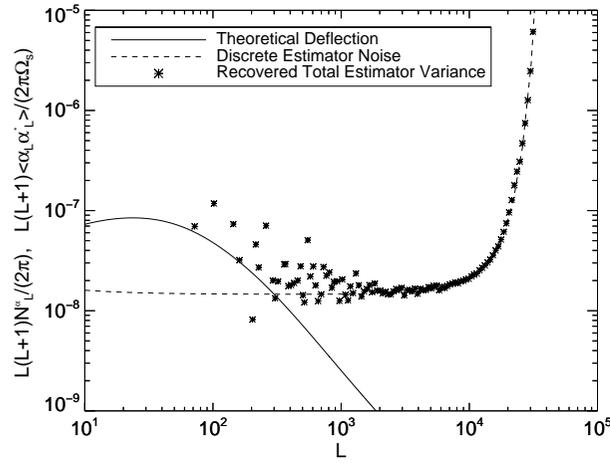


5.8.2 Estimator Noise map case (c)

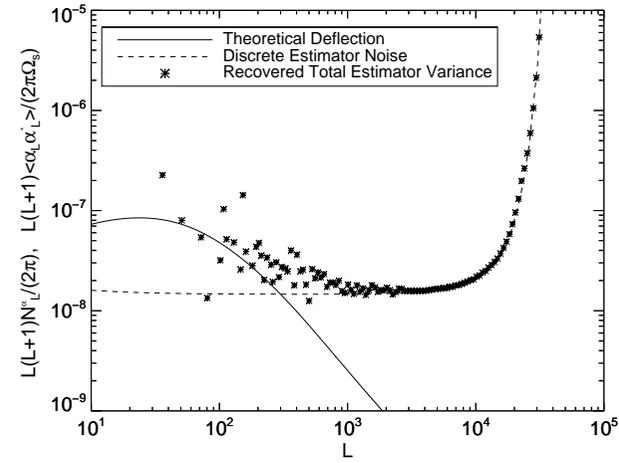


5.8.3 Estimated potential map case (c)

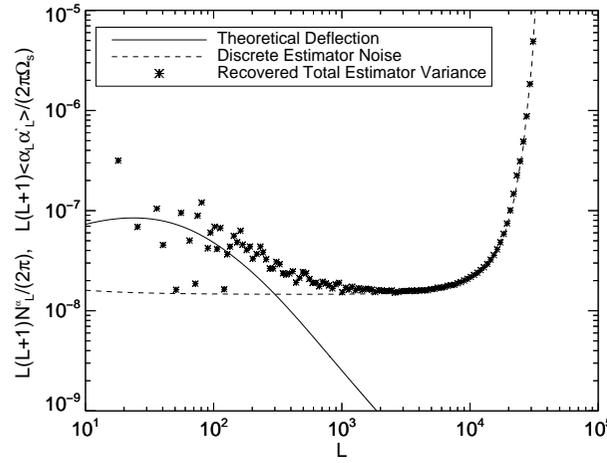
Figure 5.8: Reconstructed potential images from a realisation of the input 21 cm source box with beam cut off $L_{\text{cut}} = 13237$ at $z_s = 8$ for case (c) listed in Table 5.2. On the middle panel we see the estimator noise image produced by our estimator without any input lensing signal. On the right panel we show the reconstructed potential using the contribution of 20 k_p modes. For every case we used a SKA2-Low configuration with choices for observational time and bandwidth denoted as R2 in Table 5.1.



5.9.1 Recovered estimator variance for case (a).



5.9.2 Recovered estimator variance for case (b).



5.9.3 Recovered estimator variance for case (c).

Figure 5.9: Recovered estimator variance for the considered settings listed in Table 5.2, corresponding to images Figures 5.6.3, 5.7.3, and 5.8.3 respectively.

We computed the Fourier mode amplitude squared of these recovered potential maps, as seen in Figure 5.9. From here onwards we will show the resulting deflection field power spectra instead of the potential ones, for a better visualisation of the plots. The low- L modes are the ones mostly involved in signal reconstruction, but they suffer from a considerable sample variance, because the mode density is low in that range. We will see in Section 5.5.3 how the situation improves as we increase the number of realisations and Eq. (4.15) is recovered.

In order to quantify the accuracy of the reconstructed images we define the "Fidelity" F_ϕ as

$$F_\phi(L) = \left[\frac{|\hat{\phi}_L - \phi_L|}{|\phi_L|} \right]^{-1}, \quad (5.8)$$

the fractional difference between the estimated and the true potential at every mode L . Bigger fidelities corresponds to better reconstruction of the lensing potential. We can see from Figure 5.10 that as we increase the map's dimensions, going from case (a) (red dashed line) to case (c) (black dashed line), we get a more accurate reconstruction for modes in the intermediate range $100 \leq L \leq 1000$, corresponding to the angular range $15 \text{ arcmins} \leq \theta \leq 2.55^\circ$. This is because we are improving the resolution in Fourier space, so the modes' density in case (c) is higher than case (a), and the reconstruction is better sampled in this range. In general, we notice that large scales are more accurately reconstructed than small scales which are largely contaminated by the estimator's reconstruction noise. It can be seen that many modes with $L \lesssim 500$ have fidelity above one. Note that the three cases have comparable a comparable fidelity level.

To give a more quantitative idea about the correctness of our estimator, we report the number of modes n_i that have Fidelity bigger than one in first three columns of Table 5.3. These are computed for the ranges $L \leq 200$, $200 < L \leq 500$, and $500 < L \leq 1000$, and considering for each row the cases listed in Table 5.2. Then we report the total number n_{tot} of simulated mode and the fractional number of modes $f_i = n_i/n_{\text{tot}}$ for the L -ranges considered before. It is clearly seen how the number of well reconstructed modes increases as larger scale modes become available from case (a) to case (c). The fraction of modes in the considered ranges is always low, because the number of modes in a range δL is weighted by the total number of modes, which increases from case (a) to case (c) and so the total signal-to-noise is more or less constant.

	$n_{L \leq 200}$	$n_{200 < L \leq 500}$	$n_{500 < L \leq 1000}$	n_{tot}	$f_{L \leq 200}$	$f_{200 < L \leq 500}$	$f_{500 < L \leq 1000}$
(a)	10	31	53	211900	4.72×10^{-5}	1.46×10^{-4}	2.5×10^{-4}
(b)	37	115	194	846300	4.37×10^{-5}	1.36×10^{-4}	2.3×10^{-4}
(c)	133	464	731	3382600	3.93×10^{-5}	1.37×10^{-4}	2.16×10^{-4}

Table 5.3: The columns are the number of modes that have Fidelity bigger than 1 for $L \leq 200$, $200 < L \leq 500$, and $500 < L \leq 1000$, the total number of available modes, and the corresponditive fractional number of modes with Fidelity bigger than one for each of the considered L -ranges. Each row corresponds to the three cases listed in Table 5.2.

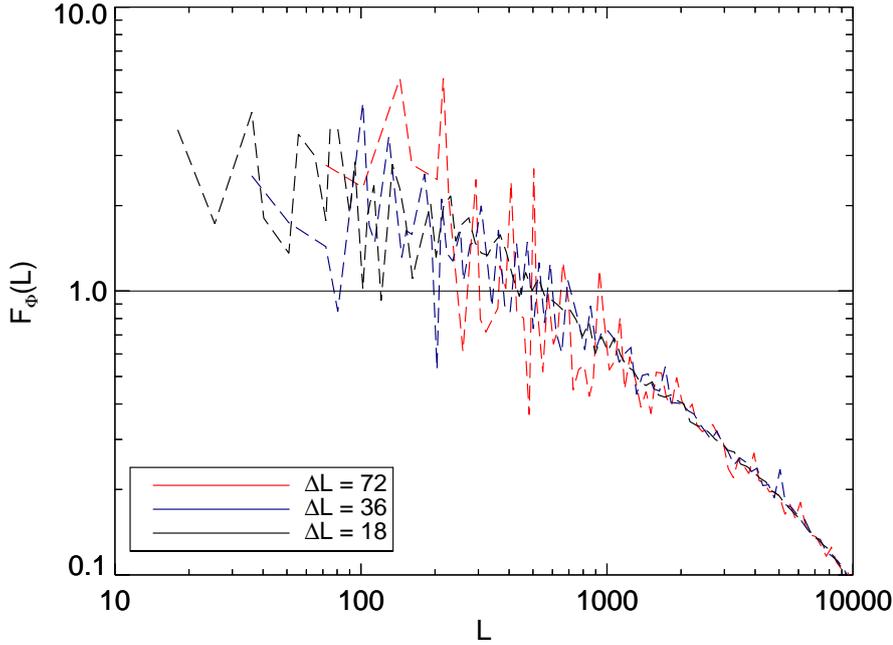


Figure 5.10: The fidelity of the reconstructed lensing potential images as a function of the multipole mode L . Red, blue, and black dashed lines are for Table 5.2 cases (a), (b), and (c) cases, respectively. The telescope model is SKA2-Low R2. The straight line with $F_\phi(L) = 1$ helps to distinguish modes with good fidelity from the ones with bad fidelity.

Note that it is possible to use fewer L -modes to reconstruct the lensing potential, but the resulting image will have a poorer resolution. For example, at $z_s = 8$ it can be seen that the estimator noise crosses the deflection field power spectrum at $L \leq 300$, so the image effective resolution will be $\Delta\theta \geq 51$ arcmins.

The forecasts presented here depend on our rather simple model for reionization and the distribution of HI at high redshift, Section 4.2. If the true reionization history varies a great deal from what we have assumed, for example reionization is extended over a large redshift range⁵ or ends well before $z=8$, then these forecasts will not be valid, since the estimator will not be optimal. We will extend this work to more complicated reionization scenarios in the future, and, for the sake of simplicity, we keep assuming that EoR has been a uniform process for redshifts around $z_s = 8$.

5.5.2 Testing Aliasing Contamination

As already stated (see Section 5.3.1), an important problem in our simulations is the aliasing effect coming from the convolution performed in the estimator Eq. (4.21). Let assume that there is a negligible aliasing effect when the non-aliased slow estimator gives the same result as the fast estimator. Hence, a comparison between the two estimators is a good tool in order to see how strong the aliasing effect is. We can visualise aliasing using the variance of

⁵Observations ensure that EoR ended at redshifts $z > 6$ (Zaroubi, 2013).

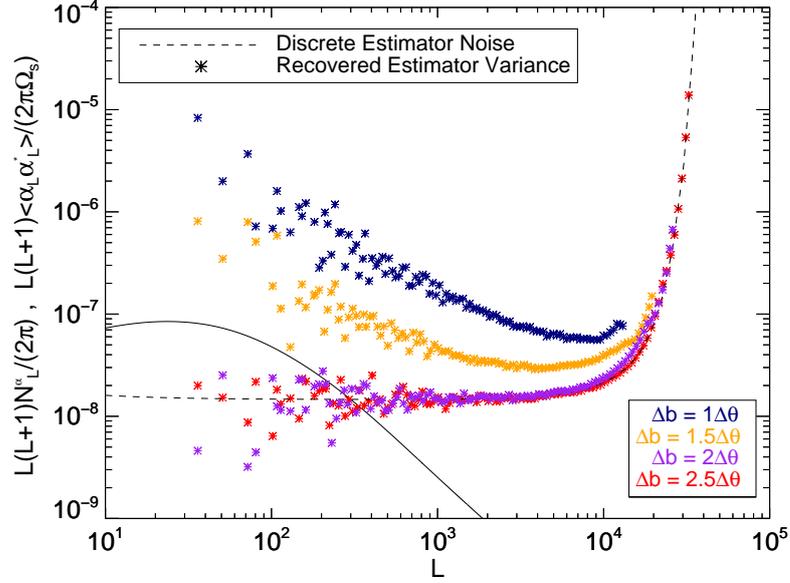
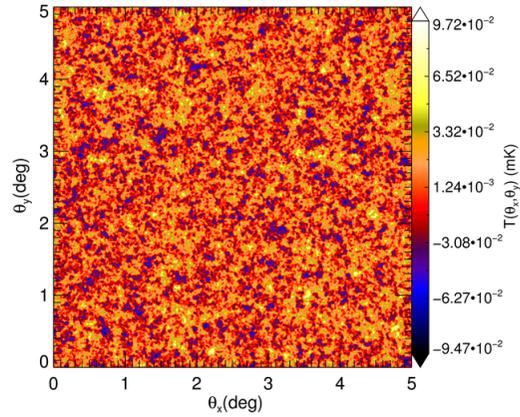


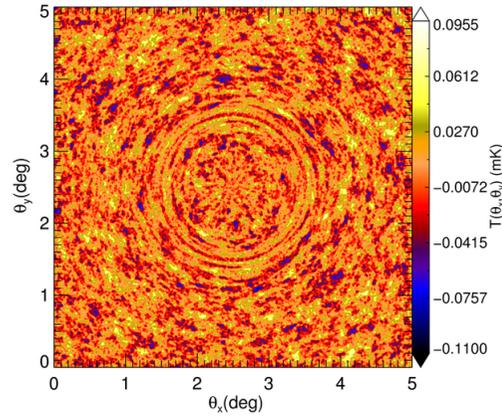
Figure 5.11: The recovered estimator variance (star points) in absence of lensing signal compared to the discrete estimator noise (solid lines), since $\langle \alpha_L \alpha_L^* \rangle / \Omega_s = N_L^\alpha$. Produced maps have $\Omega_s = 10^\circ \times 10^\circ$, $L_{\text{cut}} = 13237$ and the estimator is recovered with $k_p^{\text{max}} = 20$ for $z_s = 8$. We vary the distance between our fixed L_{cut} and the Nyquist frequency L_{Nyq} by changing the ratio between the beam resolution and the resolution of the simulation. We used a SKA2-Low configuration with choices for observational time and bandwidth denoted as R2 in Table 5.1.

the estimator. In absence of lensing, we know that, because of Eq. (4.15), the relation $\langle \hat{\phi}_L \hat{\phi}_L^* \rangle = \Omega_s N_L^\phi$ has to be satisfied. We can see from Figure 5.11 that below a certain ratio Δb between cut and Nyquist frequency, aliasing imprinting causes spurious power to be distributed over the simulated frequency range. As we can see from Figure 5.11, in order to avoid aliasing in our reconstruction simulation, we need to have a Nyquist frequency that is at least $2.5L_{\text{cut}}$. This is the reason for $\Delta b \geq 2.5\Delta\theta$. It can be seen how, for smaller values of this ratio, the aliasing effect becomes more important.

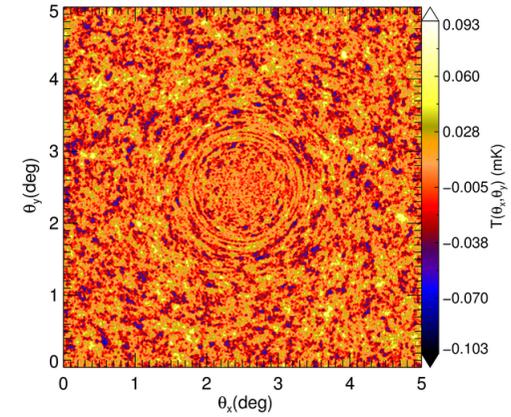
In order to study the behavior of this effect for different redshifts, we considered using a lower source redshift, $z_s = 7$, even though the Universe is unlikely to have been completely neutral at this time. The estimator noise level is lower and we have $L_{\text{cut}} \simeq 14885$. The signal-to-noise is slightly higher than the $z_s = 8$ case. The beam's resolution is higher, namely $\Delta\theta \sim 1.03$ arcmins. This means that we might need larger grids to avoid aliasing, with $L_{\text{Nyq}} \sim 37165.56$ and with $N_{\text{side}} = 730, 1460, \text{ and } 2920$ pixels for field-of-views of the cases (a), (b), and (c), respectively. If higher redshifts are considered, L_{cut} will be lower: for example at $z_s = 10$, $L_{\text{cut}} \simeq 10855$. This means that in principle smaller grids might be used, but, as we will see in few sections when multi-band detections will be studied, one needs to fix the simulated Nyquist resolution and change the ratio b between the Nyquist frequency and the cut one to obtain the proper estimator noise level. More details on behaviour of the estimator's noise at various redshifts will be given in Section 5.5.5.



5.12.1 Non lensed 21 cm temperature map



5.12.2 SIS lensed 21 cm temperature map



5.12.3 PM lensed 21 cm temperature map

Figure 5.12: Demonstrative images for a realisation of lensed 21 cm temperature maps at $k_p = 3$. All the images have been smoothed after being lensed with a beam resolution of $\Delta b = 1.15$ arcmins. The thermal sky noise has been included as well. In the left panel we can see the non lensed map; in the middle panel the 21 cm map has been lensed with a Singular Isothermal Sphere deflection field with Einstein radius $\theta_E = 1.15$ arcmins; in the right panel the 21 cm map has been lensed with a Point Mass deflection field with Einstein radius $\theta_E = 0.69$ arcseconds.

The behavior of the rule $L_{\text{Nyq}} \geq 2.5L_{\text{cut}}$ can be investigated for different redshifts. Instead of keeping the same grid dimension and change Δb , we want to equivalently stuck with $b = 2.5$ and use a smaller square grid with $N_{\text{side}} = 450$, considering $z = 12$ and $L_{\text{cut}} = 9160$. We have seen that the resulting estimator power spectrum in absence of lensing signal is weakly aliased respect to the discrete estimator noise level. So a value of $L_{\text{Nyq}} = 2.5L_{\text{cut}}$ is not enough to ensure the estimator to be aliasing-free, but b needs to be slightly higher. This is due to the approaching of the characteristic beam scale L_{cut} to the one in which the power spectrum starts to bend (see Figure 4.1 for clarity), causing the presence of more power at scales closer to the Nyquist frequency. Consequently for smaller redshift like $z \approx 6$, a b slightly smaller than 2.5 is enough to avoid aliasing in reconstructed images.

5.5.3 Other Tests and Multi-Realisation Reconstruction

Apart from the aliasing issue described in the previous subsection, we have performed further checks like comparing the recovered power spectra to the input power spectra for 21 cm radiation and thermal noises.

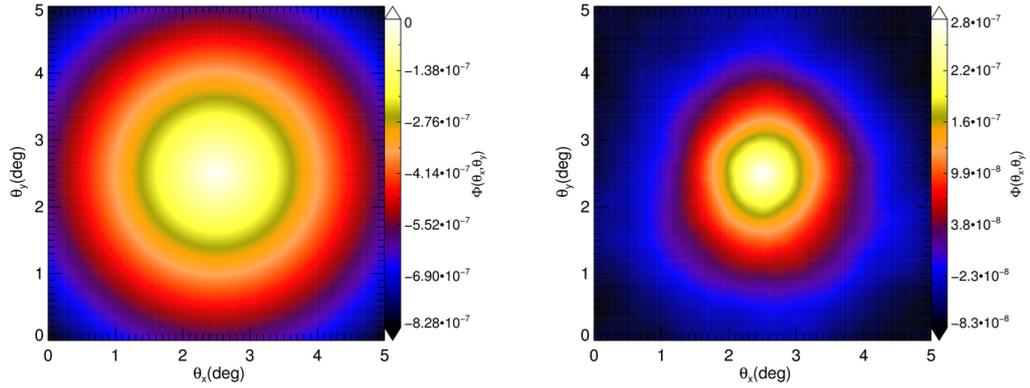
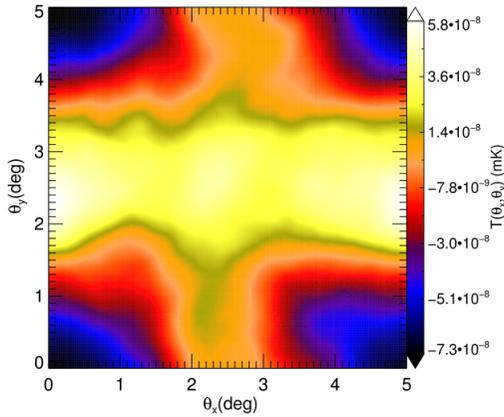
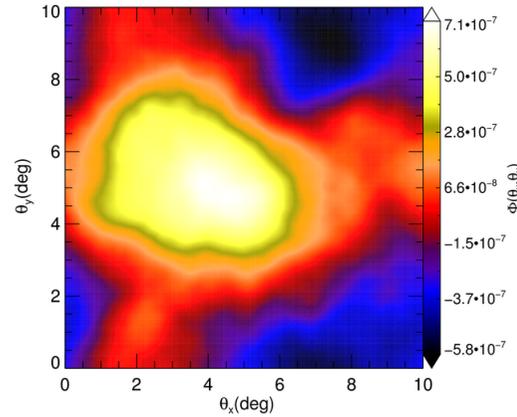


Figure 5.13: *Left panel*: the input test SIS potential with $\theta_E = 6.9$ arcsec and $\Omega_s = 5^\circ \times 5^\circ$. *Right panel*: the recovered potential obtained by using our weak lensing estimator.

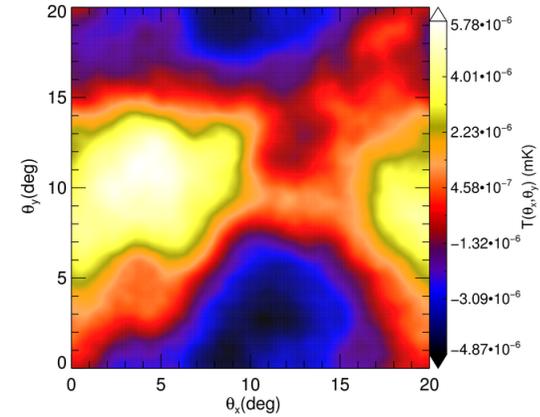
We have also tested our estimator by using strong lensing toy potential models, like the ones presented in Appendix B, to deflect the input 21 cm brightness temperature fluctuation map. This allowed us to check also more deeply the correctness of our lensing modeling, giving attention to the possible presence of artifacts due to interpolation procedure or other bugs. The results are displayed on Figure 5.12. As shown, we obtained artifacts-free images and plausible strong lensing effects in the lensed temperature map. Then, we used these lensed maps to recover the input lensing potential. The estimator in this strong regime will work as well, allowing us to have other insights about the estimator and the weak lensing assumption Eq. (4.1). The results for a SIS model with $\theta_E = 6.9$ arcsec and $\Omega_s = 5^\circ \times 5^\circ$ is shown in Figure 5.13. We can see that the shape of the recovered potential is very distorted with respect to the input one.



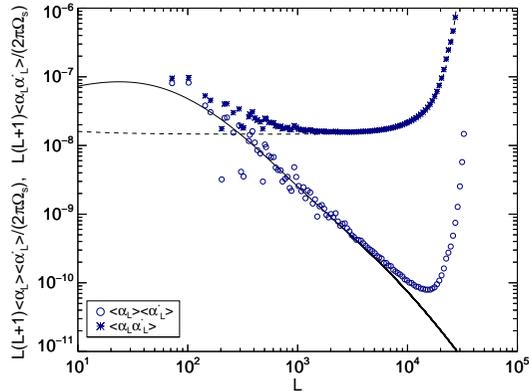
5.14.1 Averaged estimator for case (a).



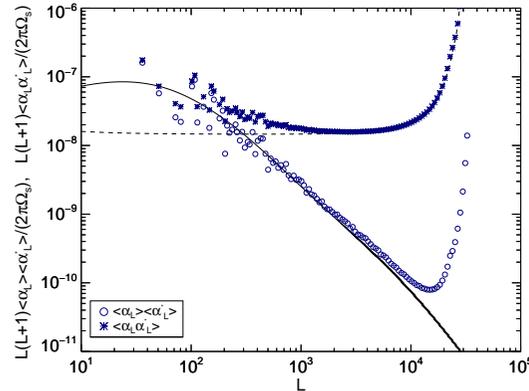
5.14.2 Averaged estimator for case (b).



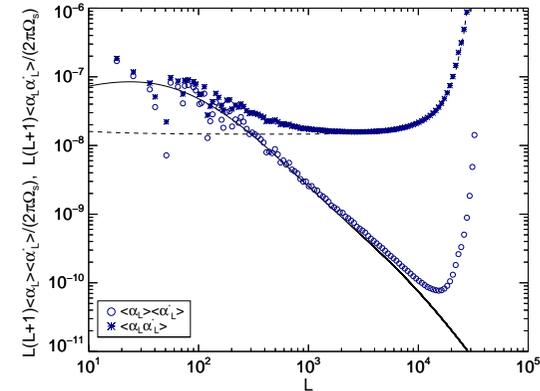
5.14.3 Averaged estimator for case (c).



5.14.4 Recovered power spectra for case (a)



5.14.5 Recovered power spectra for case (b)



5.14.6 Recovered power spectra for case (c)

Figure 5.14: *Top*: reconstructed images from $N_{\text{Sim}} = 1000$ realisations of the input 21 cm source box for the three cases listed in Table 5.2. *Bottom*: the power spectrum of the overall estimator image is displayed in the underlying row for every case, together with the averaged power spectrum over N_{Sim} . For every case we used a SKA2-Low configuration with choices for observational time and bandwidth denoted as R2 in Table 5.1.

The slope of the potential is recovered correctly up to a certain scale, in which the approximation Eq. (4.1) breaks down, and contributions from higher order terms are needed to recover the input potential. At these scales the lensing gradient is not small anymore, and such higher-order terms become more and more important, irrespectively of the magnitude of the temperature gradient. Moreover the recovered potential is more flat in the center, because of the level set by the estimator reconstruction noise.

Most importantly, we checked that our estimator is unbiased after a large number of realisations, *i.e.* the estimated potential field is equal to the true one, explicitly $\phi_L = \langle \hat{\phi}_L \rangle$, keeping the same realisation of the input lensing field. This has been achieved by generating N_{sim} realisations of the input 21 cm source field. The estimator has then been produced for each realisation as described in Section 5.5.1, being careful of generating always k_p different random realisations for thermal noise generation within every single source realisation. Then we summed these estimators to produce the total one $\sum_{N_{\text{sim}}} \hat{\phi}_{L,\text{sim}}$ shown in Figure 5.14. Here we displayed the summed estimator for the three cases listed in Table 5.2 and for $N_{\text{sim}} = 1000$. In every of them the recovered potential converges quite quickly towards the input one. The reader can compare these images with the input ones on Figures 5.6.1, 5.7.1, and 5.8.1. Of course case (c) provides a better result using less realisations than does case (a) because of the higher number of large scale modes and higher resolution in Fourier space.

Since this behavior is caused by the decreasing of the total reconstruction noise as N_{sim} increases, this test could also give a preliminary idea about the potentialities of a multi-band detection with bandwidths centered around several source redshifts. To show this, consider the bottom row of Figure 5.14, in which we show the recovered power spectra of the total estimator after N_{sim} realisations $\langle \hat{\alpha}_L \rangle \langle \hat{\alpha}_L^* \rangle / \Omega_s$ (purple circles) and the one resulting from the sum of every individual recovered power spectrum $\langle \hat{\alpha}_L \hat{\alpha}_L^* \rangle / \Omega_s$ (purple stars). The former converges to the input power spectrum Eq. (4.4), while the estimator noise is averaged out when several realisations are added and it is decreased by a factor N_{sim} . The latter recovered power spectrum is instead converging to the sum of lensing signal plus estimator noise, as expected from Eq. (4.15), and the sample variance error within every considered L -bin decreases with respect to the one displayed in Figure 5.9. In order to better appreciate the behavior pictured by the purple circles shown in the bottom row of Figure 5.14, we plot the recovered power spectrum of the total estimator for several realisations up to $N_{\text{sim}} = 100$ in Figure 5.15. Here the noise statistically decreases of a factor N_{sim} . With only 10 bands we should be able to have a larger number of modes above the noise signal in the intermediate range $100 \leq L \leq 1000$.

Therefore the same behavior could be expected when multiple 5 MHz bands are stacked together to fit a given redshift range, as anticipated by the numerical estimates made in (Pourtsidou & Metcalf, 2015) for their SKA-Low model. Remember that our multi-realisation approach explored in this section relies on a few crude approximations: the estimator noise will be slightly different from band to band, being higher for high redshifts, because noises, sources and L_{cut} depend on z . Moreover, the single estimators need to be renormalised by the estimator reconstruction noise in that band and weighted by the total-

band estimator. This topic will be discussed in Section 5.5.5.

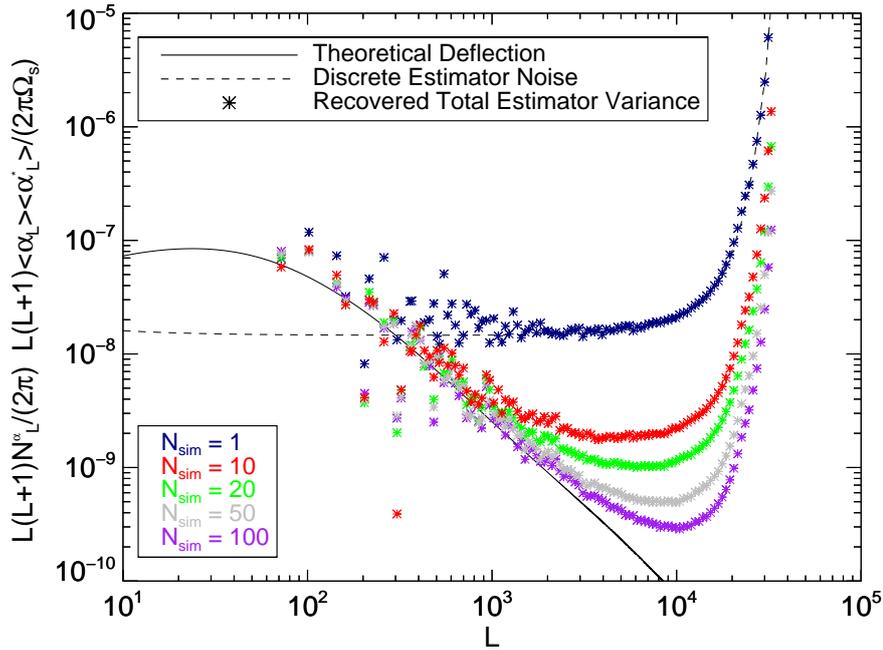


Figure 5.15: The recovered Estimator power spectrum as the number of source and noise realisations increases up to $N_{\text{sim}} = 100$. This plot is produced for the small map cases, with $\Omega_s = 5^\circ \times 5^\circ$ and using our SKA2-Low parameters.

5.5.4 Performance for different SKA configurations

As already stated in Section 4.3, the estimator noise level and, hence, reconstruction's performance, depend on several factors. Apart from the effect of redshift-space distortion and non-linear contributions in the source which tend to increase the signal-to-noise especially at smaller scales, a central role is played by the considered SKA configuration. For a fixed source redshift like $z_s = 8$, changes in the estimator noise level are mainly due to different thermal noise levels for each telescope model, causing the reconstruction noise to be considerably sensitive to changes in covering fraction, observational time, and observational bandwidth. A careful reader could note that also varying L_{max} has a big influence in determining the estimator noise level, but this effect is important only if we vary the source redshift, as we will see in the next section, or the telescope baseline length (which is considered fixed in this work).

Here we model some other possible SKA configuration beyond the one used to get the results in Section 5.5.1, in order to have a concrete idea of what a 25 square degrees survey with SKA1-Low and SKA2-Low could detect and how it depends on telescope parameters. We will always adopt the uniform Eq. (5.6) for the moment, postponing the discussion for a non-uniform noise power spectrum in Section 5.6.

The discrete estimator noise, Eq. (4.22), is plotted in Figure 5.16 for the SKA1-Low and SKA2-Low telescope designs and the survey strategies described in Table 5.1.

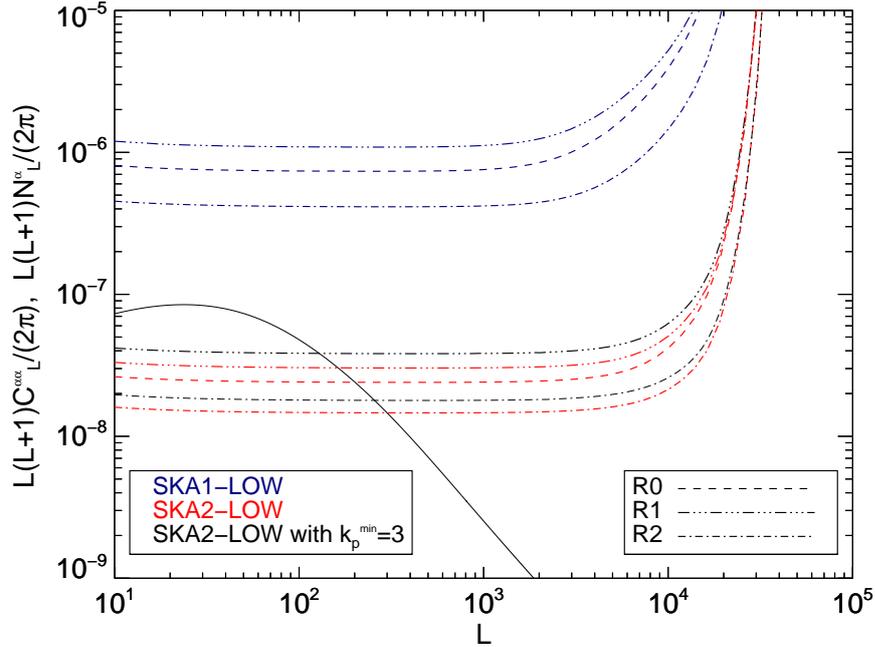


Figure 5.16: The discrete estimator noise for SKA1-Low (blue), SKA2-Low (red) and SKA2-Low with $k_p^{\min} = 3$ (black), with choices for observation time and bandwidth listed in Table 5.1. The simulated sky area is $\Omega_s = 5^\circ \times 5^\circ$ and the beam has a resolution of 1.15 arcmins. The R0 configuration is on dashed lines, the R1 one is on dashed-triple dot lines, and the R2 is on dashed-dot lines.

High-fidelity maps are possible when the noise in this plot is below the expected signal power spectrum. For every SKA1-Low case the noise is well above the signal. This means that for a single frequency band detection modeled with the thermal noise introduced in Section 5.4 is practically impossible to get high-fidelity images of the reconstructed lensing mass distribution. However, always considering a single bandwidth, for the SKA2-Low configurations the situation is much more promising. High fidelity images should be possible for the SKA2-Low experiment even in our worst case R0, whose noise level crosses the signal at $L \sim 200$, or for the SKA2-Low R1 case. The situation will be different if multiple frequency bands are stacked to fit a given redshift range. This will be treated in next section. In Figure 5.17 we can see the fidelity for SKA2-Low R2 and R0 telescope models. We can see that the quality of the reconstructed potential image is better in the first case, but there is not a big difference between the two models.

Huge contaminations can come from foregrounds (mainly due to synchrotron radiation coming from our Galaxy and others) but it is possible to remove them, as they are smooth in frequency. Foreground removal techniques will make the first few k_p modes unusable (McQuinn et al., 2006; Liu & Tegmark, 2012). The exact number of k_p modes will depend on the specific foreground removal technique. In Figure 5.16 the impact of removing the first three modes ($k_p^{\min} = 3$) is shown. The noise is increased by omitting these modes, but not drastically. In general, the more serious the foreground contamination is, the higher is the number of k_p -modes one has to remove. This is an issue that

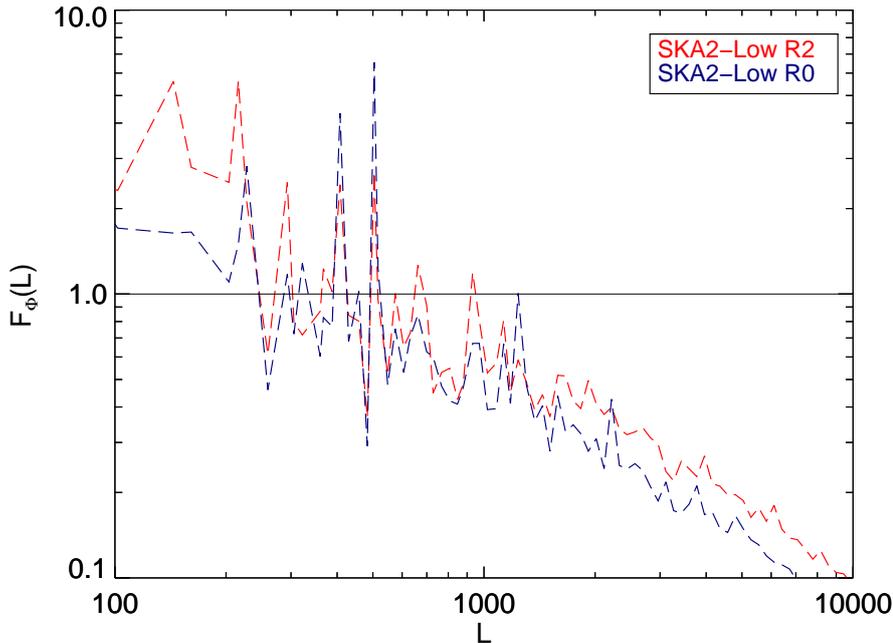


Figure 5.17: The fidelity of the reconstructed lensing potential images as a function of the multipole mode L focused in the range $100 \leq L \leq 1000$. The simulated sky area is $\Omega_s = 5^\circ \times 5^\circ$ and the beam has a resolution of 1.15 arcmins. Red and blue dashed line are the results for SKA2-Low R2 and SKA2-Low R0 survey strategies respectively. The straight line with $F_\phi(L) = 1$ helps to distinguish modes with good fidelity from the ones with bad fidelity.

needs to be investigated more deeply in the future including a specific foreground model and adopting a removal technique in our simulation code.

5.5.5 Multi-Band Reconstruction

We have previously stated that there is the intriguing possibility to stack several estimators computed at different frequency (redshift) bands. Our findings in Section 5.5.3 suggest that the combination of multiple frequency bands can aid the reconstruction of the underlying lensing potential.

Noise weighted total-band estimator

In this section we will introduce a combined multi-band noise weighted estimator that can be applied to our simulated maps extending what has already been done in Section 5.5.1. Each bandwidth $\Delta\nu$ is centered around a given source redshift z_c within the range Δz . Pourtsidou & Metcalf (2015) calculated the lensing reconstruction noise stacking 10 frequency bins of 8 MHz bandwidth spanning the redshift range $z_c \simeq 6.5 - 11$ using SKA1-Low and SKA2-Low parameters. This noise is lower than the noise obtained using a single band

(Figure 5.16), because the total estimator noise is

$$\mathcal{N}_L^{\text{tot}} = \left[\sum_{\nu} \frac{1}{\mathcal{N}_{L,\nu}^{\hat{\phi}}} \right]^{-1}. \quad (5.9)$$

This behaviour can be understood from the multi-realisation study performed in Section 5.5.3: as we stack frequency bands, the 21 cm source signal will be averaged out together with the thermal noises, and the estimator noise will go down by a factor N_{ν} , the number of stacked frequency bands. This is well shown in Figure 5.15.

The combined discrete quadratic estimator is hence noise-weighted, namely

$$\hat{\phi}_L^{\text{tot}} = \mathcal{N}_L^{\text{tot}} \sum_{\nu} \frac{\hat{\phi}_{L,\nu}}{\mathcal{N}_{L,\nu}^{\hat{\phi}}}, \quad (5.10)$$

where each single-band estimator Eq. (4.24) contributes for every frequency band and every multipole to the total band one.

Redshift Dependence

The redshift dependence of L_{cut} and thermal noise needs to be taken into account, when multiple frequency/redshift bands are stacked. The higher the redshift, the higher the thermal noise and the lower L_{cut} will be. This is shown in Figure 5.18, where the discrete estimator noise is substantially varying in the redshift range from $z_c = 6.5$ to $z_c = 12$, for a $5^\circ \times 5^\circ$ survey with $l_{\text{min}} = 72$. The adopted telescope model is SKA2-Low R2, as listed in Table 5.2, and $k_p^{\text{max}} = 20$.

It is assumed in this work that the lensing signal is not substantially varying between the first source redshift and the last one. Considering the deflection field power spectrum computed at different redshifts in Figure 5.18, we see that this approximation is valid across a substantial redshift range. In this range it is also assumed that Eq. (4.11) is valid and that our optimal estimator is derived for a Gaussian field.

In order to avoid aliasing, the Nyquist frequency is set to $2.5L_{\text{cut}}(z_c = 6.5) \approx 39690$ corresponding to a resolution of ≈ 23 arcsecs. For a higher redshift we keep this maximum frequency and we will vary $\Delta b = b\Delta\theta$. The considered redshift range $z_c = 6.5 - 12$ will correspond to beams with resolutions going from 57.7 arcsecs to 1.66 arcmins.

Note that because of Equation 5.9, the upper limit of this redshift range will not influence the total estimator noise level, since the estimator noise for $z_c \gtrsim 11$ turns to be considerably high. The discussion over the adopted lower limit will be postponed to Section 5.5.5.

Stacking Bands

The number of stacked bands depends on the adopted bandwidth which, for a given central redshift z_c , corresponds to a redshift interval $\Delta z = (1 +$

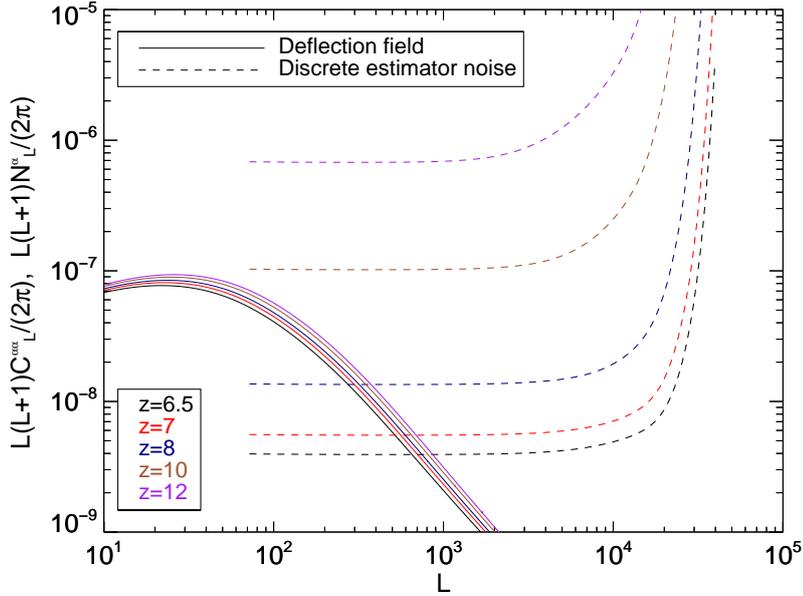


Figure 5.18: The deflection field discrete estimator noise (dashed lines) is shown for several redshifts, from $z_c = 6.5$ to $z_c = 12$, for a $5^\circ \times 5^\circ$ survey with $l_{\min} = 72$. The adopted telescope model is an SKA2-Low R2, as listed in Table 5.2, and $k_p^{\max} = 20$. We can see that the lensing signal (solid lines) does not change considerably within this redshift range, while this is not true for estimator noise. The higher is the redshift, the higher the noise will be, as $L_{\text{cut}}(z)$ decreases and thermal noise increases. We have assumed that the Universe is completely un-ionized in all cases.

$z_c)^2 \Delta\nu / \nu_{21}$. Starting from the first band centered at z_c^{\min} , the lower central redshift limit, the following band is found by decreasing the central frequency according to $\nu_c(z'_c) = \nu_c(z_c) - \Delta\nu$, where $\nu(z) = \nu_{21}/(1+z)$. The new central frequency is thus extended between the limits $\nu_{\max, \min} = \nu'_c \pm (\Delta\nu/2)$. These correspond to a new redshift interval $\Delta z = z_{\max} - z_{\min}$ with $z_{\max, \min} = (\nu_{21}/\nu_{\min, \max}) - 1$. Hence the new central redshift is $z'_c = (z_{\min} + z_{\max})/2$. For example, considering $\Delta\nu = 5$ MHz and $6.5 \leq z_c \leq 12$, we can stack 17 bands in the frequency range $\nu_c = 189.4 - 109.4$ MHz.

Consider that a R0 configuration with $\Delta\nu = 8$ MHz will stack 10 bands within the range $z = 6.5 - 12$ (in reality the maximum redshift with this bandwidth is $z \approx 11.1$). If we use a thinner bandwidth, like $\Delta\nu = 3$ MHz, this number increases to 28, reaching 82 stacked bands for $\Delta\nu = 1$ MHz. The frequency band can be chosen as thin as possible until effects due to correlations between different Fourier modes show up. In Metcalf & White (2009) it was found that the correlation between estimators at different frequency bands is not significant if $\Delta\nu \sim 1$ MHz. This means that the statistical properties of the 21 cm radiation field and noise can be assumed as constant within a band. On the other hand, a very thin band increases the thermal and estimator noises. Such a thin band poses limits on the maximum number of k_p modes we can detect within a given band as well, since $k_p^{\max} = \Delta\nu/\delta\nu$, where $\delta\nu$ is the frequency resolution of one channel. The values of the adopted bandwidths in considered survey strategies for this work are listed in Table 5.1.

SKA2-Low and SKA1-Low results

In this section we will consider only the case (a) present in Table 5.2, namely a $5^\circ \times 5^\circ$ survey with $l_{\min} = 72$, unless otherwise stated. Figure 5.19 shows the total-band discrete estimator noise Eq. (5.9) for the survey strategies listed in Table 5.1 for both SKA1-Low and SKA2-Low configurations.

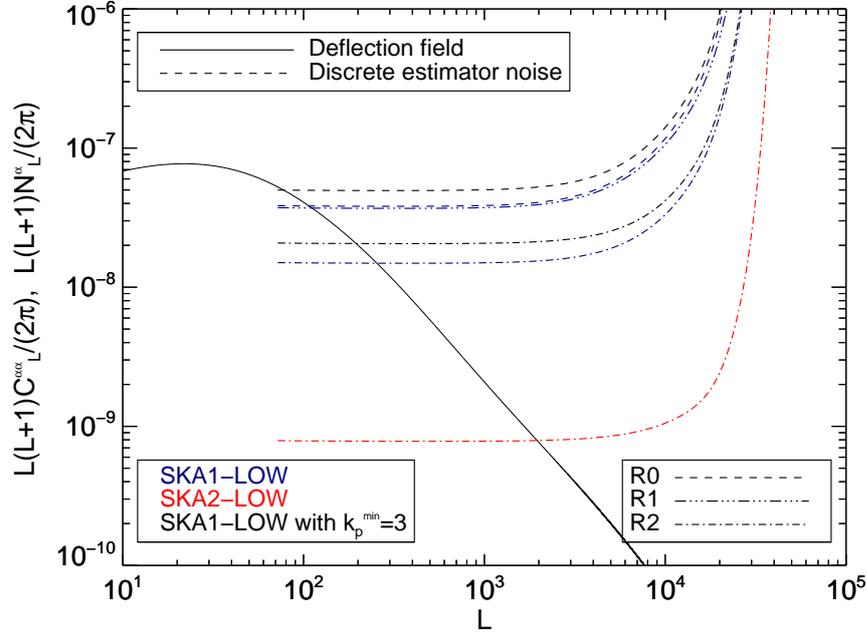


Figure 5.19: The discrete estimator noise is shown for the survey strategies listed in Table 5.1, with contributions coming from 17 stacked frequency bands that span the range $\nu_c = 189.4 - 109.4$ MHz, for a $5^\circ \times 5^\circ$ survey with $l_{\min} = 72$ and $k_p^{\max} = 20$ within each band. Blue, red, and black lines are for SKA1-Low, SKA2-Low, and SKA1-Low with $k_p^{\min} = 3$ models respectively. Dashed, dashed-triple dot, and dashed-dot denote R0, R1, and R2 survey strategies respectively.

SKA2-Low configurations already give good results in the single-band case, as seen in Section 5.5.1, and for the multi-band the situation is further improved. Considering R2 survey strategy, the noise decreases by a factor ≈ 20 , as shown in Figure 5.19, when 17 frequency bands spanning the range $\nu_c = 189.4 - 109.4$ MHz are stacked up. But the most interesting result comes from SKA1-Low detections, which now can provide hi-fi images of the reconstructed lensing potential with fidelity comparable to the one of a SKA2-Low single-band experiment if R2 survey strategy is adopted. The total-band discrete estimator noise levels for the SKA1-Low R0 and R2 models are also displayed in Figure 5.19, with the first three k_p modes have been excluded in order to take into account the impact of foreground removal techniques.

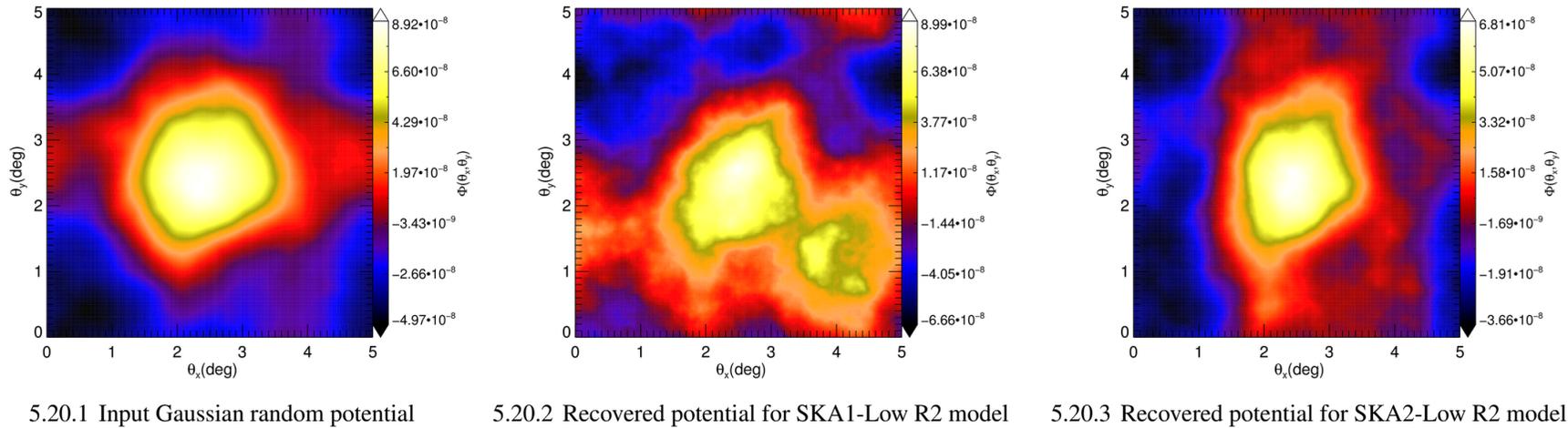


Figure 5.20: The input Gaussian random field used for the multi-band detection on the left, with $N_{\text{side}} = 780$ and $\Omega_s = 5^\circ \times 5^\circ$, and generated from a deflection field power spectrum computed at $z = 8$. On the middle and on the right, the recovered potentials obtained with SKA1-Low R2 and SKA2-Low R2 models respectively. The images have been produced by stacking 17 bands in the redshift range $z_c = 6.5 - 12$. In both cases we have excluded the first k_p modes due to foreground subtraction ($k_p^{\text{min}} = 3$).

We see that this does not avoid in principle the reconstruction of the lensing potential with high fidelity at least for our best survey strategy R2. This can not be true for R0 and R1 cases, whose modes that are over the noise level are few. So, an increment in the observational time, which would lead us to R2 configuration, would have big benefits for multiband measurements. It is therefore crucial to understand the exact number of k_p modes to remove, since this might fundamentally determine the detectability of the lensing signal for SKA1-Low experiments. Thus, specific foreground removal techniques need to be discussed with more details. This purpose will be accomplished in future works.

Note that the SKA1-Low R0 and R1 have comparable noise level because of reconstruction noise behavior at high redshifts, as previously discussed in Section 5.5.5.

We proceeded to compute the reconstructed images by following the same steps described in Section 5.5.1 for each frequency band, being careful to keep fixed the realisation for the 21 cm source. So, with the discrete estimator in one band $\Phi_{L,v}$ given by Eq. (4.24), we applied Eq. (5.10) to get the total multiband discrete estimator for the potential field. The recovered images for a SKA1-Low and SKA2-Low experiments are shown in Figure 5.20, both computed for the strategy R2 with $k_p^{\min} = 3$ and $k_p^{\max} = 20$, together with the input potential image whose resolution is $\Delta\theta \approx 23$ arcsec. The final resolution of the recovered images will depend mainly on the one set by the beam of the highest redshift band, because some modes belonging to other bands can be smoothed depending on whether they are found before or after $L_{\text{cut}}(z_c^{\max})$. In general this is equal to the beam resolution which produces the corresponding total reconstruction noise level.

The recovered square amplitude of Fourier modes is recovered in Figure 5.21, where the star points follow the multi-band analogous of Eq. (4.15) including the beam, namely

$$\left\langle \hat{\phi}_L^{\text{tot}} \left(\hat{\phi}_L^{\text{tot}} \right)^{\star} \right\rangle = \Omega_s \left(C_L^{\phi\phi} + \mathcal{N}_L^{\text{tot}} \right) \quad (5.11)$$

confirming the behavior studied in Section 5.5.3.

The fidelity for the two models considered above is shown in Figure 5.22. There is a big difference in the number of modes with Fidelity bigger of one in the range $100 \leq L \leq 1000$, because of the multi-band estimator noise behavior pictured in Figures 5.19 and 5.21. In fact, the fraction of modes with Fidelity bigger than one in the range $100 \leq L \leq 1000$ is $f_{100 \leq L \leq 1000} = 1.48 \times 10^{-4}$ for SKA1-Low R2 model, while this is five times bigger for SKA2-Low R2 model, $f_{100 \leq L \leq 1000} = 5.64 \times 10^{-4}$. The total number of modes used to reconstruct the images in Figure 5.20 is the same for both models, namely $n_{\text{tot}} = 304980$.

We conclude this section by mentioning that increasing the survey area to (b) or (c) cases in Table 5.2 would improve the results pictured in Figure 5.19 by adding contributions from more modes over the noise level.

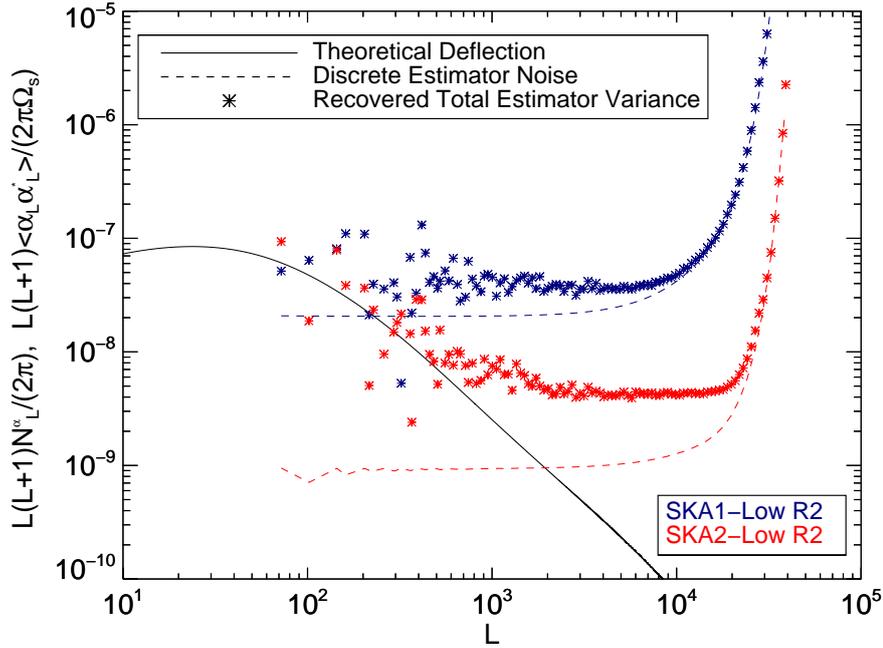


Figure 5.21: The recovered estimator power spectra (star points), as computed by stacking 17 bands for SKA1-Low R2 (blue) and SKA2-Low R2 (red) models in the redshift range $z_c = 6.5 - 12$. Both SKA configurations consider the effect of foreground subtraction by excluding the first 3 k_p modes. The survey area is $\Omega_s = 5^\circ \times 5^\circ$.

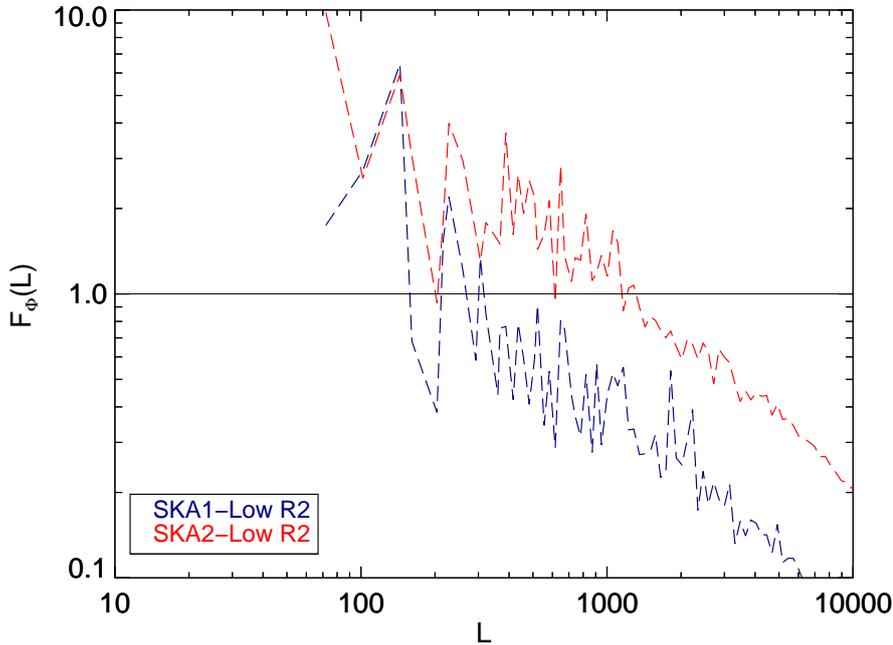


Figure 5.22: The fidelity of the multi-band reconstructed lensing potential images as a function of the multipole mode L . The simulated sky area is $\Omega_s = 5^\circ \times 5^\circ$ and the investigated redshift range is $z_c = 6.5 - 12$. Red and blue plots correspond to SKA2-Low R2 and SKA1-Low R2 survey strategies respectively, with $k_p^{\min} = 3$ because of foreground subtraction. The straight line with $F_\phi(L) = 1$ helps to distinguish modes with good fidelity from the ones with bad fidelity.

Limits on Lower Central Redshift

As stated in Section 5.5.5, the upper bound on the considered redshift range does not considerably affect the signal-to-noise because the single reconstruction noises for $z \gtrsim 11$ are very high, leading to negligible contributions in the ν -sum of Eq. (5.9). The results of multi-band approach within a SKA1-Low experiment depends mainly on the first central redshift that is chosen to define our total band. This would happen if EoR ended at earlier redshifts or if EoR is so patchy as to make hydrogen not uniformly ionized. If $z_c^{\min} = 7.5$ we should exclude 4 frequency bands from the ones considered in Section 5.5.5 for R1 and R2 strategies, while we should exclude 3 bands for R0 case. This would lead to an increase of reconstruction noise from a factor ≈ 6 for SKA1-Low R0 with $k_p^{\min} = 3$ to a factor ≈ 6.7 for SKA1-Low R2 with $k_p^{\min} = 3$. These values would make impossible the high-fidelity reconstruction for SKA1 cases with $k_p^{\min} = 3$. For SKA2-Low R2 the total-band reconstruction noise increases only by a factor ≈ 3.6 , value which keeps ensuring high-fidelity reconstruction of the underlying lensing mass distribution.

5.5.6 Lensing Power Spectrum Measurement

We are not interested in giving analytic estimates of the recovered power spectrum or provide forecasts on cosmological parameters in this work, but it is possible to investigate if some information about the power spectrum could be recovered even from a $5^\circ \times 5^\circ$ field.

As already mentioned in Section 5.5.5 and as can be seen in Eq. (4.25), if the sky fraction f_{sky} is too low, we might have cosmic variance dominated errors. Moreover, a high reconstruction noise would increase these errors, compromising the measurement of the lensing power spectrum. If we consider our single band lensing measurements for the SKA1-Low models, we see that the level of the reconstruction noise is too high to provide a significant power spectrum detection. As stated also by Pourtsidou & Metcalf (2014), the larger observed fraction of the sky planned for SKA-Mid will greatly improve this measurement, since the error increases as $f_{\text{sky}}^{-1/2}$ and the signal is detected with a much higher signal-to-noise respect to SKA1 and SKA2-Low phases. This will allow for a competitive estimate of cosmological parameters from such high-fidelity images.

We obtained better results when multiple frequency estimators are stacked together within a given redshift range. In fact, in Section 5.5.5 we have seen that for SKA2-Low experiments we can achieve a very low level for the estimator reconstruction noise when multiple frequency bands are stacked up and used simultaneously, even excluding some k_p modes because of foreground subtraction. The noise is lowered also in the SKA1-Low case, allowing for a good quality imaging at least for the survey strategy R2, and we can see that this causes improvement in the statistical detection of the power spectrum as well.

Figure 5.23 shows the deflection power spectrum with measurement error bars for $\Delta L = 72$ considering a SKA1-Low R2 model (red) and a SKA2-Low R2

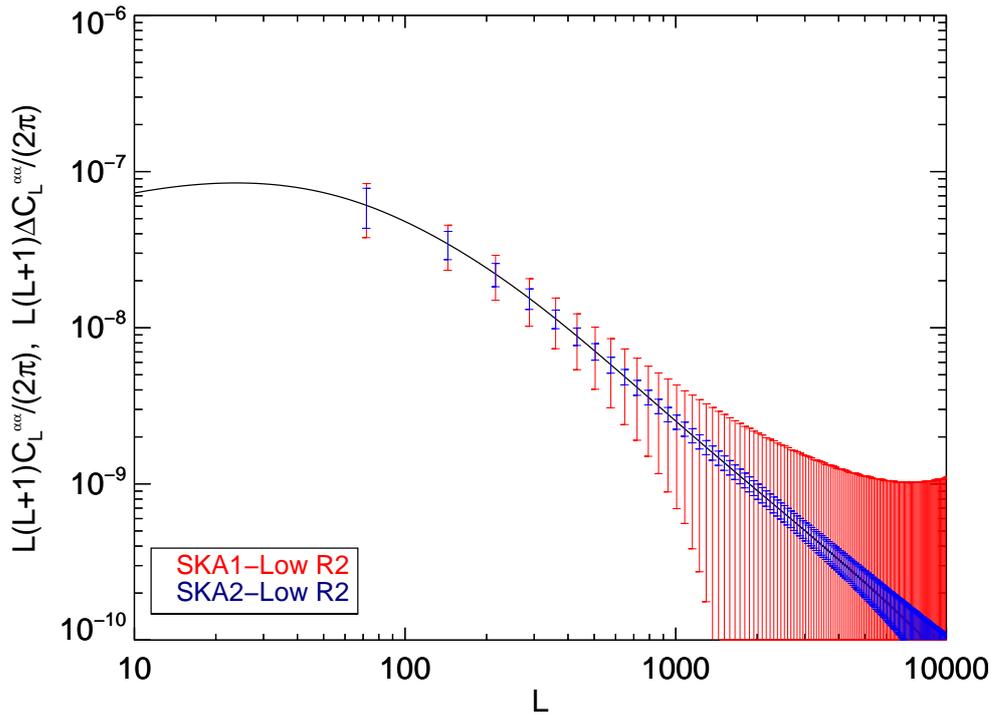


Figure 5.23: Deflection field power spectrum for $z = 8$ for a SKA1-Low R2 (red) and a SKA2-Low R2 (blue) telescope model, $\Omega_s = 5^\circ \times 5^\circ$ and $f_{\text{sky}} = 0.0006$. The error bars are computed including the total-band reconstruction noise found by stacking 17 frequency bands in the range $z_c = 6.5 - 12$ and considering $k_p^{\text{tot}} = 3$.

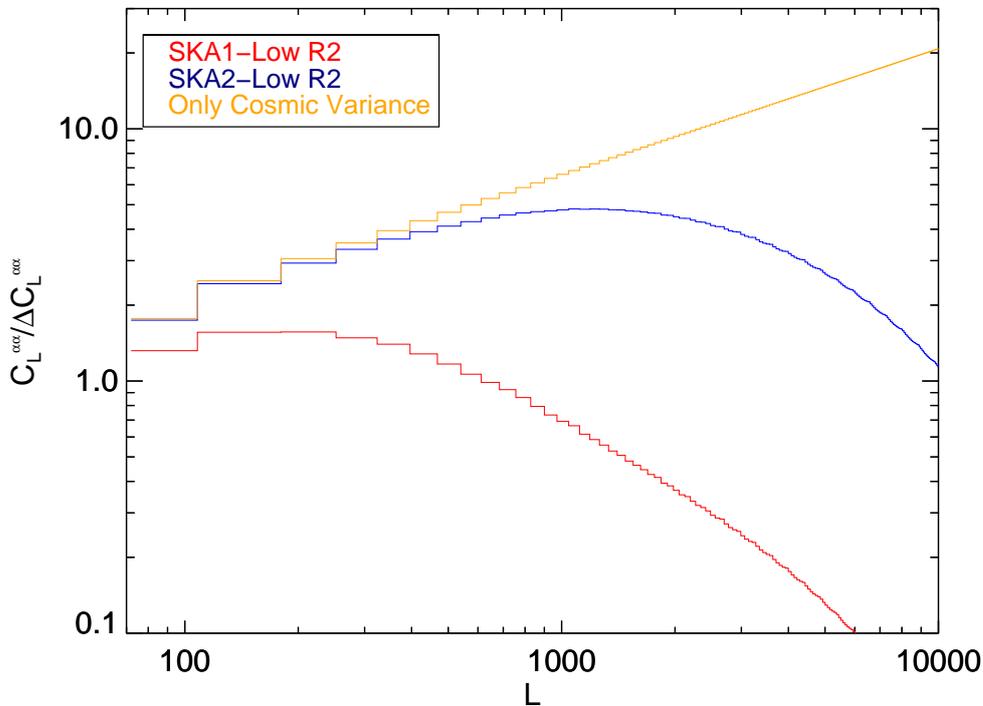


Figure 5.24: The fractional error Eq. (5.12) corresponding to the curves in Figure 5.23. The orange line indicates the fractional error due only to sample variance, namely with $\mathcal{N}_L^{\hat{\alpha}} = 0$, Eq. (5.13).

model (blue), including the multi-band estimator reconstruction noise computed in the previous section, *i.e.* by stacking 17 bands in the redshift range $z = 6.5 - 12$ and assuming $\Omega_s = 5^\circ \times 5^\circ$ ($f_{\text{sky}} = 0.0006$). We have also excluded the first three k_p modes. While the accuracy is comparable in the sample variance dominated regime at low L 's, the SKA2-Low R2 low reconstruction noise allows for much better estimates at all $L > 200$ than noisy SKA1-Low R2 model. We note that the lensing reconstruction noise level for SKA1-Low obtained by stacking multiple bands is analogous to the one obtained by considering a single-band measurement for SKA2-Low. So, a good measurement of the power spectrum, for a single-band detection with SKA-Low, can be obtained considering the phase 2, in order to decrease the noise part of the error expression Eq. (4.25).

To understand how accurate the power spectrum measurement is using these two models, we can consider the fractional error from Eq. (4.25), namely

$$\frac{C_L^{\alpha\alpha}}{\Delta C_L^{\alpha\alpha}} \approx \sqrt{\frac{2L+1}{2} f_{\text{sky}} \Delta L} \left(1 + \frac{N_L^{\hat{\phi}}}{C_L^{\alpha\alpha}} \right)^{-1}, \quad (5.12)$$

where the ratio of the power spectra is the inverse of the signal-to-noise ratio. For negligible estimator reconstruction noise we obtain the sample variance fractional error limit

$$\frac{C_L^{\alpha\alpha}}{\Delta C_L^{\alpha\alpha}} \approx \sqrt{\frac{2L+1}{2} f_{\text{sky}} \Delta L}, \quad (5.13)$$

These are shown in Figure 5.24 for the two SKA-Low models previously considered. We can see how the SKA2-Low R2 result has a comparable order to the sample variance alone result for $L \lesssim 1000$ and does much better than SKA1-Low R2 model even in the reconstruction noise dominated regime at high L 's, being more accurate by about an order of magnitude.

Power spectrum measurements can also benefit from detected signal coming from different sky patches, even for the smaller map case (a). Using the signal taken from different patches of the sky increases the Fourier mode statistics. Cosmic variance error in Eq. (4.25) could be reduced and a better $C_L^{\alpha\alpha}$ detection can be realised especially at intermediate scales, namely $100 \leq L \leq 5000$.

5.5.7 Cluster Detection

Another possible application for 21-cm lensing consists in detecting galaxy clusters lensing signal. To investigate this opportunity we generated a deflection field using the GLAMER⁶ library (Petkova et al., 2014; Metcalf & Petkova, 2014). This is a C++ library for doing gravitational lensing simulations using the output of cosmological simulations or analytic lens models or combinations of them. We used it to generate a NFW halo profile with density (Navarro et al., 1997)

$$\rho(r) = \frac{\rho_s}{r/r_s (1 + r/r_s)^2}, \quad (5.14)$$

⁶<http://glenco.github.io/glamer/>

where the scale density ρ_s is the normalisation of this profile and r_s is a scale radius. These quantities are often described in terms of the concentration parameter $c = r_{200}/r_s$, with r_{200} being the radius of the sphere in which the average density is 200 times the critical density and the enclosed mass is M_{200} . Its value is

$$r_{200} = 1.63 \times 10^{-2} \left(\frac{M_{200}}{h^{-1} M_{\odot}} \right)^{1/3} \left[\frac{\Omega_0}{\Omega(z)} \right]^{-1/3} (1+z)^{-1} h^{-1} \text{ Kpc}. \quad (5.15)$$

The mass of the cluster is linked to the concentration parameter via the relation $M = 4\pi r_s^3 \rho_s [\ln(1+c) - c/(1+c)]$. The lensing potential produced by the NFW profile is

$$\Phi_{NFW}(\theta) = 4\rho_s r_s \Sigma_{cr}^{-1} g(\theta/\theta_s), \quad (5.16)$$

where Σ_{cr} is the critical surface density, $\theta = r/\mathcal{D}(z)$ and $\theta_s = r_s/\mathcal{D}(z)$. The function $g(x = \theta/\theta_s)$ is defined as

$$g(x) = \frac{1}{2} \ln^2 \frac{x}{2} + \begin{cases} 2 \arctan^2 \sqrt{\frac{x-1}{x+1}}, & (x > 1) \\ -2 \operatorname{arctanh}^2 \sqrt{\frac{1-x}{1+x}}, & (x < 1) \\ 0, & (x = 1) \end{cases}. \quad (5.17)$$

Following work by Giocoli et al. (2014) and Sereno et al. (2014), we have simulated the deflection field produced by a plausible galaxy cluster placed in the centre of our lensed 21 cm radiation map with mass $M = 10^{15} M_{\odot}$ and concentration $c = 7$. This cluster is placed at $z = 0.5$, while the source is at $z_s = 8$.

This lensing source has been used to deflect our simulated 21 cm intensity maps, as discussed in Section 5.2. Following the procedure described in Section 5.5.1 and modeling a SKA2-Low R2 experiment, we applied these deflected maps to the estimator Eq. (4.24). We find that the NFW cluster under consideration (with a few arcseconds Einstein radius) is basically undetectable because the recovered signal is totally consistent with the estimator reconstruction noise. Analysing the input deflection field power spectrum, we found that, even for a multi-band detection constructed by stacking bands from $z_c = 6.5$ to $z_c = 12$, this is well below the estimator reconstruction noise level by four orders of magnitude. This result agrees with the one obtained in (Kovetz & Kamionkowski, 2013) for a lower redshift ($z = 7$).

It would be interesting to study this detection at lower redshift such as $z \sim 1 - 3$, where $x_H \neq 1$ and point source signal represents an important contribution to 21 cm source. For this reason we will need to take into account non-negligible Poissonian source terms in our Estimator. We could indeed place in random positions more realistic clusters in our simulated map, in order to detect the total signal coming from them, but big improvements are not expected, since at those redshifts the reconstruction signal-to-noise is lower for each mode. This case will be studied in a future work.

5.6 Adopting a More Realistic Thermal Noise Model

The results obtained for a uniform noise model through Section 5.5 are strongly dependent on the adopted thermal noise. We have just seen that in this case a positive imaging of the reconstructed mass is highly disfavoured for SKA1-Low models. Using a more realistic and updated model for the SKA-Low thermal noise will improve in general the results, since the modes mainly involved in the reconstruction will have less noise because of the core region for station distribution. We will see that high-fidelity images are indeed detectable for a multi-band experiment in a much clearer way with respect to what obtained using a uniform thermal noise. On the other hand the 21 cm maps will present a higher noise at smaller scales.

5.6.1 Non-Uniform Thermal Noise Power Spectrum

A pair of elements in an interferometer separated by a baseline of length d will measure a visibility $V(\mathbf{U}, \nu)$, where \mathbf{U} is the vector in visibility space and $U = |\mathbf{U}| = d/\lambda$. The resolution in visibility space defines the Field of View (FoV) of the telescope, that is namely $U_{\min} = d^2 U = 1/\Omega_s \sim D_{\min}^2/\lambda^2$, with D_{\min} the interferometer element diameter which in our case is a station containing a certain number N_{ant} of antennae. The visibility space is another Fourier dual space, and its relation the multipole space is $\mathbf{U} = \mathbf{l}/(2\pi)$, so that $(2\pi)^2 d^2 U = d^2 l = L_{\min}^2$. The maximum observable visibility is hence set by the baseline maximum length $U_{\max} = L_{\text{cut}}/(2\pi)$.

We define the noise power spectrum in visibility space for an interferometer in the Rayleigh-Jeans limit by averaging all visibilities falling in one visibility space resolution, namely

$$C_U^N = \left(\frac{\lambda^2(z) T_{\text{sys}}}{A_{\text{eff}}} \right)^2 \frac{d^2 U}{N_{\text{pol}} \Delta \nu t_U}, \quad (5.18)$$

(Pourtsidou et al., 2015; Bull et al., 2015) where T_{sys} has been defined in Section 5.4, and represents the sky and/or the receiver temperature, $\lambda(z) = \lambda_{21}(1+z)$, and A_{eff} is the effective area of one station. A_{eff} is usually defined as $A_{\text{eff}} = \varepsilon \pi D_{\min}^2/4$, with ε the antenna efficiency, usually a number $0.7 \lesssim \varepsilon \lesssim 1$. N_{pol} is the number of polarisation channels, and they can be added incoherently. t_U is the observation time per visibility pixel,

$$t_U = d^2 U n(\mathbf{U}, \nu) t_p = d^2 U n(\mathbf{U}, \nu) \frac{t_o}{N_p} = d^2 U n(\mathbf{U}, \nu) \frac{t_o N_b \Omega_s}{S_{\text{area}}}. \quad (5.19)$$

Here we have included the possibility to perform a mosaicking of several sky patches, using different pointings $N_p = S_{\text{area}}/\Omega_s$ to observe a given sky area S_{area} , and using a certain number N_b of beams per station with $\text{FoV} \approx \Omega_s \approx \lambda/D_{\text{station}}$ observed within a time t_p per pointing. This permits to increase the number of independent measurements on scales smaller than the telescope FoV in a given total observational time $t_o = N_p t_p$, since the number of observed modes is increased by a factor $S_{\text{area}}/(N_b \Omega_s)$. Note that $S_{\text{area}} > N_b \Omega_s$. For EoR observations

we will consider $N_p = 1$, and so the observing time per pointing will coincide with the total observation time.

$n(U)$ is the averaged baseline number density (over a 24 hrs period), and it is usually a function of (U, ν) due to rotational invariance in visibility space given by a circularly symmetric baseline distribution (Villaescusa-Navarro et al., 2014). Its normalization will be frequency dependent, since $\int n(U, \nu) d^2U \sim N_{\text{stat}}^2/2$, with N_{stat} the number of stations forming the considered baseline. The computation of $n(U, \nu)$ at different redshifts is discussed in Appendix E.

An instrument like SKA-Low uses aperture arrays and can potentially be equipped with Phased Array Feeds (PAF) inherited from ASKAP, which allow for mosaicking of different patches of the sky increasing the number of available beams. For an aperture array, under a given critical frequency ν_c the effective area of the station will be constant, *i.e.* when the array is dense, while above ν_c it scales with frequency as

$$A_{\text{eff}}(\nu) = A_{\text{eff}}(\nu_c) \begin{cases} (\nu_c/\nu)^2 & \text{for } \nu \geq \nu_c \\ 1 & \text{for } \nu < \nu_c. \end{cases} \quad (5.20)$$

Moreover the FoV scales at any frequency as $\Omega_s(\nu) = \Omega_s(\nu_c) (\nu_c/\nu)^2$.

Hence, the final form of the non-uniform thermal noise power spectrum for a bandwidth $\Delta\nu$ centered around the redshift $z + 1 = \nu_{21}/\nu$ is

$$C_{l,\Delta\nu}^{\text{N}} = \left[\frac{\lambda^2(z)}{A_{\text{eff}}(\nu_c)} \left(\frac{\nu}{\nu_c} \right)^{2\gamma} \right]^2 \frac{T_{\text{sys}}^2(\nu)}{N_{\text{pol}} \Delta\nu t_o n[U = l/(2\pi), \nu]}. \quad (5.21)$$

A complete derivation of this expression and its link with the uniform version Eq. (5.6) is given in Appendix E.1. Basically the main differences with the uniform expression previously considered consists in the presence of the frequency dependence of the effective area, the polarisation channel contribution, and above all, the specification of the baseline density distribution which causes an important decrease for the multipoles mainly involved in reconstruction (see Figure E.2). On the other hand for modes $L > 2000$, the noise rises towards very high values.

Although the baseline design and specifications for SKA-Low have been described in the official SKA document made by Dewdney (2013), there is still a certain degree of freedom in assuming the instrumental outline, since this has not been finalised yet. As already stated in Section 1.1.1, the SKA specifications have been redefined because of budget issues, causing a rebaselining of the experiment phases (Bull, 2015), leading to a considerable reduction of the elements of a baseline (and causing also the elimination of SKA-Sur plans). In particular, SKA-Low halved the number of receiving stations, but its frequency sensitivity is basically unaffected with respect to the original design plan, because of the dense core array. We will include this effect by modifying the baseline density function as $n_{\text{descop}}(U, \nu) = n(U, \nu)/4$, and so increasing the thermal noise power spectrum level by a factor 4. This choice is connected to the uniform behavior of $n(U, \nu)$ which is proportional to the number of station squared.

5.6.2 SKA1-Low and SKA2-Low Specifications

We will consider a SKA1-Low design with $D_{\min} = 35$ m diameter station, and $N_{\text{stat}} = 433$ within a maximum baseline of $D_{\max} = 4$ Km. Observational time and bandwidth are defined in R0, R1, and R2 survey strategies specified in Table 5.1. The critical frequency is 110 MHz, and the values for effective area and FoV at ν_c are 925 m^2 and 27 deg^2 respectively. The system temperature will be $T_{\text{sky}} = 1.1 \times 60 [\nu/(300 \text{ MHz})]^{-2.55}$ for the sky thermal noise, while it will be $T_{\text{rcv}} = 40$ K for the thermal noise due to the detector itself. SKA-Low is also assumed to have two polarization channels, and a single-pointing observation is performed at EoR redshifts ($z = 8$, corresponding to an observed frequency of 157.82 MHz), so $N_p = 1$. The fiducial baseline density function has been provided at $z = 8$.

SKA-Low phase 2 has still to be formally defined, and there is not a well established design for it. In general through this work we will assume an aperture array telescope with a total collecting area that is four times the one expected for SKA1-Low. This will cause the thermal noise level to be a factor 16 lower. For the moment we do not include multiple beams for SKA2-Low as well, although such an experiment would have the possibility to use $N_b = 10$ beams simultaneously. Following (Bull, 2015), we increase the sensitivity of this instrument also by decreasing the receiver noise to 15 K, although this does not imply a big change due to sky noise supremacy at these frequencies.

5.6.3 Single-Band Results for non-Uniform Noise Model

Considering our SKA-Low specifications made in the previous section, for a single-band detection simulation at $z = 8$ we have a FoV of $3.6^\circ \times 3.6^\circ$ ($f_{\text{sky}} = 3.14 \times 10^{-4}$), and hence $\Delta l = 100$. $L_{\text{cut}} = 13237$, $L_{\text{Nyq}} = 2.5 \times L_{\text{cut}}$, for which $N_{\text{side}} = 468$.

Hence we proceeded exactly as we did in Section 5.5 to produce our simulated lensed 21 cm brightness temperature and noise maps. To have a preliminary idea of the reconstruction quality, we computed the discrete estimator reconstruction noise for the same survey strategies introduced in Table 5.1. The results are shown in Figure 5.32, where we have omitted the R0 results for a better cleanliness of the plot and because they produce similar results to R2 models. With respect to the uniform thermal noise case, there is improvement by nearly an order of magnitude for the SKA1-Low models, but this is not enough to ensure a positive detection using a single frequency band, since the noise is still well above the lensing signal.

For what concerns SKA2-Low models, the noise levels for the R2 models are practically unchanged respect to the ones provided using the uniform thermal noise. This happens because of saturation limit reaching due to high thermal noise at high L_s . Indeed, this estimator noise is more steep and diverges more quickly than its uniform analogous. On the other hand, there are slight improvements for the R1 models.

In general we expect to recover images which are heavily contaminated by high-scale noise, due to the smaller FoV and, above all, to a smaller number

of lensing modes that are under the noise level which is orders of magnitude higher than in the uniform case for $L > 10000$.

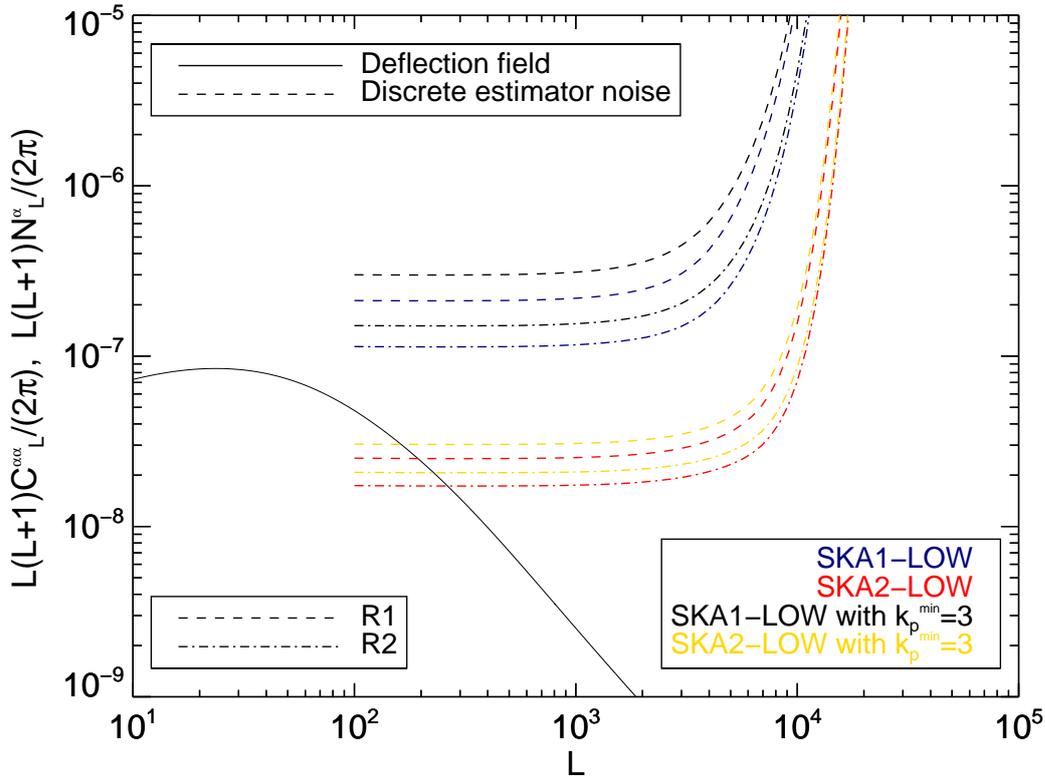
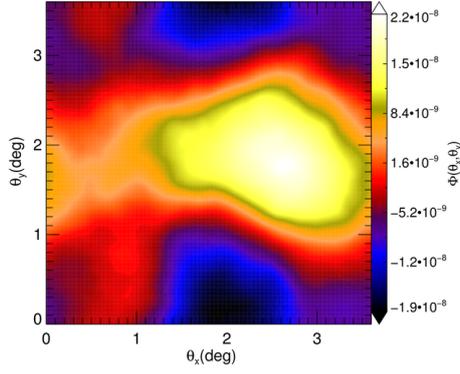


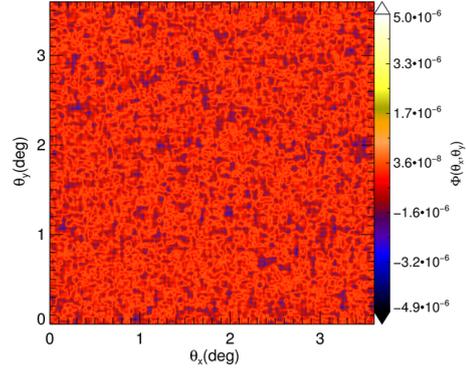
Figure 5.25: The discrete estimator noise for SKA1-Low (blue), SKA2-Low (red), SKA1-Low with $k_p^{\min} = 3$ (black), and SKA2-Low with $k_p^{\min} = 3$ (gold), with choices for observation time and bandwidth listed in Table 5.1 and for the non-uniform thermal noise power spectrum introduced in this section. The simulated sky area is $\Omega_s = 3.6^\circ \times 3.6^\circ$ and the beam has a resolution of 1.15 arcmins at $z = 8$. The R0 survey strategy results are not plotted because they produce an estimator reconstruction noise level close to R2 one. The R1 configuration is on dashed lines while the R2 is on dashed-dotted lines.

Figure 5.26 shows the input potential, estimator noise image, recovered image, and recovered square amplitude in Fourier space for SKA2-Low R2 model with $k_p^{\min} = 3$. We can notice how the high-scale noise completely overwhelms the signal, making it nearly undistinguishable from the noise image. On the other hand, the recovered power spectrum follows the theoretical profile, showing the estimator reconstruction noise feature at high multipoles which causes the high noise in the recovered image.

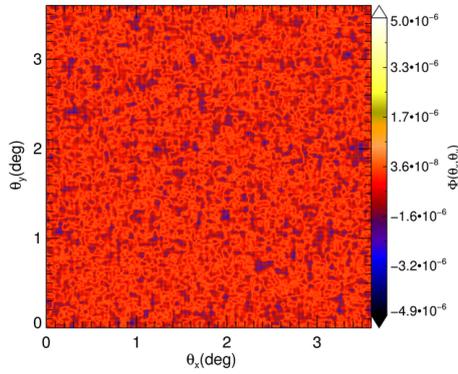
We will see in next section how it is possible to recover from these images the lensing potential with high fidelity.



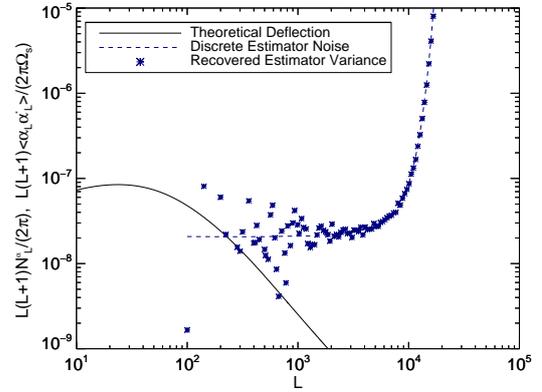
5.26.1 Input Gaussian random potential



5.26.2 Recovered estimator noise image



5.26.3 Recovered potential image



5.26.4 Recovered estimator Fourier amplitude

Figure 5.26: *Top panel*: the input potential field and the recovered estimator noise image. *Bottom panel*: the recovered estimator image and the recovered estimator Fourier amplitude. The images are computed for $N_{\text{side}} = 468$, $\Omega_s = 3.6^\circ \times 3.6^\circ$, $z = 8$, $k_p^{\text{min}} = 3$, and the SKA2-Low R2 non-uniform thermal noise model.

5.6.4 De-noising of the Reconstructed Image

Previous section's findings suggest that the image of the recovered potential can be visualised if a proper de-noising procedure is applied to the high-scale noise contaminated map. The idea is to use not all the modes available in the estimated map and/or exclude the modes involving high noise. This study considers three approaches often found in literature to de-noise contaminated images, and they are discussed in this section.

Wiener Filtering

The Wiener filter is an optimal filter, *i.e.* it makes the estimated image variance to be minimum. It is mainly used for deconvolutions or images degraded by additive noise and blurring caused by a Point Spread Function (PSF). This filter requires the second-order stationarity of both signal and noise processes.

In our case this is a satisfied condition, since both noise and signal are modeled as zero-mean processes. Moreover there must not be correlations between signal and noise, *i.e.* the noise has to be additive. This means that the input image is $S(i, j) = H(i, j)s(i, j) + N(i, j)$, where $H(i, j)$ is the blurring PSF, $s(i, j)$ is the uncontaminated image, and $N(i, j)$ is the additive noise. So if we have the Fourier transform of a pixelated image, namely $S(l, m)$, the estimated image is

$$\hat{S}(l, m) = W(l, m)S(l, m), \quad (5.22)$$

in which the Wiener filter is defined as

$$W(l, m) = \frac{H^*(l, m)}{|H|^2 + \frac{P_N(l, m)}{P_S(l, m)}}, \quad (5.23)$$

with $P_N(l, m)$ and $P_S(l, m)$ the power spectra of noise and signal respectively. So $P_N(l, m)/P_S(l, m)$ is the reciprocal of the signal-to-noise ratio. Often these two quantities are unknown in real application, so it is useful to parametrise our ignorance with a quantity, like the SNR, which can be experimentally estimated.

In our case we have no PSF, so $H(i, j) = 1$, and we know the forms of theoretical reconstruction noise and signal power spectra. So, Eq. (5.23) is simplified to

$$W_L = \frac{1}{1 + \frac{N_L^{\phi\phi}}{C_L^{\phi\phi}}}, \quad (5.24)$$

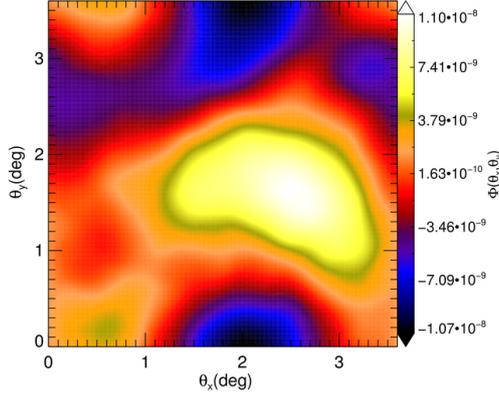
where we can substitute Eqs. (4.4) and (4.22) for signal and noise respectively. We can see that when $SNR = C_L^{\phi\phi}/N_L^{\phi\phi} \gg 1$, the filter is one, while for $SNR \ll 1$ we have $W_L \rightarrow 0$. If we apply this filter to our reconstructed potential image Figure 5.26.3, we obtain Figure 5.27.1, which looks like a high-fidelity smoothed version of the input image 5.26.1.

This can be better realised looking Figure 5.27.2, which illustrates the recovered Fourier square amplitude of the image. We can see that the noise signatures have been filtered out, but also the image has been overly-smoothed and, hence, underestimated. This is because Wiener filters are unable to reconstruct frequency components which have been degraded by noise, since they can only suppress them. In general Wiener filters are also unable to restore modes with bandlimited PSF, namely $H_L = 0$.

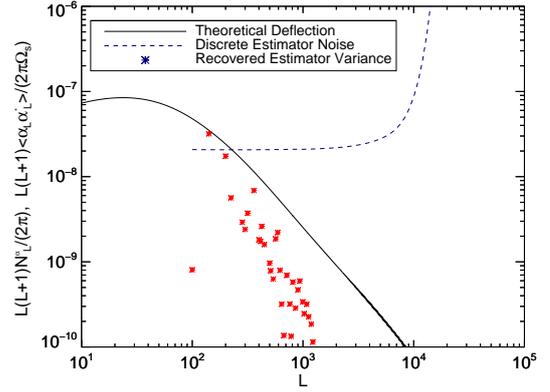
Gaussian Low-Pass Filter

The principle (and the form) of this filter is the same of the one introduced for the Beam function in Section 4.4. This filter can suppress all the frequencies bigger than a characteristic frequency L_{filt} , corresponding to a pixel resolution of σ_{filt} , preserving the small scale signal in which we are mainly interested. The difference with the previous approach consists in rejecting the idea of denoising the contaminated small-scale modes, cutting only the undesired noisy high-modes. The restored image is

$$\hat{S}(l, m) = G(l, m)S(l, m) = e^{-L(L+1)\sigma_{\text{filt}}^2/2}S(l, m), \quad (5.25)$$



5.27.1 Wiener de-noised image



5.27.2 Wiener filtered Fourier amplitude

Figure 5.27: *Left panel*: the noise reconstructed potential field obtained using the Wiener filter described in this section. *Right panel*: the filtered estimator Fourier amplitude. The recovered estimator is computed for $N_{\text{side}} = 468$, $\Omega_s = 3.6^\circ \times 3.6^\circ$, $z = 8$, $k_p^{\text{min}} = 3$, and the SKA2-Low R2 non-uniform thermal noise model.

with $\sigma_{\text{filt}} = g\Delta\theta / \sqrt{8 \ln 2}$. g is a number which quantifies the suppression respect to $\Delta\theta$, the basic resolution of the noise contaminated map.

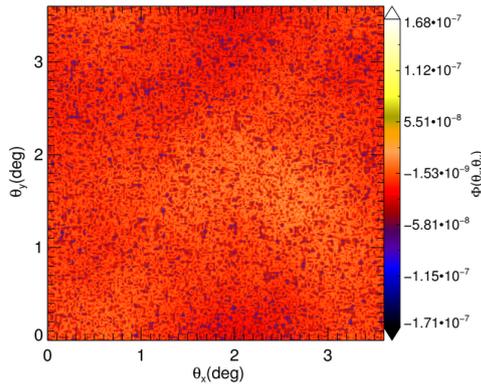
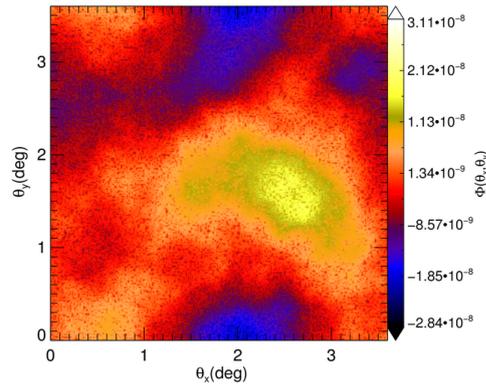
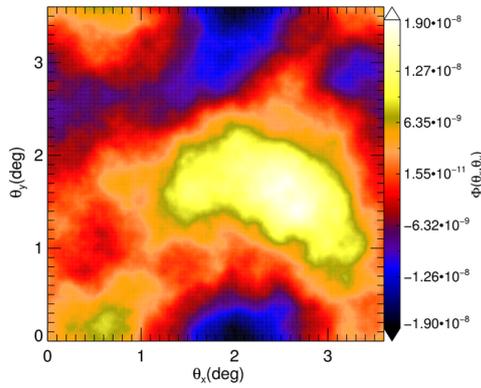
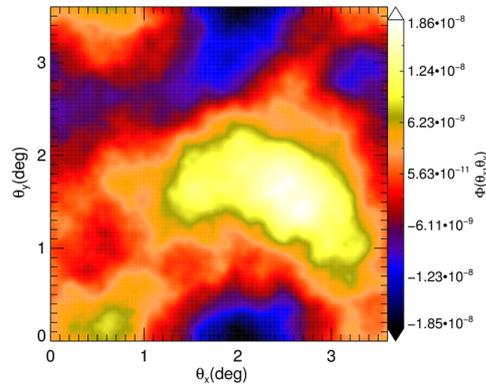
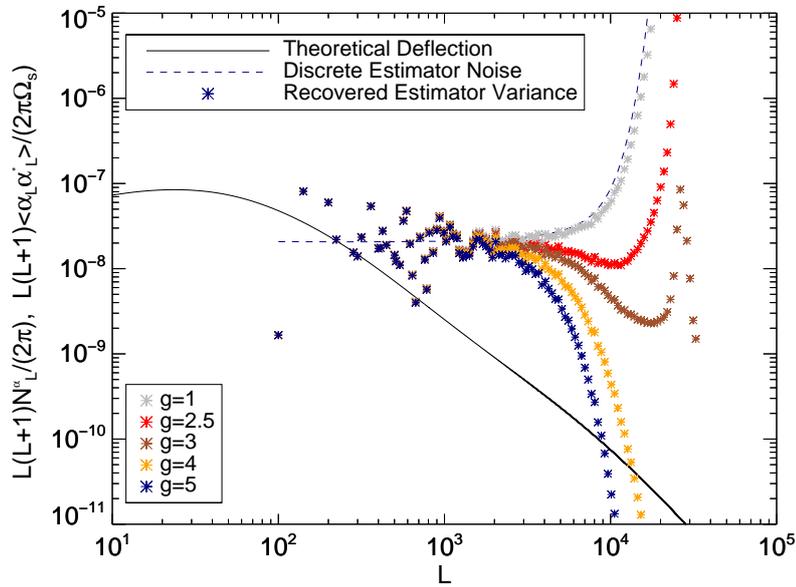
In figure 5.28 we can see the recovered images for several filter resolution σ_{filt} , or for different g 's. The basic pixel resolution is $\Delta\theta = \theta_{\text{side}}^{\text{max}} / N_{\text{side}} = 27.69$ arcsec. We can note how the high-scale signal is smoothed by looking at the denoised Fourier square amplitudes in Figure 5.28.5, in function of g .

The modes $L < L_{\text{filt}} = \sqrt{2}\pi / \sigma_{\text{filt}}$ are basically unchanged, preserving at this frequencies the total variance $\Omega_s (C_L^{\phi\phi} + N_L^{\phi})$. All the scales $L > L_{\text{filt}}$ are suppressed as the filter resolution is bigger, and so L_{filt} does not need to be too much low in order to not suppress too much low-scale modes. Figure 5.28.5 illustrates the cases corresponding to filtering multipoles of $L = 33092, 13237, 11031, 8273, 6618.5$, from lower to higher g 's. The $g = 2.5$ correspond to our well-known L_{cut} frequency, which perhaps represent the wisest choice.

Padding in Fourier Space

The last considered approach for de-noising images consists in padding the image in Fourier space with zeros for a square region in Fourier space delimited by the pixel M . When this cut is performed one need to pay serious attention to the storage of a Hermitian array, as explained in Appendix C.4.1. The square region in Fourier space has to be taken around the negative-positive ordered modes, and this means to have a zero-valued map for every index $m \leq m_{\text{Nyq}}$ when $M/2+1 \leq l \leq N-M/2$, and for $m \geq M/2+1$ when $0 \leq l \leq l \geq N-M/2+1$. It is not possible to cut directly the map in Fourier space inserting the values in a new matrix of smaller dimensions $(M-1, M/2)$.

As shown in Figure 5.29 for $M = 200$ and $M = 100$, respectively, this would considerably contaminate the denoised image with aliasing, although the high

5.28.1 Gaussian de-noised image with $g = 1.5$ 5.28.2 Gaussian de-noised image with $g = 2$ 5.28.3 Gaussian de-noised image with $g = 2.5$ 5.28.4 Gaussian de-noised image with $g = 5$ 

5.28.5 Gaussian filtered Fourier amplitude

Figure 5.28: *Top panels*: the noise reconstructed potential field obtained using the Gaussian filter described in this section with $g = 1.5, 2, 2.5, 5$, respectively. *Last bottom panel*: the filtered estimator Fourier amplitude for $g = 1, 2.5, 3, 4, 5$. The recovered estimator is computed for $N_{\text{side}} = 468$, $\Omega_s = 3.6^\circ \times 3.6^\circ$, $z = 8$, $k_p^{\text{min}} = 3$, and the SKA2-Low R2 non-uniform thermal noise model.

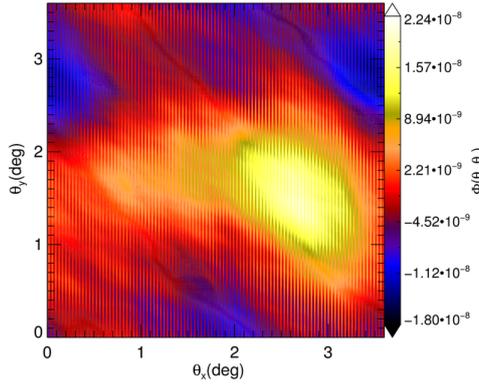
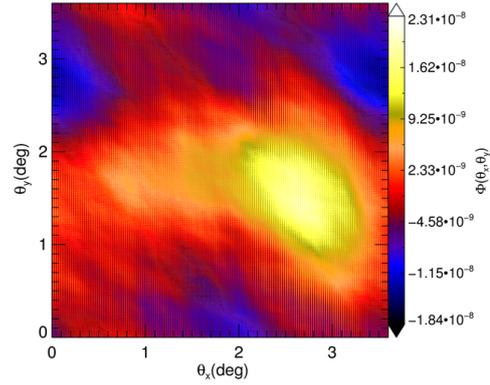
5.29.1 Aliased estimator image for $M = 200$.5.29.2 Aliased estimator image for $M = 100$.

Figure 5.29: *Left panel:* the aliased estimator imaged obtained by roughly cutting the Fourier space map around a rectangular region $(M - 1, M/2)$ with $M = 200$. The aliased estimator imaged obtained by roughly cutting the Fourier space map around a rectangular region $(M - 1, M/2)$ with $M = 100$. The recovered estimator is computed for $N_{\text{side}} = 468$, $\Omega_s = 3.6^\circ \times 3.6^\circ$, $z = 8$, $k_p^{\text{min}} = 3$, and the SKA2-Low R2 non-uniform thermal noise model.

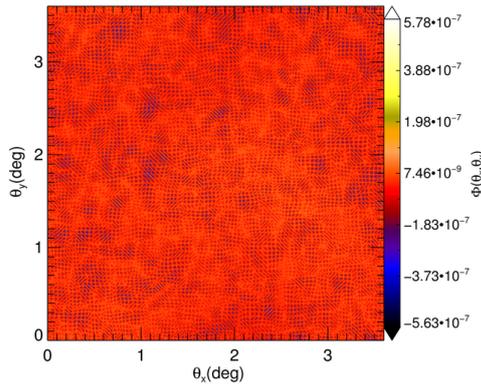
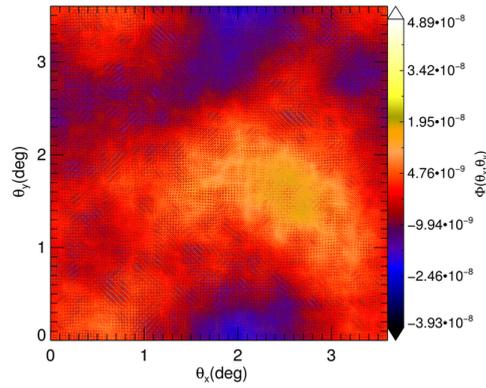
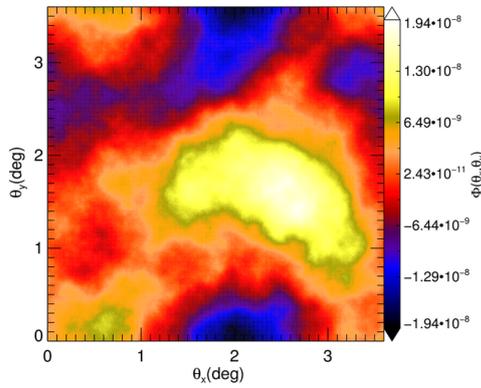
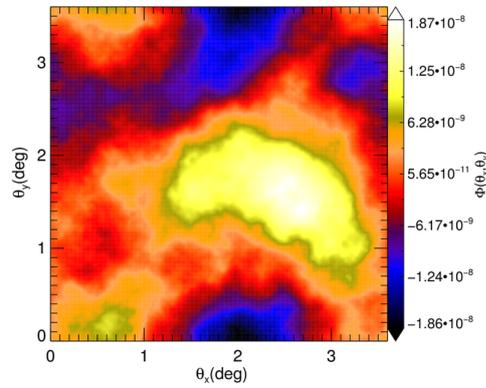
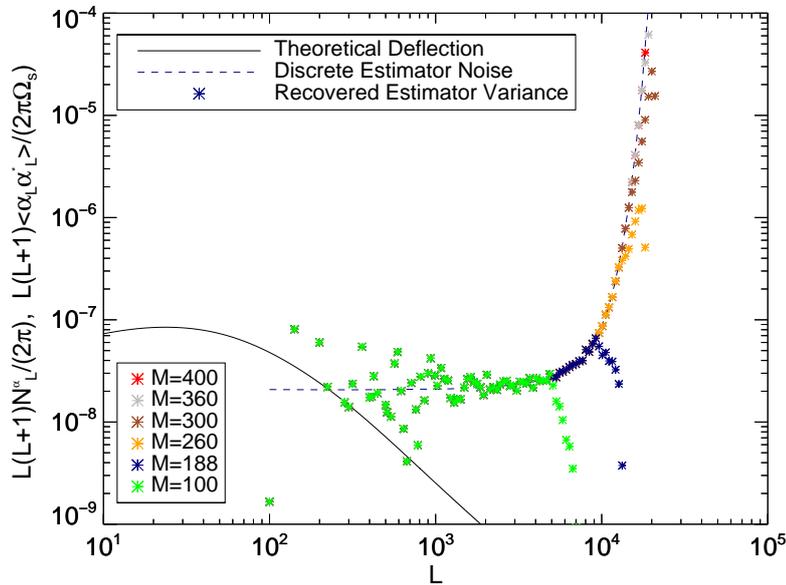
scale noise is excluded. The aliasing effect is explained in Appendix C.2 and its contamination is more serious as more modes are cut in Fourier space.

Thus, a given M cutting pixel number will correspond to various characteristic frequencies, which we will call $L_{\text{pad}} = \Delta LM / \sqrt{2}$. We display in Figure 5.30 the resulting de-noised estimator images for various M , namely $M = 400, 360, 260, 100$, corresponding to $L_{\text{pad}} = 28284, 25456, 18385, 7071$. The "2/3 truncating rule" ensures as that aliasing is avoided if $M/2 \leq 2N_{\text{Nyq}}/3$, which means $M \leq 2N_{\text{side}}/3$. If $N_{\text{side}} = 468$, $M \leq 312$ to avoid aliasing in the range $-155 \leq (l, m) \leq 155$.

In Figure 5.30.5 the padded estimator Fourier square amplitudes are computed for several choices of the padding cut pixel number M . It can be noted how the signal left at scales $L \gtrsim L_{\text{pad}}$ decreases. This happens in a sharper way respect to the Gaussian low-pass filter case. Infact, the values of the Fourier square amplitude at $L \gg L_{\text{pad}}$ assume an almost zero value ($\sim 10^{-40}$). So padding causes a sudden interruption of the signal, while the Gaussian low-pass filter has a smoother transition towards lower values.

Which of the three filtering approaches provide the highest fidelity image? To answer this question we plotted in Figure 5.31 the fidelities for the Wiener case (orange dashed line), the Gaussian case with $L_{\text{filt}} = 13230.9$ ($g = 2.5$, red dashed line), and padding case with $L_{\text{pad}} \sim L_{\text{filt}}$ which corresponds to $M = 188$ (blue dashed line).

Note how the Wiener filter has a worse fidelity in the range $100 \geq L \gtrsim 300$ than Gaussian or Padding, because of excessive smoothing, and tends to reach the limit of $S/N \rightarrow 1$ at $L > 3000$ or so. This happens because the estimated denoised image is too small with respect to the true one, leading to $F_\phi(L) = 1$. In general it seems to behave slightly better than Gaussian or padding cases at

5.30.1 Padded estimator image with $M = 400$ 5.30.2 Padded estimator image with $M = 360$ 5.30.3 Padded estimator image with $M = 300$ 5.30.4 Padded estimator image with $M = 100$ 

5.30.5 Padded estimator Fourier amplitude

Figure 5.30: *Top panels:* The potential estimator shown with high-scale modes contaminations removed with padding, for $M = 400, 360, 300, 100$, respectively. *Bottom last panel:* The padded estimator Fourier square amplitude, as computed for $M = 400, 360, 300, 260, 188, 100$. The recovered estimator is computed for $N_{\text{side}} = 468$, $\Omega_s = 3.6^\circ \times 3.6^\circ$, $z = 8$, $k_p^{\text{min}} = 3$, and the SKA2-Low R2 non-uniform thermal noise model.

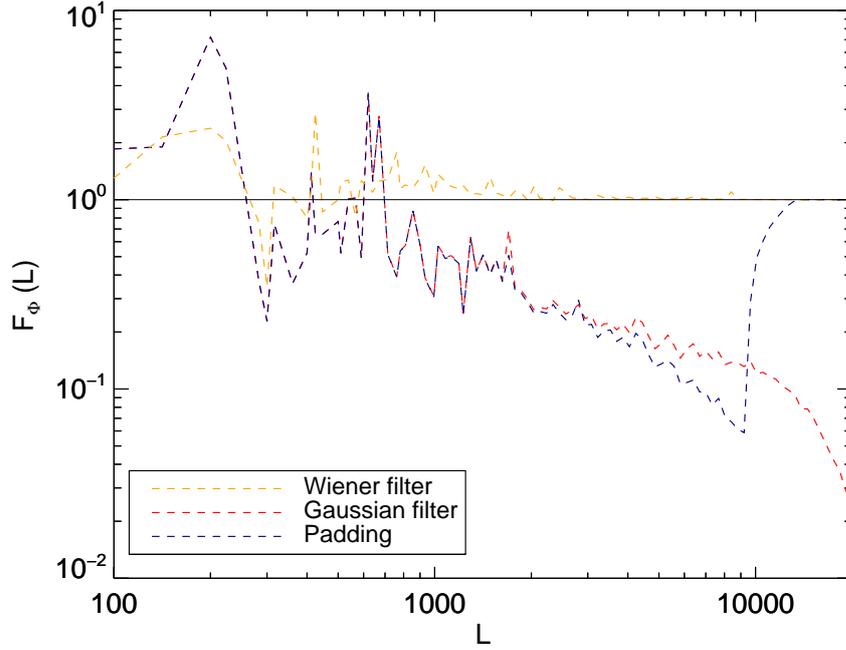


Figure 5.31: The fidelities for the filtered estimator maps obtained through the three approaches examined in this section. Wiener filter corresponds to the orange dashed line, Gaussian filter corresponds to the red dashed line, while padding corresponds to the blue dashed line. The straight line at $F_\phi(L) = 1$ helps to distinguish modes with good fidelity from the ones with bad fidelity.

intermediate scales, $300 \lesssim L \lesssim 1000$. Gaussian filter and Padding have indistinguishable fidelities up to $L \sim 3000$, and after this scale they diverge to different behaviours due to the different way of smoothing or excluding high-scale modes. It can be noted how the padding line tends to $S/N = 1$, since at those modes the padded estimator is null, while the Gaussian case goes smoothly to zero.

This plot can be compared with its analogous for a uniform thermal noise model, Figure 5.10. A little worsening in the intermediate scale is noticeable, but we obtain higher fidelities at large scales, due to the higher sensitivity available for these modes.

In summary, it seems that Wiener filtering provides the best high-fidelity image, but, as already stated, this is not equivalent to having an accurate power spectrum measurement. This would require a dedicated estimator which is beyond the scope of this work. We need to point out that for Gaussian filter and padding, the signal is not denoised: high-scale noisy modes which disturb the real space image are simply excluded, leaving the modes $L < L_{\text{filt,pad}}$ still noise contaminated. While Wiener filtering is unambiguously defined to be optimal, Gaussian filter and padding are kind of ill-defined, since the filtering scale can be arbitrarily chosen.

It should be interesting to test alternative filters similar to the Wiener. The results shown here encourage to see if other approaches, like modified Wiener filters or wavelets for example, could also detect with good accuracy the power spectrum and provide a high-fidelity of the potential simultaneously. We plan

to explore these approaches in future works.

5.6.5 Multi-Band Results for non-Uniform Noise Model

At this point we proceeded to compute our results stacking 14 redshift bands in the range $z = 7 - 11.6$, corresponding to a frequency range of $\nu_c = 177.55 - 112.55$ with 5 MHz bandwidths, following the steps already described in Section 5.5.5, and using the total-band estimator Eq. (5.10) and the total band noise Eq. (5.9). We will consider only Table 5.1 R1 and R2 survey strategies, since R0 8 MHz case provides results similar to R2 one. We have increased z_{\min} in order to be more conservative about the EoR ending period.

With respect to the previously adopted noise model, there is a particular aspect concerning the FoV. As seen in Section 5.6, the FoV is frequency dependent. A more complete description of the beam would also include a cutoff at large scales induced by the PSF of the telescope. This means that the estimator in Fourier space will have a different grid dimension at each band. In reality this would be true also within each frequency band, for each k_p mode. On the other hand the estimator noise level does not depend on the FoV (which sets the resolution in Fourier space), and so the reconstruction general properties are not greatly affected from the signal coming from larger scales, which can slightly change the results for detecting more or less large scale modes in the lensing potential spectrum or seeing more or less features in the recovered images, as demonstrated when we considered the Table 5.2 (b) and (c) cases in Figures 5.6, 5.7, and 5.9). For the moment we will consider that the FoV is fixed from one band to the other, and we will keep assuming that the properties of the lensing and telescope do not change within a single band.

So we will consider a $\Omega_s = 5^\circ \times 5^\circ$ survey area (set by the highest observed central redshift), which implies $\Delta L = 72$. The smallest observable redshift fixes the Nyquist mode to $L \sim 37267$, since $L_{\text{Nyq}} \gtrsim 2.5L_{\text{cut}}$, with $L_{\text{cut}} = 14884.7$, corresponding to $\Delta b = b\Delta\theta = 1.02$ arcmin. Considering that at $z_{\max} = 11.6$, $L_{\text{cut}} = 9430.89$, we will vary b from band to band, reaching the final beam resolution at z_{\max} of $\Delta b = 1.62$ arcmin. For each band we used $k_p^{\max} = 20$ modes and $N_{\text{side}} = 732$.

The computed multi-band estimator noise levels are presented in Figure 5.32, for SKA1 and SKA2-Low R1 and R2 telescope models. As made in Section 5.5.5, the effect of foreground contamination has been included considering the estimator noise computed with $k_p^{\min} = 3$. Comparing this plot with its uniform thermal noise analogous, it is noticed a half-magnitude improvement for the SKA1-Low results, even including the effect of foreground removal and without considering additive bands coming from the unexplored redshift range $6.5 \leq z < 7$, reaching an imaging regime with SNR analogous to the SKA2-Low in single-band we have seen for $z = 8$. Increasing the observational time from model R1 to model R2 causes the noise level to be decreased by nearly a factor 2. For what concerns SKA2-Low, the level of the total-band is basically unchanged respect to the uniform case, due to the noise saturation limit. In this case, the SNR is not greatly affected by increasing the observational time

from R1 to R2 strategy.

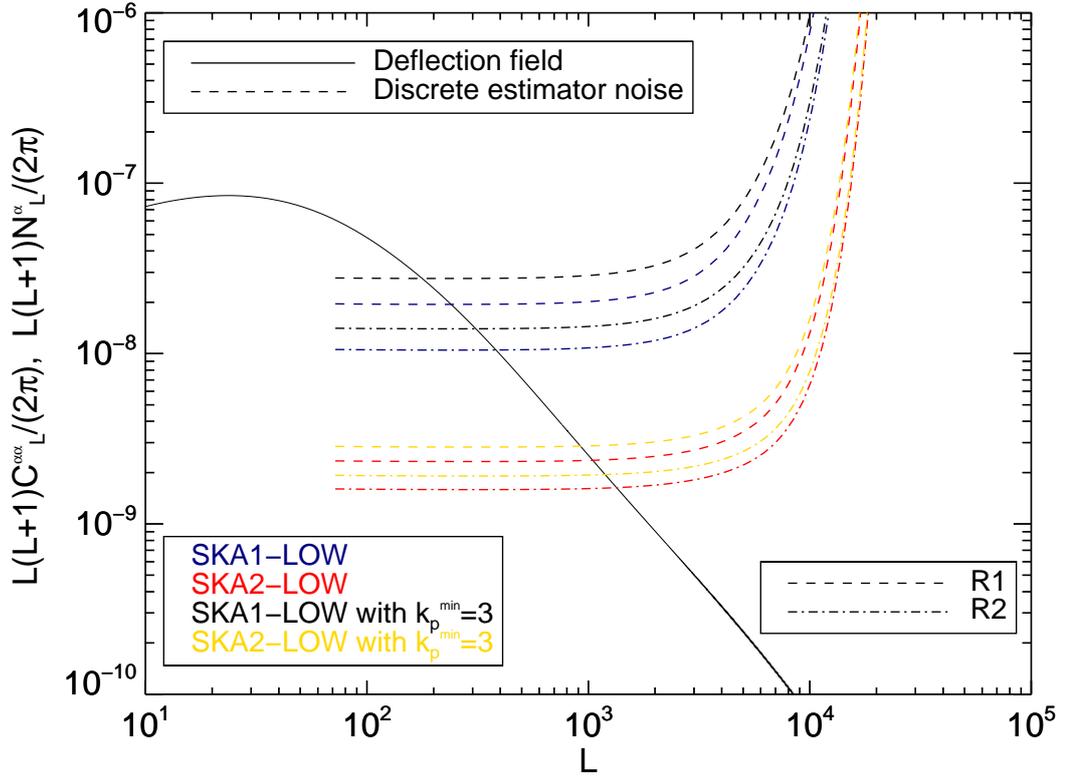
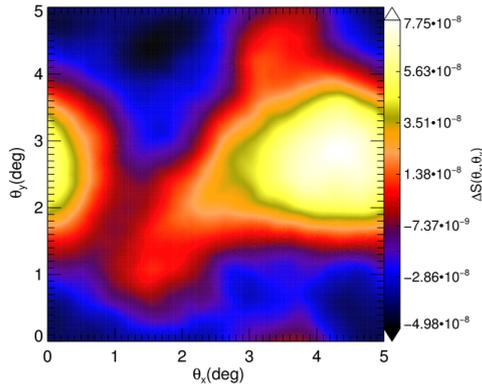


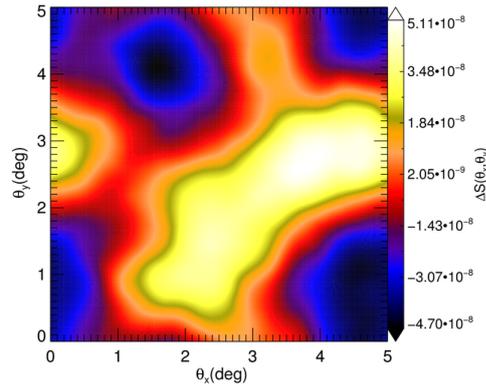
Figure 5.32: The multi-band discrete estimator noise for SKA1-Low (blue), SKA2-Low (red), SKA1-Low with $k_p^{\min} = 3$ (black), and SKA2-Low with $k_p^{\min} = 3$ (gold), with choices for observation time and bandwidth listed in Table 5.1 and for the non-uniform thermal noise power spectrum introduced in this chapter. The explored redshift range goes from $z = 7$ to $z = 11.6$. The simulated sky area is $\Omega_s = 5^\circ \times 5^\circ$ and $k_p^{\max} = 20$. The R0 survey strategy results are not plotted because they produce an estimator reconstruction noise level close to R2 one. The R1 configuration is on dashed lines while the R2 is on dashed-dot lines.

Following the procedure described in Section 5.5.5 we proceeded to compute the estimated potential images excluding the first 3 k_p modes from the total-band estimator for each survey strategy and frequency band. As for the single-band results, the resulting images are maps dominated by small-scale noise. Figure 5.33 hence shows the denoised maps computed following the Wiener filtering procedure described in Section 5.6.4, which can be compared to the input potential map in Figure 5.33.1. Here we have considered the SKA1-Low R2 and SKA2-Low R1 models for our discussion. The differences between the SKA1-Low images can be barely noticed, mainly due to Wiener filter smoothing. On the other hand SKA2-Low model reproduces the input structures with more accuracy than the two SKA1-Low models.

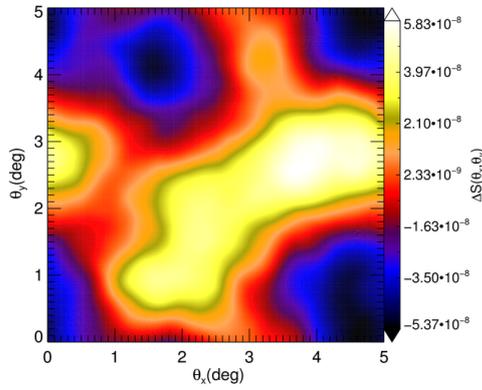
This behaviour can be further investigated by looking at the fidelities displayed in Figure 5.33.5 for the examined models. Thanks to the Wiener filtering procedure, we see that the fidelity is generally above 1, and obviously the SKA2-Low model map has a better quality, explaining the accuracy with which the



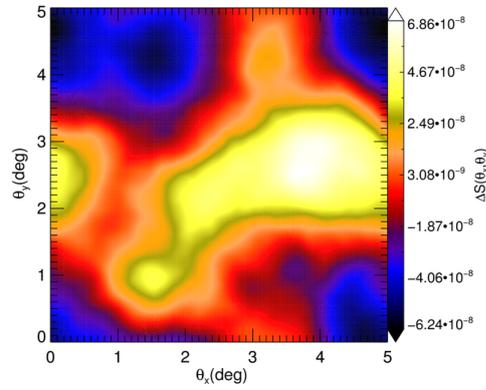
5.33.1 Input potential field



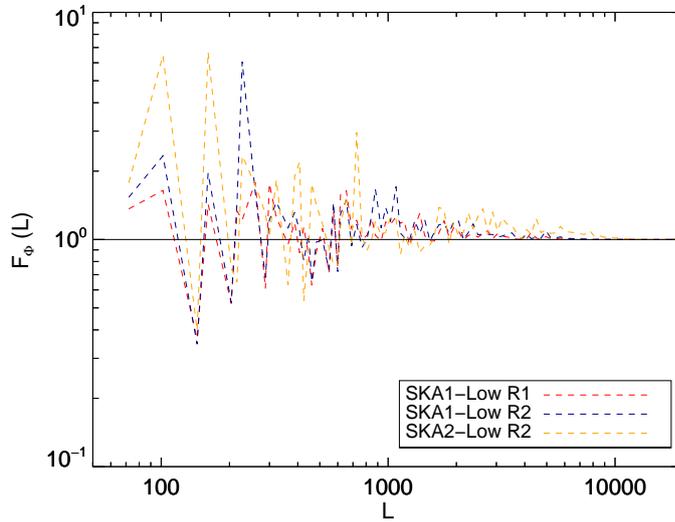
5.33.2 Denoised estimator SKA1-Low R1



5.33.3 Denoised estimator SKA1-Low R2



5.33.4 Denoised estimator SKA2-Low R1



5.33.5 Fidelities for considered models

Figure 5.33: From top to bottom: the input potential field, the denoised estimators for SKA1-Low R1, SKA1-Low R2, SKA2-Low R1 models, and respective fidelities. The recovered estimators are computed for $N_{\text{side}} = 732$, $\Omega_s = 5^\circ \times 5^\circ$, $k_p^{\text{min}} = 3$, non-uniform thermal noise model, and by combining 14 maps from the redshift range $z = 7 - 11.5$. Denoising is performed using a Wiener filter.

structures are reproduced in Figure 5.33.4. The two SKA1-Low reconstructed potentials seem to have similar fidelities, but the R2 model produces a slightly better image than R1 one, as could be foreseen from Figure 5.32.

As a last remark, note that a higher number of modes will be above the noise level if the FoV is increased. This can be accomplished by using more beams simultaneously for a SKA-Low telescope, as explained in Section 5.5.6, and mosaicking an area of the sky $S_{\text{area}} \geq N_b \Omega_s(z)$. Table 5.2 cases (b) and (c) indeed show a preliminary idea of the high quality images that can be recovered with such mosaicked FoV using additive large scale Fourier modes. Adding these modes does not improve the total signal-to-noise and the quality (fidelity) of the image will remain unchanged.

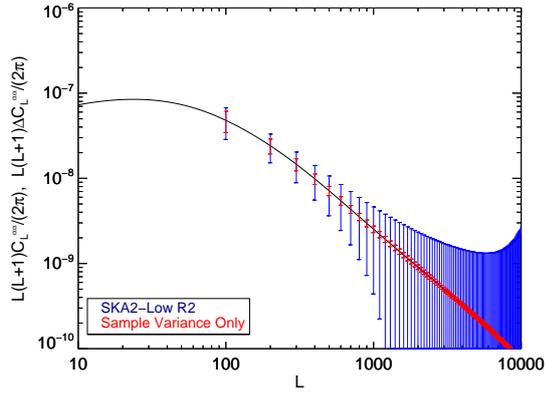
5.6.6 Power Spectrum Measurement Estimate

As demonstrated on Section 5.5.6, we can understand how accurately the power spectrum can be measured with a $\Omega_s = 5^\circ \times 5^\circ$ survey like the one considered for a multi-band detection simulated in Section 5.6.5 or with a $\Omega_s = 3.6^\circ \times 3.6^\circ$ survey for a single-band detection simulated in Section 5.6.3.

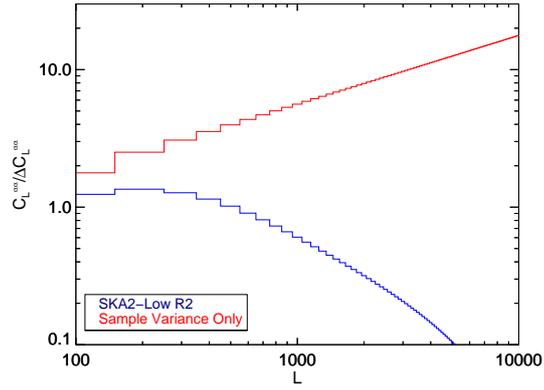
Let us first discuss the single-band results shown in Figure 5.34.1, in which the deflection power spectrum measurement errors are plotted for the SKA2-Low R2 model with a 5 MHz bandwidth centered around a redshift of $z = 8$. In this case $f_{\text{sky}} = 3.14 \times 10^{-4}$ and the multipole resolution is $\Delta L = 100$. We assumed that the foreground cleaning made unusable the first 3 k_p modes. The result is obtained by applying Eq. (4.25), and this is compared to the errors given by the sample variance limit for $\mathcal{N}_L^{\hat{\alpha}} \rightarrow 0$. The accuracy of the measurement can be further observed in Figure 5.34.2, where we show the fractional error ratio Eq. 5.12 compared to the sample variance fractional error Eq. (5.13). It can be noticed once again that a good fidelity image does not correspond to an accurate measurement of the power spectrum even in the region where the reconstruction noise is small compared to the deflection field signal.

Things are different for the multi-band results. Figure 5.34.3 shows the measurement error bars obtained for our most conservative survey strategy R1, for both SKA1-Low (red) and SKA2-Low (blue) telescope models. As described in Section 5.6.5, we used 14 bands in the redshift range $z = 7 - 11.6$, with $f_{\text{sky}} = 6 \times 10^{-4}$, $\Delta L = 72$, and $k_p^{\text{min}} = 3$. We can notice that even if in Figure 5.33.5 the images fidelities were more or less comparable, here SKA2-Low improves considerably the results with respect to SKA1-Low model. This can be better appreciated in Figure 5.34.4, showing the fractional error ratio of the above mentioned models, compared to the sample variance limit result (orange). We can see that the SKA2-Low result can measure the power spectrum with an accuracy comparable to the sample variance one for $L \lesssim 1000$. The phase 2 for SKA-Low improves considerably the accuracy also in the estimator reconstruction noise limited regime at high L 's.

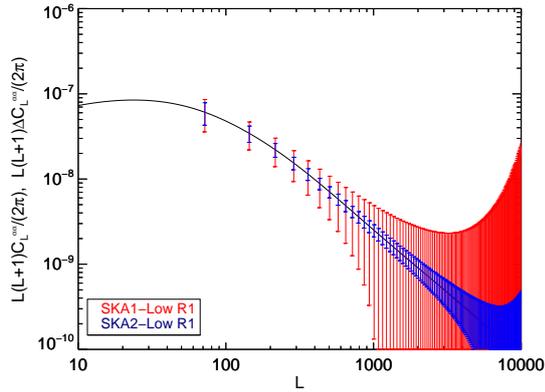
In general if these results are compared with the uniform noise model ones pictured in Figure 5.24, it can be noticed once again how the situation is slightly worse mainly because of the more steep reconstruction noise for $L > 1000$.



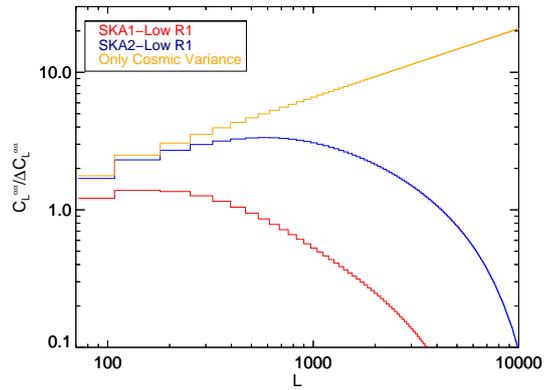
5.34.1 Single Band Errors



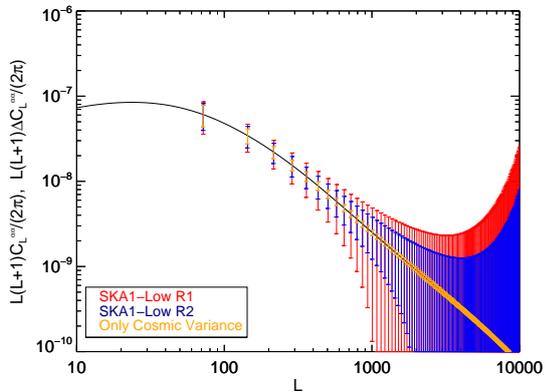
5.34.2 Single Band Fractional Errors



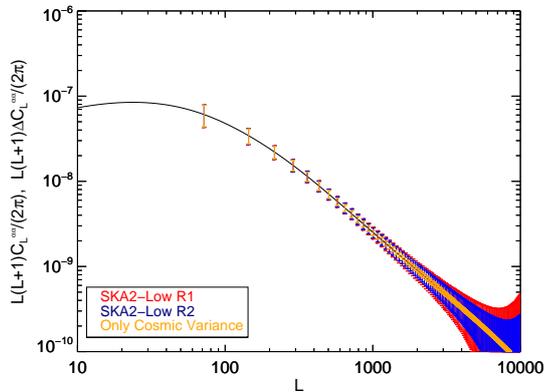
5.34.3 Multi-band Errors SKA-Low R1



5.34.4 Multi-band Fractional Errors SKA R1



5.34.5 Multi-band Errors SKA1-Low R1/R2



5.34.6 Multi-band Errors SKA2-Low R1/R2

Figure 5.34: *First row panels:* on the left the deflection field power spectrum and measurement error bars for a SKA2-Low R2 model, considering a single band detection at $z = 8$, with $f_{\text{sky}} = 3.14 \times 10^{-4}$, $\Delta L = 100$, and $k_p^{\text{min}} = 3$. On the right we see the fractional error for the same experiment (blue), compared to the sample variance limit (red). *Second row panels:* on the left the deflection field power spectrum and measurement error bars for a SKA1-Low (red) and a SKA2-Low R2 (blue) models, considering a multiple frequency band detection in the range $z = 7 - 11.6$, with $f_{\text{sky}} = 6 \times 10^{-4}$, $\Delta L = 72$, and $k_p^{\text{min}} = 3$. On the right we see the fractional error for the same experiments, compared to the sample variance limit (orange). *Third row panels:* on the left the deflection field power spectrum and measurement error bars for a SKA1-Low R1 (red) R2 (blue) models. On the right the same thing for SKA2-Low R1 (red) R2 (blue) models. The cosmic variance limit result is represented by the orange bars.

We also investigated possible benefits in changing the survey strategy by doubling the observational time from model R1 to R2. It can be seen from Figure 5.34.5 for SKA1-Low model and from Figure 5.34.6 for SKA2-Low model that the improvement is minimal. SKA1-Low still is far away from the sample variance limit, while SKA2-Low gets a bit closer to it, but the total accuracy is slightly improved.

The situation can be improved by considering larger surveyed areas of the sky (like the ones explored by SKA-Mid) in order to have a larger f_{sky} , or by detecting the signal for different patches in the sky. This latter kind of observation can increase the statistics of a given mode range by lowering the sample variance error, especially in the intermediate L -range $200 \lesssim L \lesssim 1000$. This can be done in a reasonable amount of time, even with SKA-Low aperture array.

5.7 Resuming of Results and Future Developments (pt. I)

In this section we have achieved a large number of results, which can be resumed in the following points:

- we have implemented a simulation pipeline and a theoretical framework capable of dealing with issues that can not be treated analytically, like the simulation of a telescope beam, the non-uniform visibility space coverage, the non-linearity of the lensing source field, non-Gaussianity of the 21 cm lensing field, foreground subtraction techniques, and the discreteness of visibility measurements;
- by taking advantage of the 21 cm source signal division into multiple statistically independent maps along the frequency direction, we have demonstrated how the lensing mass distribution can be reconstructed with high fidelity using a three dimensional optimal quadratic lensing estimator in Fourier space;
- we have demonstrated that the weak lensing assumption widely used in the CMB case, is also valid for the 21 cm field as well at the scales considered in this study;
- we have successfully implemented and tested the 3D Fourier space quadratic estimator in our simulation code, taking into account the beam of the telescope (set by the baseline maximum dimensions) and the discreteness of visibility measurements;
- we have showed that the discrete 21 cm weak lensing quadratic estimator can be employed by using a single frequency band or combining multiple frequency band measurements as well;
- we have implemented a more realistic telescope model with respect to the one widely used in literature and studied the estimator reconstruc-

tion technique for various survey strategies, including the possible degradation of the first k_p modes caused by foreground cleaning techniques;

- we have implemented a simple de-noising procedure in order to filter out the small-scale noise which is likely to strongly contaminate the estimated signal;
- we have found that an SKA-Low interferometer, modeled with our realistic noise telescope model, should obtain high-fidelity images of the underlying mass distribution in its phase 1 only if several bands are stacked up together, covering a redshift range that goes from $z = 7$ to $z = 11.6$ and with a total resolution of 1.6 arcmin. Moreover, the SKA-Low phase 2, modeled in order to improve the sensitivity of the instrument by almost an order of magnitude, should be capable of providing reconstructed images with good quality even when the signal is detected within a single frequency band. In this case the reconstructed image has a resolution of 1.15 arcmin at $z = 8$, within a field of view of 13 deg^2 ;
- in the case of multi-band detection of the lensed 21 cm signal made with an SKA2-Low telescope model we found constraints close to the sample variance ones in the range $L < 1000$, even for a small field of view such as a 25 deg^2 survey area. Good constraints have been found also for SKA1-Low in multi-band detection, and for SKA2-Low in single band detection;
- We have also explored the possibility to detect even a cluster lensing signal coming from a redshift $z = 0.5$ with a mass of $M = 10^{15} M_\odot$, but we found their signal to be overwhelmed by the estimator reconstruction noise by several orders of magnitude, going well below the saturation limit of the noise imposed by sample variance also for multi-band analysis.

Nevertheless, other issues need to be further investigated in order to improve these important results. These improvements towards a more realistic description of the 21 cm lensing detection can be easily implemented in our numerical framework. This is designed to handle the following points, which we plan to address in future works:

- our work would benefit from improvements on the modeling of a more realistic reionization history. For example, it would be very beneficial to run our simulations including a non-instantaneous, non Gaussian Epoch of Reionization. It is very likely that EoR was a non-homogeneous process expanded over a considerable redshift range, and the detected signal strongly depends on the number density of ionized regions which are causing inhomogeneities in the 21 cm temperature signal;
- another interesting extension of our work would be to implement foreground contamination and study how foreground removal techniques can affect the accuracy of imaging and detecting the lensing signal. In fact, foreground cleaning schemes, depending on the degree of foreground

contamination, could require the rejection of a large number of k_p modes. In particular, they could leave a residual signal and cross-correlations among different frequency bands, whose importance needs to be investigated numerically;

- we plan to take into account more realistic lensing distribution, like the ones provided by simulation codes like MOKA⁷. However, we do not expect to see much different results with respect to the ones obtained in this work, since these features appear only at small scales and we are reconstructing mainly a large scale signal.
- we also need to test how the frequency dependence of the FoV could influence our results, including a PSF term into a more accurate expression of the beam. However, we do not expect this issue to influence too much the results proposed here, because of the negligible effect of the largest scales on the reconstruction noise level and on the total signal-to-noise. On the other hand, a real telescope PSF needs to be handled numerically, since it can be a very complicated function for real radio telescopes;
- we plan to forecast constraints on the cosmological parameters, given the results obtained here for the power spectrum measurement statistics. We also plan to simulate the effect in power spectrum measurements of measuring the signal from different patches of the sky.

⁷<https://cgiocoli.wordpress.com/research-interests/moka/>

Chapter 6

Reconstruction at Low Redshifts

The HI intensity mapping introduced briefly in Section 1.4 allows for measuring the distribution of neutral hydrogen before and during EoR. Actually many efforts are produced to use such a technique to study whether positive and accurate detection of BAOs at low redshifts can be performed with SKA (Chang et al., 2010, 2008; Bull et al., 2015; Masui et al., 2010; Pober et al., 2013), but it represents a unique opportunity to detect 21 cm gravitational lensing as well.

We have studied in the previous chapters how to image the recovered mass distribution from temperature fluctuation treated as an unresolved background. The strength of intensity mapping consists in the possibility of detecting this three-dimensional field without resolving (in angular resolution, not frequency) individual objects such as high-redshift galaxies at an excellent resolution. Pourtsidou & Metcalf (2014) showed that the signal due to these sources, which can be modeled as a Poisson source term, do not represent an obstacle to lensing measurement, since they contribute to an improvement of the lensing signal itself. Lensing can be measured without resolving these objects and the Poisson shot noise, previously treated as a source of only noise (Zhang & Pen, 2006, 2005; Yang & Zhang, 2011), contributes to increase the signal as well. This technique can be extended to any redshift below EoR, with various degrees of signal-to-noise and combining redshift to perform tomographic studies, like the evolution of the growth function (which is at a much higher accuracy compared to the one obtained from galaxy surveys).

Pourtsidou & Metcalf (2015) developed the formalism to include point sources in our already explored theoretical framework for lensing reconstruction of a continuous source on a discrete grid, especially at redshifts $z \sim 2 - 3$. They found that an SKA-Mid interferometer could detect a possible evolution of the HI mass function or interacting dark energy models. Imaging capabilities for a post-EoR detection have been explored also in Hilbert et al. (2007), where the authors found that high-quality images can be detected using a SKA-like telescope model, far exceeding the signal-to-noises of any map made using galaxy survey.

Moreover, better constraints on lensing power spectrum measurement can be reached because of the higher sky area surveyed, as explained on previous chapters. Further improvements in measuring the lensing power spectrum

can be achieved by considering the HI or galaxy density fields in cross correlation with the detected convergence field. Poursidou et al. (2015) found that this considerably improves the 21cm lensing detection prospects and excellent results can be achieved within frequencies observed by SKA-Mid and MeerKAT. Cross-correlating the galaxy and HI densities will also give us information about the galaxy-HI correlation coefficient. A significant advantage of cross correlating HI intensity mapping and optical galaxy surveys is the alleviation of the issues arising from systematic effects.

In this chapter we will review this theoretical framework, in order to apply it to our already presented code and produce simulations for recovering lensing at these redshifts. We will describe how we generated unclustered discrete point sources, in order to move towards a more realistic simulation of clustered discrete sources of our HI signal. The aim is to investigate and simulate to what extent a SKA-Mid settings is capable of recovering the mass distribution and measure the lensing power spectrum.

6.1 Weak Lensing Estimator for Unclustered Sources

In Section 4.3 we derived the estimator from the expansion Eq. (4.1). In the case of randomly placed sources this might not be valid, since the potential variation may be larger than the size of the individual sources. We hence need to find a discrete estimator from a different assumption and we will follow Poursidou & Metcalf (2015) derivation.

6.1.1 Pure Poisson Noise Power Spectrum

Let us consider a three-dimensional grid of dimensions $N_{\text{vol}} = N_{\perp}^2 N_{\parallel}$. The volume of one cell i is $\delta V = V_{\text{tot}}/N_{\text{vol}}$. The total number of objects with given luminosity \mathcal{L} populating the volume V_{tot} is $N_{\text{gal}} = \bar{\eta} V_{\text{tot}}$, with $\bar{\eta}$ the average number density of objects.

Usually this number density is drawn from a given luminosity (or mass) distribution function which express the comoving number of objects per luminosity (mass) bin. Our objects are the HI sources and we will assume that they follow the Schechter luminosity (mass) function, whose properties are explored in Appendix F. This represents an excellent fit to the observed data in the local Universe, and it is

$$\frac{dn(\mathcal{L}, z)}{d\mathcal{L}} d\mathcal{L} = \Phi^*(z) \left(\frac{\mathcal{L}}{\mathcal{L}^*(z)} \right)^{\alpha} e^{-\mathcal{L}/\mathcal{L}^*(z)} d\left(\frac{\mathcal{L}}{\mathcal{L}^*(z)} \right). \quad (6.1)$$

This function is parametrized by a low-mass slope α , a characteristic luminosity $\mathcal{L}^*(z)$, and a normalization Φ^* . The choice of these parameters represents the biggest uncertainty in this study, since they are well measured only in the local Universe. As will be better explained in Section 6.5.1, a conservative no-evolution model¹ for the HI is adopted, using $\alpha = -1.3$, $\mathcal{L}^* = cM^* =$

¹Poursidou & Metcalf (2015) studied the properties of lensing reconstruction assuming various HI

$c3.47 h^{-2} 10^9 M_\odot$ (we assume that $\mathcal{L}/M = \text{const}$), and $\Phi^\star = 0.0204 h^3 \text{Mpc}^{-3}$. These values have been reported from the HIPASS survey (Zwaan et al., 2003).

If we define the average occupation number in a cell to be $\bar{n}_i = \bar{\eta} \delta V = N_{\text{gal}}/N_{\text{vol}}$, we can define the surface brightness fluctuation in one cell as the integrated luminosity contribution

$$\delta S(\mathbf{i}) = \sum_{\mathcal{L}} (n_i^{\mathcal{L}} - \bar{n}_i) \mathcal{L} = \sum_{\mathcal{L}} \delta n_i^{\mathcal{L}} \mathcal{L}, \quad (6.2)$$

where we defined $\delta n_i^{\mathcal{L}}$ to be the fluctuation in the number of sources with luminosity \mathcal{L} in a cell \mathbf{i} in real space. Hence, the average flux in a cell is

$$\bar{S} = \bar{n}_i \langle \mathcal{L} \rangle = \bar{\eta} \delta V \langle \mathcal{L} \rangle. \quad (6.3)$$

Note that $\langle \delta n_i^{\mathcal{L}} \rangle = 0$.

If (i_\perp, i_\parallel) are the components of the three-dimensional cell \mathbf{i} , we can define the DFT of the above expression, namely

$$\begin{aligned} \delta \tilde{S}(\mathbf{j}) &= \frac{\Omega_s}{N_{\text{vol}}} \sum_{\mathbf{i}_\perp} \sum_{i_\parallel} \delta S(\mathbf{i}) e^{-i2\pi \mathbf{i}_\perp \cdot \mathbf{j}_\perp / N_\perp^2} e^{-i2\pi i_\parallel j_\parallel / N_\parallel} \\ &= \frac{\Omega_s}{N_{\text{vol}}} \sum_{\mathbf{i}_\perp} \sum_{i_\parallel} \sum_{\mathcal{L}} \mathcal{L} \delta n_i^{\mathcal{L}} e^{-i2\pi \mathbf{i}_\perp \cdot \mathbf{j}_\perp / N_\perp^2} e^{-i2\pi i_\parallel j_\parallel / N_\parallel}, \end{aligned} \quad (6.4)$$

where $\mathbf{j} = (j_\perp, j_\parallel)$ represents a three-dimensional cell in Fourier space.

The correlation between different modes is

$$\begin{aligned} \langle \delta \tilde{S}(\mathbf{j}) \delta \tilde{S}^\star(\mathbf{j} - \mathbf{m}) \rangle &= \frac{\Omega_s^2}{N_{\text{vol}}^2} \left\langle \sum_{\mathbf{i}_\perp} \sum_{i_\parallel} \sum_{\mathcal{L}} \mathcal{L} \delta n_i^{\mathcal{L}} e^{-i2\pi \mathbf{i}_\perp \cdot \mathbf{j}_\perp / N_\perp^2} e^{-i2\pi i_\parallel j_\parallel / N_\parallel} \right. \\ &\quad \times \left. \sum_{\mathbf{i}'_\perp} \sum_{i'_\parallel} \sum_{\mathcal{L}'} \mathcal{L}' \delta n_{i'}^{\mathcal{L}'} e^{i2\pi \mathbf{i}'_\perp \cdot (\mathbf{j}_\perp - \mathbf{m}_\perp) / N_\perp^2} e^{i2\pi i'_\parallel (j_\parallel - m_\parallel) / N_\parallel} \right\rangle. \end{aligned} \quad (6.5)$$

Since $\langle \delta n_i^{\mathcal{L}} \rangle = 0$ and only the $\mathbf{i} = \mathbf{i}'$ and $\mathcal{L} = \mathcal{L}'$ terms contribute, we can write down

$$\begin{aligned} \langle \delta \tilde{S}(\mathbf{j}) \delta \tilde{S}^\star(\mathbf{j} - \mathbf{m}) \rangle &= \frac{\Omega_s^2}{N_{\text{vol}}^2} \left\langle \sum_{\mathbf{i}_\perp} \sum_{i_\parallel} \sum_{\mathcal{L}} \mathcal{L}^2 (\delta n_i^{\mathcal{L}})^2 e^{-i2\pi \mathbf{i}_\perp \cdot \mathbf{m}_\perp / N_\perp^2} e^{-i2\pi i_\parallel m_\parallel / N_\parallel} \right\rangle \\ &= \frac{\Omega_s^2}{N_{\text{vol}}^2} \left\langle \sum_{\mathbf{i}_\perp} \sum_{i_\parallel} \sum_{\mathcal{L}} \mathcal{L}^2 \bar{n}_i e^{-i2\pi \mathbf{i}_\perp \cdot \mathbf{m}_\perp / N_\perp^2} e^{-i2\pi i_\parallel m_\parallel / N_\parallel} \right\rangle, \end{aligned} \quad (6.6)$$

where the second moment of the number counts in a cell is given by a Poisson distribution $\langle (\delta n_i^{\mathcal{L}})^2 \rangle = \bar{n}_i$, as demonstrated in Appendix G.1.

At this point we include lensing effect on our brightness fluctuation field along the direction which is perpendicular to the line of sight. We know that

evolution models, finding that an SKA-like instrument should be able to provide high-fidelity images even assuming no evolution. Moreover, other models derived from damped Ly α systems are possible. See (P eroux et al., 2003) for one example.

lensing conserves the surface brightness, so if a cell of fixed angular size is magnified by a factor μ_i , the galaxies within it will be brighter by a factor μ_i . In order to keep the surface brightness, the true volume of the cell and the average number of galaxies have to be a factor $1/\mu_i$ smaller. This means:

$$\begin{aligned}
\langle \delta\tilde{S}(\mathbf{j})\delta\tilde{S}^*(\mathbf{j}-\mathbf{m}) \rangle &= \frac{\Omega_s^2}{N_{\text{vol}}^2} \langle \sum_{i_{\perp}} \sum_{i_{\parallel}} \sum_{\mathcal{L}} \mu_{i_{\perp}} \mathcal{L}^2 \bar{n}_i e^{-i2\pi i_{\perp} \cdot \mathbf{m}_{\perp}/N_{\perp}^2} e^{-i2\pi i_{\parallel} m_{\parallel}/N_{\parallel}} \rangle \\
&= \frac{\Omega_s^2}{N_{\text{vol}}^2} \bar{\eta} \delta V \langle \mathcal{L}^2 \rangle \langle \sum_{i_{\perp}} \mu_{i_{\perp}} e^{-i2\pi i_{\perp} \cdot \mathbf{m}_{\perp}/N_{\perp}^2} \rangle \langle \sum_{i_{\parallel}} e^{-i2\pi i_{\parallel} m_{\parallel}/N_{\parallel}} \rangle \\
&= \frac{\Omega_s^2}{N_{\text{vol}}} \bar{\eta} \delta V \langle \mathcal{L}^2 \rangle \tilde{\mu}(\mathbf{m}_{\perp}) \delta_{m_{\parallel}}^{\text{K}}, \tag{6.7}
\end{aligned}$$

where $\tilde{\mu}(\mathbf{m}_{\perp})$ is the DFT of the magnification and the Kroenecher delta indicates the absence of correlations along the line-of-sight direction.

Let us now consider the dimensionless brightness fluctuation field $\Delta S = \delta S/\bar{S} = S/\bar{S} - 1$. Its modes correlation in Fourier space is

$$\begin{aligned}
\langle \Delta\tilde{S}(\mathbf{j})\Delta\tilde{S}^*(\mathbf{j}-\mathbf{m}) \rangle &= \frac{\langle \delta\tilde{S}(\mathbf{j})\delta\tilde{S}^*(\mathbf{j}-\mathbf{m}) \rangle}{\bar{S}^2} = \frac{\Omega_s^2}{N_{\text{vol}}} \bar{\eta} \delta V \langle \mathcal{L}^2 \rangle \tilde{\mu}(\mathbf{m}_{\perp}) \delta_{m_{\parallel}}^{\text{K}} \frac{1}{\bar{\eta}^2 (\delta V)^2 \langle \mathcal{L} \rangle^2} \\
&= \frac{\Omega_s^2}{N_{\text{vol}}} \frac{1}{\bar{\eta} \delta V \langle \mathcal{L} \rangle^2} \tilde{\mu}(\mathbf{m}_{\perp}) \delta_{m_{\parallel}}^{\text{K}} = \frac{\Omega_s^2 \langle \mathcal{L}^2 \rangle}{N_{\text{gal}} \langle \mathcal{L} \rangle^2} \tilde{\mu}(\mathbf{m}_{\perp}) \delta_{m_{\parallel}}^{\text{K}}. \tag{6.8}
\end{aligned}$$

Now we can define the angular shot noise power spectrum

$$C^{\text{shot}} = \frac{\Omega_s \langle \mathcal{L}^2 \rangle}{N_{\text{gal}} \langle \mathcal{L} \rangle^2}, \tag{6.9}$$

which depends on the ratio of the moments of the luminosity function defined in Appendix F.2. In the end we get

$$\langle \Delta\tilde{S}(\mathbf{j})\Delta\tilde{S}^*(\mathbf{j}-\mathbf{m}) \rangle = \Omega_s C^{\text{shot}} \tilde{\mu}(\mathbf{m}_{\perp}) \delta_{m_{\parallel}}^{\text{K}}. \tag{6.10}$$

6.1.2 Quadratic Estimator for Magnification

We define the quadratic estimator for the magnification as

$$\hat{\mu}(\mathbf{m}_{\perp}) = \frac{1}{\Omega_s C^{\text{shot}}} \frac{1}{N_{\text{vol}}} \sum_{j_{\perp}} \sum_{j_{\parallel}} \Delta\tilde{S}(\mathbf{j})\Delta\tilde{S}^*(\mathbf{j}-\mathbf{m}_{\perp}), \tag{6.11}$$

This optimal estimator is trivially unbiased

$$\langle \hat{\mu}(\mathbf{m}_{\perp}) \rangle = \frac{\tilde{\mu}(\mathbf{m}_{\perp})}{N_{\text{vol}}} \sum_{j_{\perp}} \sum_{j_{\parallel}} 1 = \tilde{\mu}(\mathbf{m}_{\perp}) \tag{6.12}$$

and the estimator filter is a function of only m_{\perp} .

We can write the variance of this estimator

$$\begin{aligned} \langle |\hat{\mu}(\mathbf{m}_\perp)|^2 \rangle &= \frac{1}{(\Omega_s C^{\text{shot}})^2} \frac{1}{N_{\text{vol}}^2} \sum_{j_\perp} \sum_{j_\parallel} \sum_{j'_\perp} \sum_{j'_\parallel} \langle \Delta \tilde{\mathcal{S}}(j) \Delta \tilde{\mathcal{S}}^*(j - \mathbf{m}_\perp) \Delta \tilde{\mathcal{S}}^*(j') \Delta \tilde{\mathcal{S}}(j' - \mathbf{m}_\perp) \rangle \\ &= \frac{1}{N_{\text{gal}}} \frac{\langle \mathcal{L}^4 \rangle}{\langle \mathcal{L}^2 \rangle^2} \left(1 + 3 \frac{N_{\text{gal}}}{N_{\text{vol}}} \right) + 2 \frac{N_{\text{vol}} - 1}{N_{\text{vol}}^2}. \end{aligned} \quad (6.13)$$

This computation involves the higher moments of the Poisson number counts in cells, and more details can be found in Appendix G.1.1. This expression is finite and different from zero when $N_{\text{vol}} \rightarrow \infty$, since only the first term survives.

As already pointed out by Pourtsidou & Metcalf (2015), this estimator is linked to the potential estimator because in the weak lensing limit we have $\mu \simeq 1 + 2\kappa = 1 - \nabla^2 \phi$. In Fourier space this implies

$$\hat{\phi}(\mathbf{L}) = \frac{\hat{\mu}(\mathbf{L})}{L^2}, \quad (6.14)$$

with $\mathbf{L} = \Delta L \mathbf{m}_\perp$. It can be realised that this filter is the same that one would get if a constant power spectrum for the Gaussian case filter is assumed, but obtaining a different noise. The estimator reconstruction noise in this case is found in the limit $N_{\text{vol}} \rightarrow \infty$, namely

$$N_L^{\hat{\mu}} = \frac{\lim_{N_{\text{vol}} \rightarrow \infty} \langle |\hat{\mu}(\mathbf{m}_\perp)|^2 \rangle}{\Omega_s} = \frac{1}{\Omega_s N_{\text{gal}}} \frac{\langle \mathcal{L}^4 \rangle}{\langle \mathcal{L}^2 \rangle^2} = L^4 N_L^{\hat{\phi}}. \quad (6.15)$$

6.1.3 Including the Beam for Unclustered Sources

Including a beam function W_l , just as we have done in Section 4.4, is trivial. The quadratic estimator Eq. (6.11) is modified into

$$\hat{\mu}(\mathbf{m}_\perp) = \frac{1}{\Omega_s C^{\text{shot}}} \frac{1}{N_{\text{vol}}} \sum_{j_\perp} \sum_{j_\parallel} \Delta \tilde{\mathcal{S}}(j) \Delta \tilde{\mathcal{S}}^*(j - \mathbf{m}_\perp), \quad (6.16)$$

with $C^{\text{shot}} = W_l C^{\text{shot}}$, and $\Delta \tilde{\mathcal{S}}$ is the DFT of the source field which includes the beam function, namely

$$\Delta \tilde{\mathcal{S}}(j) = W(j) \Delta \tilde{\mathcal{S}}(j). \quad (6.17)$$

The noise reconstruction noise can be again evaluated from the variance of this estimator,

$$N_L^{\hat{\mu}} = \frac{1}{\Omega_s N_{\text{gal}}} \frac{1}{|W_L|^2} \frac{\langle \mathcal{L}^4 \rangle}{\langle \mathcal{L}^2 \rangle^2}. \quad (6.18)$$

6.2 Simulating Unclustered Point Sources

In this section we will describe the sampling of luminosity points drawn from the Schechter luminosity function distribution through the transformation method (Press et al., 2002).

The method relies on the probability integral transform, which states that if x is a continuous random variable with probability distribution $f(x)$ and cumulative $F(x) = \int_{-\infty}^x dx f(x)$, then the random variable $y = F(x)$ has a uniform distribution in the range $[0, 1]$. Also the inverse property is true: if y is a random uniform number belonging to the range $[0, 1]$ and x has cumulative $F(x)$, then the random variable $F^{-1}(y)$ has the same distribution as x . Notice that the cumulative function has to be invertible, *i.e.* monotonic and right-continuous. In our case the integral of the Schechter function for $\alpha < -1$ is the Gamma function $\Gamma(\alpha + 1)$ which diverges at $x \rightarrow 0$. Our truncated Schechter function (see Appendix F.1 for details) by the way is monotonic and continuous, so we can apply the transformation method.

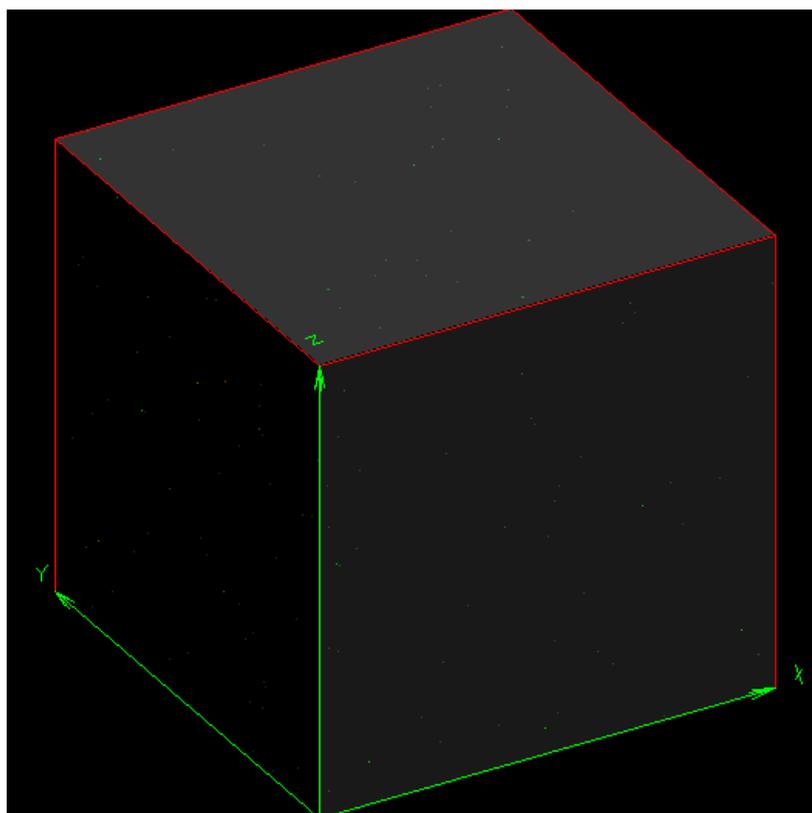
Let us consider a large number of $\mathcal{L}/\mathcal{L}^*$ bins in logarithmic space. We can compute the cumulative of the Schechter function for each bin and table the results. These can thus be normalized to one, in order to have numbers uniformly distributed within the range $[0, 1]$. At this point it is possible to assign a $\mathcal{L}/\mathcal{L}^*$ bin to any point in the grid by randomly generating a number U between the range $[0, 1]$. Every uniform number will correspond to a given luminosity bin from the tabled values of the cumulative. This is made for voxels which are randomly selected in the simulated grid with volume N_{vol} . The process will continue until the number of randomly sampled voxels is equal to the total number of galaxies N_{gal} . This value is set by the integral for the entire $\mathcal{L}/\mathcal{L}^*$ range multiplied by the physical volume of the simulated survey $V(z) = \Omega_s \mathcal{D}^2(z) \mathcal{L}(z)$, with $\mathcal{D}(z)$ and $\mathcal{L}(z)$ the comoving angular distances along the perpendicular and parallel direction respect to the line of sight, respectively.

Moreover, there is also the possibility that a given voxel is selected more than once: this could happen if the resolution of the grid is poor or because of rounding effects. In this case we will sum the values falling inside that voxel.

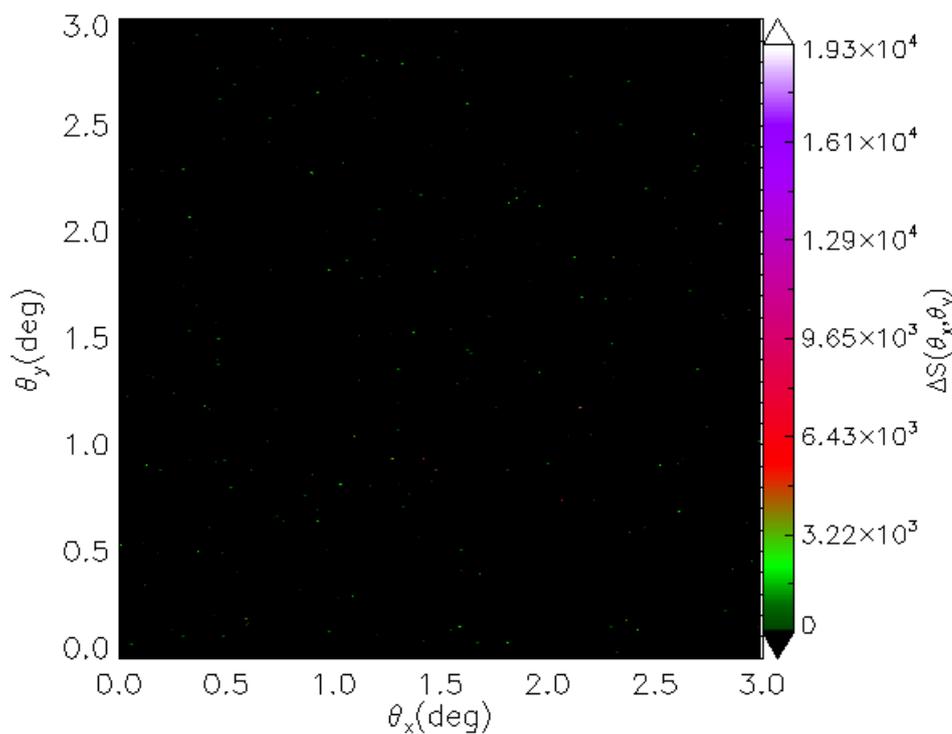
We can see in Figure 6.1 the dimensionless three-dimensional brightness fluctuation field for a volume of 512 voxels with $\Omega_s = 3^\circ \times 3^\circ$ over a bandwidth of 5 MHz centered around a source redshift of $z = 2.5$. In the same picture we show a slice extracted out of it.

6.3 Weak Lensing Estimator for Clustered Sources

We are going to include the Poisson signal due to point sources into our Gaussian formalism explored in Chapter 4 and Appendix D, to form a quadratic estimator for 21 cm weak lensing and reproduced the results already obtained by Pourtsidou & Metcalf (2015). This is because HI is found mainly within galaxies, and we need to model the discreteness of these objects with a Poisson distribution drawn from a Gaussian one, in order to simulate the clustering of galaxies. We will not seek an optimal estimator, but one of the same form as the one we previously used. This means that the estimator filter is not found by imposing the minimum variance requirement and that the estimated potential amplitude is slightly biased.



6.1.1 Unclustered sources field



6.1.2 Unclustered source map

Figure 6.1: *Top panel*: the simulated brightness fluctuation field for unclustered sources, produced for $N_{\text{vol}} = 512^3$ voxels, $\Omega_s = 3^\circ \times 3^\circ$, and a bandwidth of 5 MHz centered around $z = 2.5$. *Bottom panel*: a 2D map extracted out of the above cube.

If we consider the two-point correlation function for a brightness fluctuation clustered point in a pixelized grid, we have, by considering Eq. (6.2)

$$\begin{aligned}\langle \delta S(\mathbf{x}) \delta S^*(\mathbf{x}') \rangle &= \left\langle \sum_{\mathcal{L}} \mathcal{L} \delta n_{\mathbf{x}}^{\mathcal{L}} \sum_{\mathcal{L}'} \mathcal{L}' \delta n_{\mathbf{x}'}^{\mathcal{L}'} \right\rangle \\ &= \langle \mathcal{L}^2 \rangle \bar{\eta} \delta V \delta_{\mathbf{x}\mathbf{x}'}^{\mathbf{K}} + (\bar{\eta} \delta V \langle \mathcal{L} \rangle)^2 \xi_{\mathbf{x}\mathbf{x}'},\end{aligned}\quad (6.19)$$

in which, differently from the case studied in Section 6.1.1, the $i \neq i'$ term gives the clustering correlation. As can also be seen in Appendix G.2.1, dividing by $\bar{S}^2 = (\bar{\eta} \delta V \langle \mathcal{L} \rangle)^2$ we have

$$\langle \Delta S(\mathbf{x}) \Delta S^*(\mathbf{x}') \rangle = \frac{1}{\bar{\eta} \delta V} \frac{\langle \mathcal{L}^2 \rangle}{\langle \mathcal{L} \rangle^2} \delta_{\mathbf{x}\mathbf{x}'}^{\mathbf{K}} + \xi_{\mathbf{x}\mathbf{x}'}, \quad (6.20)$$

where the first part comes from Poisson fluctuations in the number counts and the second from density fluctuations. If we write the correlation between Fourier modes by applying Eq. (6.4), we get

$$\begin{aligned}\langle \Delta S_{\mathbf{k}} \Delta S_{\mathbf{k}'}^* \rangle &= \frac{\Omega_s^2}{N_{\text{vol}}^2} \left[\sum_{\mathbf{x}} \sum_{\mathbf{x}'} \frac{1}{\bar{\eta} \delta V} \frac{\langle \mathcal{L}^2 \rangle}{\langle \mathcal{L} \rangle^2} e^{-i\mathbf{k} \cdot \mathbf{x}} e^{i\mathbf{k}' \cdot \mathbf{x}'} \delta_{\mathbf{x}\mathbf{x}'}^{\mathbf{K}} + \sum_{\mathbf{x}} \sum_{\mathbf{x}'} \xi_{\mathbf{x}\mathbf{x}'} e^{-i\mathbf{k} \cdot \mathbf{x}} e^{i\mathbf{k}' \cdot \mathbf{x}'} \right] \\ &= \Omega_s^2 (P^{\text{shot}} + P_{\mathbf{k}}) \delta_{l,l'}^{\mathbf{K}} \delta_{k_p,k_p'}^{\mathbf{K}} \\ &= \Omega_s (C^{\text{shot}} + C_{l,k_p}) \delta_{l,l'}^{\mathbf{K}} \delta_{k_p,k_p'}^{\mathbf{K}}\end{aligned}\quad (6.21)$$

where in the second passage we applied what has been done for Eq. (D.23), and in the last one we used the angular power spectrum definitions Eqs. (D.24) and (6.9). All the details can be found in Appendix G.2.1.

At this point we proceed as usual by writing down the lensing correlations for this field and excluding the $L = 0$ modes, namely

$$\langle \tilde{\Delta S}_{l,k_p} \tilde{\Delta S}_{l-L,k_p'}^* \rangle = L \cdot [l C_{l,k_p} + (L-l) C_{l-L,k_p} + L C^{\text{shot}}] \Phi(L) \delta_{k_p,k_p'}^{\mathbf{K}}, \quad (6.22)$$

up to first order. The estimator we want to find will be non-optimal, and it has the form

$$\hat{\Phi}(L) = f(L) \sum_{k_p} \sum_l \tilde{\Delta S}_{l,k_p} \tilde{\Delta S}_{l-L,k_p'}^*. \quad (6.23)$$

Note the difference of this expression with Eq. (4.13). By requiring the estimator to be unbiased, we find a filter

$$\begin{aligned}f(L) &= \left\{ \sum_{k_p} \sum_l L \cdot [l C_{l,k_p} + (L-l) C_{l-L,k_p} + L C^{\text{shot}}] \right\}^{-1} \\ &= \left\{ \left[\sum_{k_p} \sum_l L \cdot l C_{l,k_p} + L \cdot (L-l) C_{l-L,k_p} \right] + N_{\text{vol}} L^2 C^{\text{shot}} \right\}^{-1}.\end{aligned}\quad (6.24)$$

The variance of this estimator is $\mathcal{V} = \langle |\hat{\Phi}(L)|^2 \rangle$, and explicitly

$$\begin{aligned}\mathcal{V} &= f^2(L) \sum_l \sum_{l'} \sum_{k_p} \sum_{k_p'} \langle \tilde{\Delta S}_{l,k_p} \tilde{\Delta S}_{l-L,k_p}^* \tilde{\Delta S}_{l',k_p'} \tilde{\Delta S}_{l'-L,k_p'}^* \rangle \\ &= f^2(L) (\mathcal{I}_0 + \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3 + \mathcal{I}_4)\end{aligned}\quad (6.25)$$

where the four point correlation function is computed in Appendix G.2.3 and provides five non null contributions \mathcal{I}_i . The computations are lengthy but not difficult, and the following expressions for the \mathcal{I}_i terms have been simplified by sending $N_{\parallel} \rightarrow \infty$. Thermal noise is considered by sending $C_{l,k_p} \rightarrow C_{l,k_p}^T = C_{l,k_p} + N_l^{\text{Sky}} + N_l^{\text{Rcv}}$.

As made in Appendix G.2.2 and G.2.3, we can define the Poisson trispectrum and bispectrum

$$T^{\text{shot}} = \frac{\Omega_s^3 \langle \mathcal{L}^4 \rangle}{N_{\text{gal}}^3 \langle \mathcal{L} \rangle^4}, \quad B^{\text{shot}} = \frac{\Omega_s^2}{N_{\text{vol}}^2} \frac{1}{\eta^2 \delta V^2} \langle \mathcal{L}^3 \rangle \quad (6.26)$$

respectively, and, remembering the definition of C^{shot} , we can write the \mathcal{I}_i contributions as

$$\mathcal{I}_0 = \sum_l \sum_{l'} \sum_{k_p} \sum_{k'_p} \Omega_s T^{\text{shot}} = N_{\text{vol}}^2 \Omega_s T^{\text{shot}}, \quad (6.27)$$

$$\begin{aligned} \mathcal{I}_1 &= \sum_l \sum_{l'} \sum_{k_p} \sum_{k'_p} \Omega_s B^{\text{shot}} \left[C_{l'-L,k'_p}^T + C_{l',k'_p}^T + C_{l-L,k_p}^T + C_{l,k_p}^T \right] \\ &= 2N_{\text{vol}} \Omega_s B^{\text{shot}} \sum_l \sum_{k_p} \left[C_{l-L,k_p}^T + C_{l,k_p}^T \right], \end{aligned} \quad (6.28)$$

$$\begin{aligned} \mathcal{I}_2 &= \sum_l \sum_{l'} \sum_{k_p} \sum_{k'_p} \Omega_s (C^{\text{shot}})^2 \left[C_{L,0}^T + C_{l-l',k_p-k'_p}^T + C_{l+l'-L,k_p+k'_p}^T \right] \\ &= \Omega_s (C^{\text{shot}})^2 \left\{ N_{\text{vol}}^2 C_{L,0}^T + \sum_l \sum_{l'} \sum_{k_p} \sum_{k'_p} \left[C_{l-l',k_p-k'_p}^T + C_{l+l'-L,k_p+k'_p}^T \right] \right\} \end{aligned} \quad (6.29)$$

$$\begin{aligned} \mathcal{I}_3 &= \sum_l \sum_{l'} \sum_{k_p} \sum_{k'_p} \Omega_s^2 C^{\text{shot}} \left[C_{l',k'_p}^T \delta_{l,l'}^{\text{K}} \delta_{k_p,k'_p}^{\text{K}} \delta_{l,l'}^{\text{K}} \delta_{k_p,k'_p}^{\text{K}} + C_{L-l',-k'_p}^T \delta_{l,L-l'}^{\text{K}} \delta_{k_p,-k'_p}^{\text{K}} \delta_{l,L-l'}^{\text{K}} \delta_{k_p,-k'_p}^{\text{K}} \right. \\ &\quad \left. + C_{l',k'_p}^T \delta_{L-l,l'}^{\text{K}} \delta_{k'_p,-k_p}^{\text{K}} \delta_{L-l,l'}^{\text{K}} \delta_{k'_p,-k_p}^{\text{K}} + C_{l'-L,k'_p}^T \delta_{l,l'}^{\text{K}} \delta_{k_p,k'_p}^{\text{K}} \delta_{l,l'}^{\text{K}} \delta_{k_p,k'_p}^{\text{K}} \right] \\ &= 2\Omega_s^2 C^{\text{shot}} \sum_l \sum_{k_p} \left[C_{l-L,k_p}^T + C_{l,k_p}^T \right], \end{aligned} \quad (6.30)$$

$$\begin{aligned} \mathcal{I}_4 &= \sum_l \sum_{l'} \sum_{k_p} \sum_{k'_p} \Omega_s^2 \left[C_{l,k_p}^T C_{l'-L,k'_p}^T \delta_{l,l'}^{\text{K}} \delta_{k_p,k'_p}^{\text{K}} \delta_{l,l'}^{\text{K}} \delta_{k_p,k'_p}^{\text{K}} + C_{l,k_p}^T C_{l',k'_p}^T \delta_{l,L-l'}^{\text{K}} \delta_{k_p,-k'_p}^{\text{K}} \delta_{l,L-l'}^{\text{K}} \delta_{k_p,-k'_p}^{\text{K}} \right] \\ &= 2\Omega_s^2 \sum_l \sum_{k_p} C_{l-L,k_p}^T C_{l,k_p}^T. \end{aligned} \quad (6.31)$$

We note that \mathcal{I}_0 is a pure Poisson term, while \mathcal{I}_4 is a pure Gaussian term. The others come from a mixed contribution of the twos. We can write down these quantities also in the continuous limit, in order to perform numerically these integrals and gain some computational time. So the filter is

$$f(\mathbf{L}) = \frac{1}{\Omega_s} \left\{ \sum_{k_p} \int \frac{d^2 l}{(2\pi)^2} \left[\mathbf{L} \cdot l C_{l,k_p} + \mathbf{L} \cdot (\mathbf{L} - l) C_{|\mathbf{L}-l|,k_p} + L^2 C^{\text{shot}} \right] \right\}^{-1},$$

while the variance contributions become

$$\mathcal{I}_0 = N_{\parallel}^2 \Omega_s^3 T^{\text{shot}} \left(\int \frac{d^2 l}{(2\pi)^2} \right)^2, \quad (6.32)$$

$$\mathcal{I}_1 = 2N_{\parallel} \Omega_s^3 B^{\text{shot}} \left(\int \frac{d^2 l}{(2\pi)^2} \right) \sum_{k_p} \int \frac{d^2 l}{(2\pi)^2} [C_{l-L, k_p}^T + C_{l, k_p}^T], \quad (6.33)$$

$$\begin{aligned} \mathcal{I}_2 = \Omega_s^3 (C^{\text{shot}})^2 & \left\{ N_{\parallel}^2 \left(\int \frac{d^2 l}{(2\pi)^2} \right)^2 C_{L,0}^T \right. \\ & \left. + \sum_{k_p} \sum_{k'_p} \int \frac{d^2 l}{(2\pi)^2} \int \frac{d^2 l'}{(2\pi)^2} [C_{l-l', k_p-k'_p}^T + C_{l+l'-L, k_p+k'_p}^T] \right\}, \end{aligned} \quad (6.34)$$

$$\mathcal{I}_3 = 2\Omega_s^3 C^{\text{shot}} \sum_{k_p} \int \frac{d^2 l}{(2\pi)^2} [C_{l-L, k_p}^T + C_{l, k_p}^T], \quad (6.35)$$

$$\mathcal{I}_4 = 2\Omega_s^3 \sum_{k_p} \int \frac{d^2 l}{(2\pi)^2} C_{l-L, k_p}^T C_{l, k_p}^T, \quad (6.36)$$

in which we will ignore the $C_{L,0}$ contribution in \mathcal{I}_2 because of foreground cleaning. All these terms have a Ω_s^3 while the filter squared carries a Ω_s^2 factor: the total variance contribution will have a surviving factor Ω_s , and to get the estimator noise we can consider \mathcal{V}/Ω_s .

As already pointed out by Pourtsidou & Metcalf (2015), the dominant term is represented by \mathcal{I}_4 , since it involves the fourth moment of the luminosity function. In fact, the main role in having a low reconstruction noise level is played by the C^{shot} contribution. The \mathcal{I}_2 term is the smallest contribution to the variance. As its pure Gaussian analogous, it is flat up to scales where the thermal noise becomes important and saturates for a certain k_p^{max} . For example, at a redshift $z = 2$, $k_p^{\text{max}} \sim 40$. Hence, another important contribution comes from interferometer specifications to determine the thermal noise spectra.

The most important feature of Poisson fluctuations is that they contribute to both noise and estimator signal. Also in this case we need to determine exactly how many k_p modes have to be neglected because of foreground removal techniques. If the contamination is not serious, only the first modes can be neglected, producing a small deterioration in the signal-to-noise level.

6.3.1 Including the Beam for Clustered Sources

Just like for the unclustered case, the DFT of the beam function, $W(j)$, can be trivially included. The resulting non-optimal estimator is

$$\hat{\Phi}(L) = \mathcal{F}(L) \sum_{k_p} \sum_l \Delta \tilde{\mathcal{S}}_{l, k_p} \Delta \tilde{\mathcal{S}}_{l-L, k_p}^*, \quad (6.37)$$

where $\Delta \tilde{\mathcal{S}}(j) = W(j) \Delta \tilde{\mathcal{S}}(j)$ and the filter is

$$\mathcal{F}(L) = \left\{ \sum_{k_p} \sum_l W_L W_{l-L}^* L \cdot [l C_{l, k_p} + (L-l) C_{l-L, k_p} + L C^{\text{shot}}] \right\}^{-1}. \quad (6.38)$$

6.4 Telescope Model

While for the previous section we considered only a SKA-Low experiment for EoR observations, the telescope aimed to observe the post-EoR frequencies ($\nu_c \sim 450$ MHz) treated in this work is SKA-Mid. A quick overview on SKA-Mid has been given on Section 1.1.2. As already stated SKA-Mid can observe in single-dish (autocorrelation) mode or in interferometer mode.

6.4.1 Single-Dish Mode Thermal Noise

If we consider a single dish effective area A_{eff} , the noise RMS per pixel for a Gaussian process is

$$\sigma_N = \frac{2k_{\text{BT}_{\text{sys}}}}{A_{\text{eff}} \sqrt{N_{\text{pol}} \Delta\nu t_p}}, \quad (6.39)$$

where T_{sys} is the system temperature, thermal sky noise and/or receiver noise, N_{pol} takes into account the possibility of having more than one uncorrelated polarization channels, $\Delta\nu$ is the bandwidth centered around a given observational frequency, and t_p is the time per pointing. Following an analogous argumentation to the one developed in Appendix E, we can write the expression for the noise angular power spectrum

$$C_{l,\Delta\nu}^N = \left[\frac{\lambda^2(z)}{A_{\text{eff}} \Omega_s(\nu)} \right]^2 \frac{T_{\text{sys}}^2(\nu)}{N_{\text{pol}} \Delta\nu t_o} \frac{S_{\text{area}}}{N_b}, \quad (6.40)$$

where we used $t_p = t_o/N_p = t_o \Omega_s/S_{\text{area}}$, since for a given survey area S_{area} , N_p pointings are needed given a total observation time t_o . The instantaneous FoV of the telescope is hence increased using focal plane arrays with multiple phased feeds (PAF). $S_{\text{area}} \geq N_b \Omega_s$, since nothing is gained from observing same parts of the sky. As seen in Appendix E.1, below a certain critical frequency the beams will overlap for PAFs (in order to achieve uniformity on the noise across sky maps), so

$$\Omega_s(\nu) = \Omega_s(\nu_c) \begin{cases} (\nu_c/\nu)^2 & \text{for } \nu \leq \nu_c \\ 1 & \text{for } \nu > \nu_c \end{cases} \quad (6.41)$$

So, the pixel size corresponds to the instantaneous FoV with FWHM

$$\Omega_s \approx \frac{\pi}{8} \left(1.3 \frac{\lambda}{D_{\text{dish}}} \right)^2 [\text{sr}] \approx \text{FoV} \quad (6.42)$$

or smaller. The effective area of one dish is

$$A_{\text{eff}} = \epsilon \pi D_{\text{dish}}^2 / 4, \quad (6.43)$$

where the antenna efficiency $\epsilon \sim 0.7 - 0.8$. So the factor in round brackets in Eq. 6.40 is of order $\sim 1/\epsilon^2$. So, we can write

$$C_{l,\Delta\nu}^N \approx \frac{T_{\text{sys}}^2(\nu)}{\epsilon^2 N_{\text{pol}} \Delta\nu t_o} \frac{S_{\text{area}}}{N_b} \begin{cases} (\nu_c/\nu)^2 & \text{for } \nu \leq \nu_c \\ 1 & \text{for } \nu > \nu_c \end{cases}, \quad (6.44)$$

(Bull, 2015; Santos et al., 2015). So the power spectrum is insensitive to the way we pack the feeds for mosaicking, since this is connected to the pixel resolution that is canceled out in this final expression.

If more than one dish is considered, the total power spectrum is modified with a further factor $1/N_{\text{dish}}$, since the signal can be added incoherently.

6.4.2 Interferometer Mode Thermal Noise

The interferometer thermal noise model is analogous to what we have developed for SKA-Low in Section 5.6, with the main difference that SKA-Mid is not an aperture array, so the collecting area is not frequency dependent. The station elements are hence substituted by N_{dish} dishes of diameter D_{dish} , which cover a primary beam $\Omega_s \approx F\theta V$ and are distributed in visibility space with density function $n(U, \nu)$.

Thus, including the possibility of having multiple beams N_b with PAFs and multiple pointings of the sky $N_p = S_{\text{area}}/[N_b\Omega_s(\nu)]$, the thermal noise power spectrum is

$$C_{l,\Delta\nu}^N = \left[\frac{\lambda^2(z)}{A_{\text{eff}}} \right]^2 \frac{T_{\text{sys}}^2(\nu)}{N_{\text{pol}}\Delta\nu t_o N_b n[U = l/(2\pi), \nu]} \frac{S_{\text{area}}}{\Omega_s(\nu)}, \quad (6.45)$$

(?) (compare with Eq. (5.21)). Note that the time per pointing is increased at lowest frequencies (for a fixed t_o) and the full beam is N_b times the single pixel feed Ω_s .

6.4.3 SKA1 and SKA2-Mid Specifications

The original SKA-Mid design (Dewdney, 2013) considers 2 bands for SKA1-Mid, the first, B1, covering a frequency range $\nu = 350\text{--}1050$ MHz ($z \sim 3.06\text{--}0.35$), while the second, B2, a frequency range of $\nu = 950\text{--}1760$ ($z \sim 0.5\text{--}(0)$). The band B1 has dishes of 15 m diameter each with effective area of 133 m^2 , while B2 dishes have an effective area of 150 m^2 . The primary beam of B1 and B2 SKA1-Mid bands is 1.78 deg^2 and 0.48 deg^2 respectively and the critical frequency is placed at half of the frequency range, so 700 MHz ($z \sim 1$) and 1 GHz ($z \sim 0.42$) respectively. The instrumental noise due to the receiver is $T_{\text{rcv}} = 23$ K for B1 and $T_{\text{rcv}} = 15.5$ K for B2 dishes and there are $N_{\text{pol}} = 2$ polarization channels for each band.

The recent rebaselining reduced the number of MID receiver elements of the 30%, reducing the number of dishes from $N_{\text{dish}} = 190$ to $N_{\text{dish}} = 130$ in both bands.

We plan to mainly use SKA1-Mid B1 in interferometer mode, using a bandwidth of $\Delta\nu = 40$ MHz ($\Delta z \sim 0.345$) centered at $z = 2.5$, a total survey area of 25000 deg^2 and a total observation time of 4000 hrs. Note that with respect to the EoR redshifts case, we can use a larger bandwidth, since the convergence power spectrum variation within this redshift range is small. At $z = 8$ we had to use a much thinner bandwidth. $N_b = N_p = 1$ is going to be preliminarily used.

With this approximation Eq. (6.45) becomes

$$C_{l,\Delta\nu}^N = \left[\frac{\lambda^2(z)}{A_{\text{eff}}} \right]^2 \frac{T_{\text{sys}}^2(\nu)}{2\Delta\nu t_o n(l,\nu)} \quad (6.46)$$

(Pourtsidou et al., 2015). The observation frequency at $z = 2.5$ is $\nu = 405.83$ MHz (corresponding to $\lambda = 0.74$ m), and the sky noise temperature is $T_{\text{sky}} \approx 30.55$ K, and so it has comparable magnitude with the receiver noise for B1 band.

The fiducial dish distribution function for SKA1-Mid is pictured in Figure 6.2.

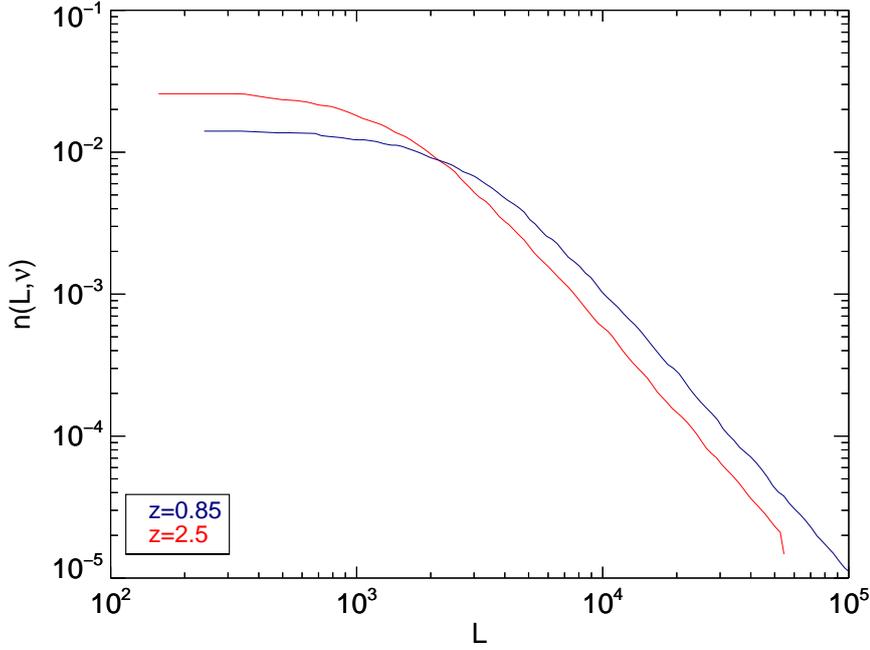


Figure 6.2: The array distribution in visibility space computed at $z = 2.5$ (red) compared to the fiducial one at $z = 0.85$ (blue).

As can be seen $l_{\text{min}} = 300$, corresponding to $\sqrt{\Omega_s} = 1.2^\circ$. Following Eq. (6.42) we deduce that this $n(U, \nu)$ is computed at a fiducial redshift of $z \approx 0.85$, corresponding to a fiducial frequency of $\nu_f = 768$ MHz. From $l_{\text{max}} \approx 10^5$, we deduce that the baseline length at this redshift is of ≈ 6500 m. With these data we computed, using Eq. (E.9), the baseline array density at $z = 2.5$ plotted on Figure 6.2, with a FoV $\Omega_s \approx 4.36 \text{ deg}^2$, $l_{\text{min}} \approx 172$, and $l_{\text{max}} \approx 52735$.

Again the effect of rebaselining can be modeled on the $n(U, \nu)$, since it is proportional to the square of the number of dishes. This means that the noise Eq. 6.46 is increased by a factor $(190/130)^2 \sim 2.14$.

The thermal noise is shown in Figure 6.3, for both sky and receiver noise at $z = 2.5$.

6.5 The 21 cm Power Spectrum Model at Post-EoR Redshifts

In the previous chapter we used Eqs. (4.5) and (4.11) to define the 21 cm temperature brightness fluctuation power spectrum. This expression is only

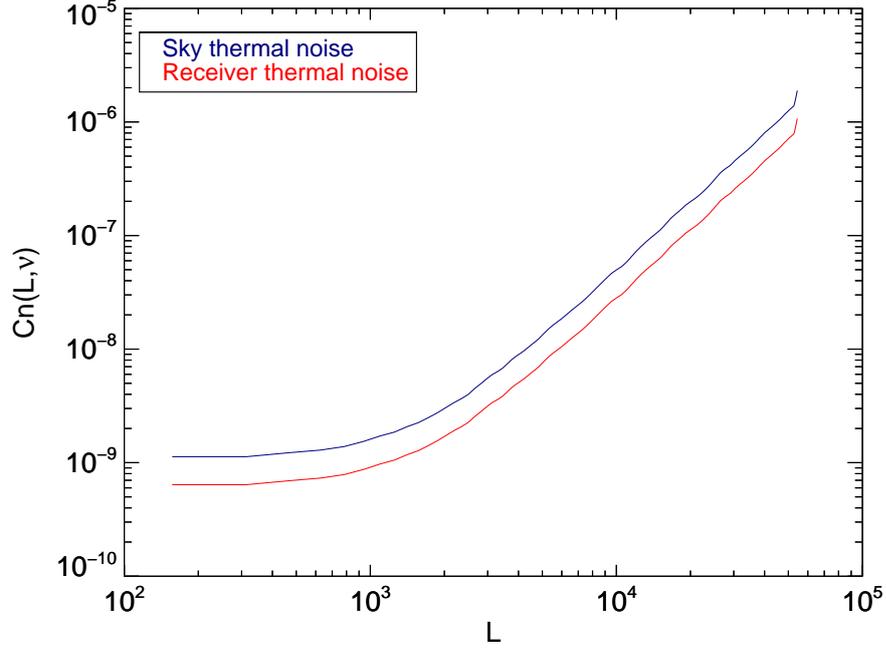


Figure 6.3: The thermal noise power spectrum contributions from sky synchrotron (blue) and receiver (red) in $[mK]^2$ units. This has been computed for $z = 2.5$, $\Delta\nu = 40$ MHz, 4000 hrs, and the SKA1-Mid specifications described in this section.

valid at high redshifts ($z > 6$) and for $T_s \gg T_{\text{CMB}}$ ($z < 15$, which is still our case). We will consider a slightly different expression at the redshifts considered in this section ($z \sim 2.5$).

To find it let us consider Eq. (3.9), *i.e.* the expression for the optical depth of a HI cloud. If we write the HI number density as $n_{\text{HI}}(z) = \rho_{\text{HI}}(z)/M_{\text{HI}}$, with M_{HI} the mass of an HI atom, we can write the mean temperature brightness fluctuation from the second line of Eq. (3.10), namely

$$\bar{T}(z) = \frac{3}{32\pi} \frac{hc^3 A_{10}}{k_B v_{21}^2 M_{\text{HI}}} \left(\frac{\rho_{\text{HI}}(z)}{1+z} \right) \left(\frac{dv_{\parallel}}{dr_{\parallel}} \right)^{-1}, \quad (6.47)$$

where all the quantity of interest here have been already defined in Section 3.1. Now consider that for the comoving volume element dV

$$\frac{dV(z)}{dz d\Omega} = \frac{c\chi^2}{H(z)} = c \left(\frac{\chi}{1+z} \right)^2 \frac{(1+z)^2}{H(z)} = c \mathcal{D}^2(z) \frac{(1+z)^2}{H(z)}, \quad (6.48)$$

where we applied Eq. 2.20, the definition of the angular diameter distance. Since $dV(z) = \mathcal{D}^2(z) d\Omega dr_{\parallel}$, we can write

$$\frac{dr_{\parallel}}{dz} = \frac{c}{H(z)} (1+z)^2. \quad (6.49)$$

Knowing that

$$\frac{dz}{1+z} = \frac{dv}{c}, \quad (6.50)$$

the gradient of the proper velocity along the LoS can be written as

$$\frac{dv_{\parallel}}{dr_{\parallel}} = \frac{H(z)}{(1+z)^3}. \quad (6.51)$$

Now the mean 21 cm temperature brightness fluctuation is

$$\bar{T}(z) = \left(\frac{3}{32\pi} \frac{hc^3 A_{10}}{k_B v_{21}^2 M_{\text{HI}}} \right) \frac{(1+z)^2 \rho_{\text{HI}}(z)}{H(z)}. \quad (6.52)$$

If we define the density parameter of HI at the redshift z as

$$\Omega_{\text{HI}}(z) = \frac{\rho_{\text{HI}}(z)}{\rho_{\text{crit}}}, \quad (6.53)$$

where $\rho_{\text{crit}} = 3H_0^2 / (8\pi G) = 2.7755h^2 10^{11} M_{\odot}$ is the present day critical density of the Universe. Substituting into Eq. (6.52), we finally obtain

$$\bar{T}(z) = \left(\frac{9}{256\pi^2} \frac{hc^3 A_{10} H_0}{k_B v_{21}^2 M_{\text{HI}} G} \right) \frac{(1+z)^2 \Omega_{\text{HI}}(z)}{E(z)} = (180 \text{ mK}) \times \Omega_{\text{HI}}(z) \frac{(1+z)^2 h}{E(z)}, \quad (6.54)$$

(Battye et al., 2012).

With Eq. 6.54 we can now write down the expression for our 21 cm temperature brightness fluctuation angular power spectrum, analogously to Eq. (4.11) and including the redshift space distortion term,

$$C_{l,k_p} = \frac{P_{\Delta T}(k, z)}{\mathcal{D}^2 \mathcal{L}} = \frac{\bar{T}^2(z)}{\mathcal{D}^2 \mathcal{L}} (1 + f\mu_k^2)^2 P_{\delta}(k, z) \quad (6.55)$$

where $k = \sqrt{(l/\mathcal{D})^2 + (2\pi k_p/\mathcal{L})^2}$, $\mu_k = k_{\parallel}/k$ and f is the logarithmic derivative of the linear growth rate with respect to the scale factor $a = (1+z)^{-1}$, namely

$$f = \frac{d \ln D}{d \ln a} \simeq \Omega_m^{0.55}(z), \quad (6.56)$$

where

$$\Omega_m(z) = \frac{\Omega_m(1+z)^3}{\Omega_m(1+z)^3 + (1 - \Omega_m - \Omega_{\Lambda})(1+z)^2 + \Omega_{\Lambda}}. \quad (6.57)$$

6.5.1 HI Evolution

As seen from Eq. (6.53), the density parameter involves the knowledge of HI density at redshift z . Since at post-EoR redshifts the HI is mostly confined within galaxies, $\rho_{\text{HI}}(z)$ is related to the Schechter mass function introduced in Appendix F.1 through

$$\rho_{\text{HI}}(z) = \Phi^*(z) M^*(z) \int \left(\frac{M}{M^*} \right)^{\alpha+1} e^{-M/M^*} d\left(\frac{M}{M^*} \right) = \Phi^*(z) M^*(z) \Gamma(\alpha + 2). \quad (6.58)$$

As stated in Section 6.1.1, an important systematic uncertainty in this work could come from the assumption made on $\Phi^*(z)$ and $M^*(z)$, since their measurements are available only in the local Universe. To take into account a redshift evolution of these parameters we can write

$$\Omega_{\text{HI}}(z) = \Omega_{\text{HI}}^{\text{no-ev}} \frac{\Phi^*(z)}{\Phi^*(z=0)} \frac{M^*(z)}{M^*(z=0)}, \quad (6.59)$$

where $\Omega_{\text{HI}}^{\text{no-ev}} = 4.9 \times 10^{-4}$ is the value of the density parameter assuming no redshift evolution of the Schechter function's parameters.

As pointed out by Santos et al. (2015), at low redshifts $\Omega_{\text{HI}}(z)$ is measured using 21 cm observations directly from galaxies. At high redshifts, it is estimated by computing the HI associated with Damped Ly α systems observed in absorption in quasar spectra. These systems are easy to identify, given their prominent damping wings in both high-resolution and low-resolution data even at low signal-to-noise, and a HI column density is inferred by Voigt profile fitting. This is in turn easily translated into a value for $\Omega_{\text{HI}}(z)$. Present constraints infer a constant $\Omega_{\text{HI}}(z)$ at $z = 2 - 4$, while at higher redshift this value is expected to increase, as the Universe is becoming more neutral. For a recent summary of observed trends we refer the reader to consult Padmanabhan et al. (2015).

Pourtsidou & Metcalf (2015) used Eq. (6.54) to study the possibility of having high SNR measurements for lensing reconstruction adopting a few HI redshift evolution models. They found that the no-evolution scenario is the most conservative choice, because either an increase in $\Phi^*(z)$ or $M^*(z)$ increases the detectability of of lensing. They obtain an high SNR even in this case, and such result motivates us to assume this scenario for our purposes.

6.5.2 A Bias for HI Tracing

In Section 4.2 we have already discussed the possibility concerning the presence of a bias between matter and HI power spectra in Eq. (4.11). This bias would change, up to linear order, the observed power spectrum according to

$$P_{\Delta T}(k, z) \propto b_{\text{HI}}^2(z) P_{\delta}(k, z). \quad (6.60)$$

The signal will then be completely specified once we find a prescription for the HI density and bias function $b_{\text{HI}}(z)$. This can be obtained by making use of the halo mass function, dn/dM and relying on a model for the amount of HI mass in a dark matter halo of mass M , namely $M_{\text{HI}}(M)$, so that

$$b_{\text{HI}}(z) = \rho_{\text{HI}}^{-1} \int_{M_{\text{min}}}^{M_{\text{max}}} dM \frac{dn(M, z)}{dM} M_{\text{HI}}(M, z) b(M, z), \quad (6.61)$$

where $b(M, z)$ is the bias function for the halo. Some concerns might arise due to the possible stochastic behaviour of the function $M_{\text{HI}}(M)$ or its dependence with the environment (so that it would be a function of position also). However, given the low resolution pixels used in HI intensity mapping experiments,

we expect a large number of HI galaxies per pixel, which should average-down any fluctuations and allow us to take the above deterministic relation for the mass function (Santos et al., 2015). For example, for the typical scales we are interested in Cosmology, one needs angular/frequency resolutions of around 1 degree and 5 MHz respectively, which translates into a comoving volume of $\sim 105\text{Mpc}^3$. In each volume element, we expect a total of around 106 DM halos with mass between $10^8 - 10^{15} M_\odot$, and ~ 31000 with masses between 5×10^9 and $1 \times 10^{12} M_\odot$ (where the latter range corresponds to halos expected to contain most of the HI mass). This supports our assumption of a position-independent HI mass function due to the averaging over many halos. Some level of stochasticity could still increase the shot noise of the signal, but this is expected to be quite small. Moreover a more evolved option would be to consider the proportionality of the mass function to the mass as a function of redshift. This would at least guarantee the fit to the density measurements by construction.

Another issue is whether we can assume that the bias is scale dependent. Again, as long as we restrict ourselves to large scales, this should be a reasonable assumption since we are averaging over many galaxies. Results from simulations show that the bias can be safely assumed constant for $k < 1h/\text{Mpc}$ at high redshifts (while at $z < 1$, it should be safe for $k < 0.1h/\text{Mpc}$). Note that this bias can also be modelled using a variety of relatively simple prescriptions on top of the outputs of large volume and high resolution hydrodynamic or N-body simulations (Villaescusa-Navarro et al., 2014).

For the moment we will keep assuming, within our conservative no-evolution model, that the HI traces unbiasedly the matter fluctuations, considering also that the reconstructed signal is mainly at large scales, neglecting a possible redshift and scale dependence. This assumption will be tested more accurately on future works.

6.5.3 The Total Clustered Angular Power Spectrum

From what has been derived in Sections 6.1.1 and 6.5 the total power spectrum for clustered sources is

$$C_{l,k_p}^{\text{clust}} = C_{l,k_p} + \bar{T}^2(z)C^{\text{shot}}, \quad (6.62)$$

where C_{l,k_p} is defined in Eq. (6.55), and C^{shot} is defined in Eq. (6.9).

The result is pictured in Figure 6.4 for $z = 2.5$, $\Delta\nu = 40$ MHz and compared to the power spectra produced by sky and receiver thermal noises described in Section 6.4.3. We can see that the first $k_p = 20$ modes are dominated by the clustering contribution and are above the noise level up to $L \sim 1000$, which is, more or less, the point in which the signal turns to be Poisson noise dominated. All modes with $k_p \gtrsim 35$ appear to be under the thermal noise level and will be unusable for lensing reconstruction.

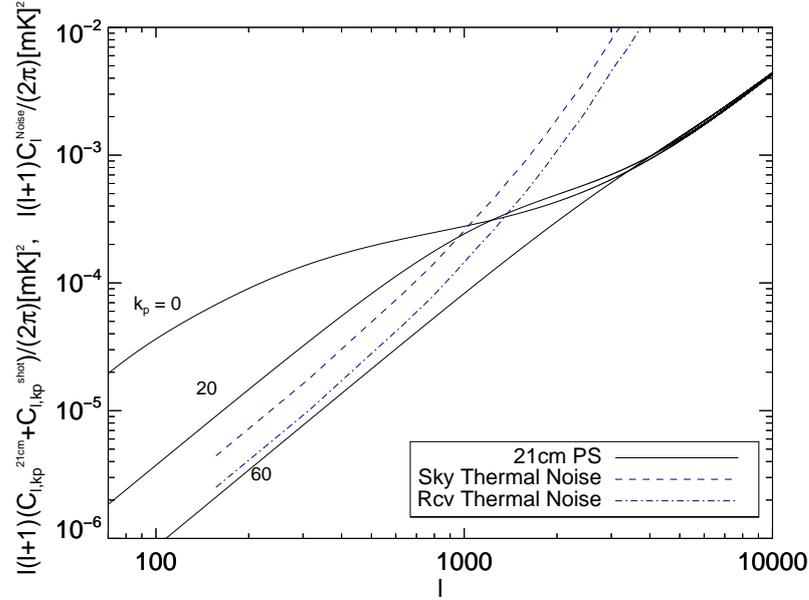


Figure 6.4: The 21 cm power spectrum including the Poisson noise contribution in solid line, computed for $z = 2.5$ and using an observational bandwidth of $\Delta\nu = 40$ MHz. This has been also compared to the SKA1-Mid sky (dashed line) and receiver (dot-dashed line) power spectra with specifications described in Section 6.4.3.

6.6 Resuming of Results and Future Developments (pt. II)

Here we will resume the main results obtained in this chapter, in order to list briefly the future plans for simulating post-EoR 21 cm lensing detections. In this chapter:

- we have adopted the theoretical framework constructed at EoR redshifts in order to review the discrete weak lensing reconstruction formalism at post-EoR redshifts studied by Pourtsidou & Metcalf (2015). Including the point source contribution as an additive discrete Poisson noise to a clustering Gaussian three-dimensional signal, they showed that it contributes to improve the lensing reconstruction signal-to-noise;
- because of this formalism, the expression of a discrete non-optimal Fourier space quadratic estimator can be found in order to reconstruct the underlying lensing mass, for both clustered and unclustered sources. Finally, we have modified these expression by including the beam of the telescope, as for the analogous estimators implemented at EoR redshifts;
- unclustered point sources have been simulated in order to study the behaviour of this estimator and to move towards more realistic clustered signal reconstruction simulations;
- in order to explore the post-EoR epoch, we have also modeled in our code a more accurate power spectrum for the 21 cm brightness temperature

fluctuation field at low redshifts, with a formalism which will allow for studying different HI evolutionary models;

- we developed a thermal noise model for the lower frequency band SKA-Mid in interferometer mode, including the possibility to observe simultaneously different parts of the sky.

Moreover, we need to complete the studies performed in this chapter, and we plan to address the following points in the near future:

- we will simulate the clustered field by drawing random temperature points from the total distribution at each pixel of the three-dimensional grid. This will be lensed, contaminated with noise, and it will include the effect of the telescope's beam;
- the estimators for unclustered and clustered sources can be tested and computed within our simulation code, in order to compute results analogous to the ones obtained in the previous chapter, for single and multiple frequency band detections, and for different survey strategies;
- we plan to investigate the effects on our results caused by important and controversial issues like the possible evolution of the HI signal or the inclusion of a bias which modifies the proportionality relation between matter density fluctuations and neutral hydrogen;
- the effects of foreground subtraction need to be investigated at post-EoR redshifts, even if they are much less important at these frequencies;
- we will provide an estimate on the accuracy achieved for lensing power spectrum measurements at these redshift, with the reconstruction noise computed using the SKA-Mid model developed in this chapter. With SKA-Mid larger survey areas become available and better constraints on power spectrum accuracy can be obtained. Also in this case, it is indeed possible to point different sky patches in a reasonable time, improving the statistics of large-scale modes.
- we aim to cross-correlate the lensing field with galaxies or HI density fields, since it has been found that such a measurement could produce excellent results for SKA-Mid in interferometer mode and even using the MeerKAT pathfinder.

Chapter 7

Conclusions

In this work we have seen how 21 cm lensing can be a leading cosmological probe during the next decade. Using the forthcoming observations from the SKA and other radio telescopes, a huge amount of cosmological information can potentially be extracted over a wide range of redshifts, in order to constrain the standard Λ CDM paradigm, and shed a new light on the dark sector of our Universe. The innovative technique of Intensity Mapping treats the 21 cm brightness temperature fluctuations as a continuous three-dimensional field, opening up the possibility of using alternative analysis methods similar to those successfully applied to the CMB. Thanks to the narrow channel bandwidths of modern radio receivers, one automatically measures redshifts with high precision too, bypassing one of the most difficult aspects of performing a galaxy redshift survey.

We investigated the potentialities offered by the weak gravitational lensing of the 21 cm brightness temperature fluctuation field. We considered two different cosmological epochs, namely a typical EoR redshift ($z = 8$) in which the HI is fully ionized, and at a post-EoR redshift ($z = 2.5$), when HI is mostly found within galaxies.

To this purpose, we implemented a simulation pipeline capable of dealing with issues that can not be treated analytically, like the simulation of a telescope beam, the non-uniform visibility space coverage, the non-linearity of the lensing source field, and the discreteness of visibility measurements. Moreover, in the theoretical and numerical framework established in this work, it is possible to include and investigate other complicated issues regarding our ignorance about the reionization process history, like the non Gaussianity of the 21 cm source in the considered EoR redshift range. In fact, it is very likely that EoR was a non-homogeneous process expanded over a considerable redshift range, and the detected signal strongly depends on the number density of ionized regions which are causing inhomogeneities in 21 cm temperature signal that is not possible to investigate analytically.

Another important non-analytic issue our code is designed to handle concerns foreground subtraction techniques. With the pipeline developed in this work we can implement foreground contamination and study how foreground removal techniques can affect the accuracy of our results. These methods would produce residuals and cross-correlations among different frequencies,

and their influence can be treated only numerically.

After having briefly discussed the SKA telescope and the concept of Intensity Mapping, in the first part of this work we have reviewed the standard cosmological model and its connections with gravitational lensing, in order to introduce the main quantities and observables useful to study the weak gravitational lensing of 21 cm radiation. For the same reason, we then introduced the physics of the 21 cm line, briefly discussing its evolution through the cosmological epochs and the great efforts made by the astrophysics community to probe the Epoch of Reionization.

Then, in the main part of this work we have presented the weak gravitational lensing of the 21 cm brightness temperature fluctuation field at typical EoR redshifts. By taking advantage of the 21 cm source signal division into multiple statistically independent maps along the frequency direction, we have demonstrated how the lensing mass distribution can be reconstructed with high fidelity using a three dimensional optimal quadratic lensing estimator in Fourier space. This would provide a great opportunity to correlate mass with visible objects and test the dark matter paradigm.

We successfully implemented the 3D Fourier space quadratic estimator in our simulation code, taking into account the beam of the telescope (set by the baseline maximum dimensions) and the discreteness of visibility measurements, paving the way for future numerical studies aimed to investigate more realistic issues. We showed that the discrete 21 cm estimator can be employed by using a single frequency band or by combining multiple frequency band measurements as well.

Considering the current SKA plans, we studied the performance of the quadratic estimator for detections aimed to observe EoR redshifts, for different observational strategies and comparing two thermal noise models for the SKA-Low array: the first is widely used in the literature but assumes a uniform visibility space distribution, while the second takes into account a more realistic array density. These noise models have been added to simulated lensed 21 cm brightness temperature fluctuation maps, produced by interpolating on the grid the lensed positions of the temperature maps. To accomplish this task we followed the weak lensing assumption widely used in the CMB case, and valid for the 21 cm field as well at the scales considered in this study.

Assuming a realistic non-uniform distribution of SKA-Low stations, we found that an SKA-Low interferometer should obtain high-fidelity images of the underlying mass distribution in its phase 1 only if several bands are stacked together, covering a redshift range from $z = 7$ to $z = 11.6$ and with a total resolution of 1.6 arcmin. We also implemented a simple de-noising procedure in order to filter out the small-scale noise which is likely to strongly contaminate the estimated signal. The SKA-Low phase 2, modeled in order to improve the sensitivity of the instrument by almost an order of magnitude, should be capable of providing reconstructed images with good quality even when the signal is detected within a single frequency band. In this case the reconstructed image has a resolution of 1.15 arcmin at $z = 8$, within a field of view of 13 deg^2 .

Considering the serious effect that foregrounds could have on these detections (by making the first few k_p modes unusable), we discussed the limits of

these results as well as the possibility of measuring an accurate lensing power spectrum. In the case of multi-band detection of the lensed 21 cm signal made with an SKA2-Low telescope model we found constraints close to the sample variance ones in the range $L < 1000$, even for a small field of view such as a 25 deg^2 survey area. Good constraints have been found also for SKA1-Low in multi-band detection, and for SKA2-Low in single band detection.

We also explored the possibility to detect a cluster lensing signal coming from redshift $z = 0.5$ with a mass of $M = 10^{15} M_{\odot}$, but we found their signal to be overwhelmed by the estimator reconstruction noise by several orders of magnitude, going well below the saturation limit of the noise imposed by sample variance also for multi-band analysis.

In the last part of this work we adopted the theoretical framework constructed at EoR redshifts in order to review the weak lensing discrete reconstruction formalism at post-EoR redshifts, namely at $z = 2.5$. With Intensity Mapping, discrete point sources need to be resolved only in frequency and can be added incoherently to the clustered 21 cm signal. Following the study performed by Pourtsidou & Metcalf (2015), their contribution is included as an additive discrete Poisson noise to a clustering Gaussian three-dimensional signal, demonstrating that they contribute to improve the lensing reconstruction signal-to-noise by computing an expression for a discrete non-optimal Fourier space quadratic estimator for both clustered and unclustered sources.

Unclustered point sources have been simulated in order to study the behaviour of this estimator and to move towards more realistic clustered signal reconstruction simulations. The estimator for unclustered and clustered sources can be tested within our simulation code, and, as made for the EoR observations, we have modified the estimator in order to include the beam of the telescope and we considered a non-uniform antennae distribution in our simulation framework. In order to explore the post-EoR epoch, we have also modeled in our code a more accurate power spectrum for the 21 cm brightness temperature fluctuation field, with a formalism which will allow for studying different HI evolution models.

We developed a thermal noise model for the lower frequency band SKA-Mid in interferometer mode, including the possibility to observe simultaneously different parts of the sky. With SKA-Mid larger survey areas become available and better constraints on power spectrum accuracy can be obtained. It is indeed possible to point different sky patches in a reasonable time, improving the statistics of large-scale modes. Moreover, it has been studied that cross-correlating the lensing field with galaxies or HI density field measurements could produce excellent results for SKA-Mid and even using the MeerKAT prototype. The code developed in this work is capable of treating this issue as well, providing more accurate estimates on what SKA could measure.

To conclude, weak gravitational lensing of 21 cm emission could offer an exciting opportunity to enhance our understanding of the Epoch of Reionization. Furthermore, detecting 21 cm lensing in post-EoR redshifts using the intensity mapping method would be an important science achievement of this technique, and it would give us the chance to detect lensing from higher redshifts than traditional optical galaxy surveys. The above possibilities are very

important for cosmology, but they are challenged by problems like foreground contamination and surveys limitations. However, the nascent field of 21cm cosmology holds great promise for the future, and the simulation we developed can be a useful tool in order to investigate and resolve some of the aforementioned issues. As soon as those issues are dealt with, the SKA and its pathfinders can provide us with new data which will push the boundaries of our understanding, contributing to make times like these unique and exciting for the entire astrophysical community.

Appendix A

Gaussian Random Fields

In this appendix the properties of Gaussian Random Fields (GRF) will be explored, as well as the Fourier Transform convention adopted through this entire work.

A.1 Fourier Transforms Convention

Throughout this work we will adopt the following definition for the Fourier Transform (FT)

$$\delta(\mathbf{k}) = \int d^3x \delta(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}, \quad (\text{A.1})$$

while the Inverse Fourier Transform is defined as

$$\delta(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} \delta(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}. \quad (\text{A.2})$$

Both real and Fourier space functions are uniquely defined as representations of a function defined in Hilbert space.

A.1.1 Hermitianity Condition

Since $\delta(\mathbf{x})$ is real, its Fourier dual satisfies the Hermitianity condition

$$\delta(-\mathbf{k}) = \delta^*(\mathbf{k}). \quad (\text{A.3})$$

A.1.2 Dirac Delta Function

We define the Dirac delta function as

$$\delta^{\text{D}}(\mathbf{x} - \mathbf{x}') = \int_{-\infty}^{+\infty} \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}, \quad (\text{A.4})$$

such that

$$\delta(\mathbf{x}) = \int_{-\infty}^{+\infty} d^3x' \delta^{\text{D}}(\mathbf{x} - \mathbf{x}') \delta(\mathbf{x}'). \quad (\text{A.5})$$

A.1.3 Convolution Theorem

We define $h(\mathbf{x})$ as the convolution of two functions f and g

$$h(\mathbf{x}) = [f \otimes g](\mathbf{x}) = \int d^3 x' f(\mathbf{x}') g(\mathbf{x} - \mathbf{x}'). \quad (\text{A.6})$$

Let us compute the Fourier Transform of this quantity,

$$\begin{aligned} h(\mathbf{k}) &= \int d\mathbf{x} \left[\int d^3 x' f(\mathbf{x}') g(\mathbf{x} - \mathbf{x}') \right] e^{-i\mathbf{k} \cdot \mathbf{x}} \\ &= \int d\mathbf{x} \left[\int d^3 x' f(\mathbf{x}') \int d^3 x'' g(\mathbf{x}'') \delta^{\text{D}}(\mathbf{x} - \mathbf{x}' - \mathbf{x}'') \right] e^{-i\mathbf{k} \cdot \mathbf{x}} \\ &= \int d\mathbf{x} \left[\int d^3 x' f(\mathbf{x}') \int d^3 x'' g(\mathbf{x}'') \int \frac{d^3 k'}{(2\pi)^3} e^{i\mathbf{k}' \cdot (\mathbf{x} - \mathbf{x}' - \mathbf{x}'')} \right] e^{-i\mathbf{k} \cdot \mathbf{x}} \\ &= \int \frac{d^3 k'}{(2\pi)^3} \int d\mathbf{x} e^{i\mathbf{x} \cdot (\mathbf{k}' - \mathbf{k})} \int d^3 x' f(\mathbf{x}') e^{-i\mathbf{k}' \cdot \mathbf{x}'} \int d^3 x'' g(\mathbf{x}'') e^{-i\mathbf{k}' \cdot \mathbf{x}''} \\ &= \int d^3 k' \delta^{\text{D}}(\mathbf{k} - \mathbf{k}') f(\mathbf{k}') g(\mathbf{k}') = f(\mathbf{k}) g(\mathbf{k}). \end{aligned} \quad (\text{A.7})$$

We can extend this property to the convolution of n functions, namely

$$h(\mathbf{x}) = [f_1 \otimes \cdots \otimes f_n](\mathbf{x}) = \int d^3 x_1 f_1(\mathbf{x}_1) \cdots \int d^3 x_n f_n(\mathbf{x}_n) \delta^{\text{D}}\left(\mathbf{x} - \sum_{i=1}^n \mathbf{x}_i\right), \quad (\text{A.8})$$

with Fourier Transform

$$h(\mathbf{k}) = f_1(\mathbf{k}) \cdots f_n(\mathbf{k}). \quad (\text{A.9})$$

Inversely, a real space product $h(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$ can be seen as a Fourier space convolution. Therefore,

$$h(\mathbf{k}) = \int \frac{d^3 k'}{(2\pi)^3} f(\mathbf{k}') g(\mathbf{k} - \mathbf{k}') = (2\pi)^{-3} [f \otimes g](\mathbf{k}). \quad (\text{A.10})$$

Again, we can generalise the Fourier transformation of the product of n real functions $h(\mathbf{x}) = f_1(\mathbf{x}) \cdots f_n(\mathbf{x})$ with the convolution in Fourier space

$$\begin{aligned} h(\mathbf{k}) &= \int \frac{d^3 k_1}{(2\pi)^3} f_1(\mathbf{k}_1) \cdots \int d^3 k_n f_n(\mathbf{k}_n) \delta^{\text{D}}\left(\mathbf{k} - \sum_{i=1}^n \mathbf{k}_i\right) \\ &= (2\pi)^{-3(n-1)} [f_1 \otimes \cdots \otimes f_n](\mathbf{k}). \end{aligned} \quad (\text{A.11})$$

A.2 GRF Statistics

We have seen in Section 2.1.5 that, from Inflation theory, the primordial matter fluctuations are assumed to be a homogeneous and isotropic Gaussian Random Field (GRF). This means that the probability distribution of these fluctuations is Gaussian. Following (Coles & Lucchin, 2002), let us consider a large

number N of realisations of our periodic volume denoted by V_1, V_2, \dots, V_N , and a Fourier field

$$\delta_{\mathbf{k}} = |\delta_{\mathbf{k}}| e^{i\vartheta_{\mathbf{k}}} = \mathcal{R}(\delta_{\mathbf{k}}) + i\mathcal{I}(\delta_{\mathbf{k}}), \quad (\text{A.12})$$

which across this ensemble of realisations. This field has a homogeneous and isotropic distribution, so that both real and imaginary parts of the field are mutually independent and have the following Gaussian probability distribution

$$\mathcal{P}[\mathcal{R}(\delta_{\mathbf{k}}), \mathcal{I}(\delta_{\mathbf{k}})] = \sqrt{\frac{V}{2\pi\sigma_k^2}} e^{-[\mathcal{R}(\delta_{\mathbf{k}}), \mathcal{I}(\delta_{\mathbf{k}})]^2 V / 2\sigma_k^2}, \quad (\text{A.13})$$

where the variance $\sigma_k^2 = P_\delta/2 = |\delta_{\mathbf{k}}|^2/2$ is the power spectrum of the field. This is equivalent to assuming that the phases $\vartheta_{\mathbf{k}}$ are mutually independent and randomly distributed over the uniform range $[0, 2\pi]$, so that the moduli of Fourier field amplitudes are Rayleigh distributed. Namely

$$\mathcal{P}(|\delta_{\mathbf{k}}|, \vartheta_{\mathbf{k}}) d|\delta_{\mathbf{k}}| d\vartheta_{\mathbf{k}} = \frac{|\delta_{\mathbf{k}}| V}{2\pi P_\delta} e^{-|\delta_{\mathbf{k}}|^2 V / 2P_\delta} d|\delta_{\mathbf{k}}| d\vartheta_{\mathbf{k}}. \quad (\text{A.14})$$

Because of homogeneity and isotropy, the field will depend only on $k = |\mathbf{k}|$, and not on its direction. The Fourier inverse transform of this Rayleigh distribution is Gaussian,

$$\mathcal{P}[\delta(\mathbf{x})] d\delta(\mathbf{x}) = \sqrt{\frac{1}{2\pi\sigma^2}} e^{-\delta^2(\mathbf{x})/2\sigma^2} d\delta(\mathbf{x}). \quad (\text{A.15})$$

We can assume for the central limit theorem that, if the phases are randomly distributed, the field will be Gaussian for a large number of modes, since $\delta(\mathbf{x})$ is a sum over a large number of Fourier modes. For a large number N of volumes V_i , *i.e.* a large number of realisations of the Universe, one will find that $\delta(\mathbf{k})$ varies from one to the other in both amplitude and phase. Although the mean value of the perturbation across the statistical ensemble is null, its variance is not and will be proportional to the power spectrum of these fluctuations, carrying information only on the amplitude of these perturbations, not on their spatial structure. The definition of the power spectrum is

$$\langle \delta(\mathbf{k}) \delta^*(\mathbf{k}') \rangle = \langle |\delta(\mathbf{k})|^2 \rangle = (2\pi)^3 \delta^{\text{D}}(\mathbf{k} - \mathbf{k}') P_\delta(k), \quad (\text{A.16})$$

the ensemble average, over all universes, of the square amplitude of the Fourier mode $\delta(\mathbf{k})$. For the reality condition Eq.(A.3), we also have

$$\langle \delta(\mathbf{k}) \delta(\mathbf{k}') \rangle = \langle |\delta(\mathbf{k})|^2 \rangle = (2\pi)^3 \delta^{\text{D}}(\mathbf{k} + \mathbf{k}') P_\delta(k). \quad (\text{A.17})$$

So, main requirements for a GRF are variance proportional to field power spectrum and randomly distributed phases.

If $\delta(\mathbf{x})$ is a GRF, the N -variate joint distribution of a set $\delta_i = \delta(x_i)$ can be written as a multivariate Gaussian distribution

$$\mathcal{P}_N(\delta_i) = \sqrt{\frac{\|C^{-1}\|}{(2\pi)^N}} e^{\mathbf{V}^T \cdot C^{-1} \cdot \mathbf{V} / 2}, \quad (\text{A.18})$$

where $C = \langle \delta_i \delta_j \rangle$, and V is the column vector made from δ_i . Here $\langle \delta_i \rangle = 0$, but its variance is the correlation function

$$\langle \delta(\mathbf{x}_i) \delta(\mathbf{x}_j) \rangle = \xi(|\mathbf{x}_i - \mathbf{x}_j|) = \xi r_{ij}, \quad (\text{A.19})$$

where the average is taken over all the spatial positions. $\xi(r)$ depends only the distance modulus between two connected points because of homogeneity and isotropy.

A.3 The Wiener-Khintchine Theorem

The two-point correlation function $\xi(r)$ is related to the power spectrum $P(k)$. The two-point correlation function of density fluctuations is

$$\xi(r) = \frac{\langle [\rho(\mathbf{x}) - \bar{\rho}] [\rho(\mathbf{x} + \mathbf{r}) - \bar{\rho}] \rangle}{\bar{\rho}^2} = \langle \delta(\mathbf{x}) \delta(\mathbf{x} + \mathbf{r}) \rangle, \quad (\text{A.20})$$

where $\bar{\rho} = \langle \rho \rangle$, $\mathbf{x} + \mathbf{r} = \mathbf{x}'$, and the average is taken over all possible positions \mathbf{x} . Let us write the power spectrum using FT definitions

$$\begin{aligned} P(k) &= \langle \delta(\mathbf{k}) \delta^*(\mathbf{k}') \rangle \\ &= \left\langle \int d^3 x \int d^3 x' \delta(\mathbf{x}) e^{-i\mathbf{k} \cdot \mathbf{x}} \delta^*(\mathbf{x}') e^{i\mathbf{k}' \cdot \mathbf{x}'} \right\rangle \\ &= \int d^3 x \int d^3 x' \langle \delta(\mathbf{x}) \delta(\mathbf{x} + \mathbf{r}) \rangle e^{-i\mathbf{k} \cdot \mathbf{x}} e^{-i\mathbf{k}' \cdot (\mathbf{x} + \mathbf{r})} \\ &= \int d^3 r \int d^3 x \langle \delta(\mathbf{x}) \delta(\mathbf{x} + \mathbf{r}) \rangle e^{-i(\mathbf{k} - \mathbf{k}') \cdot \mathbf{x}} e^{-i\mathbf{k}' \cdot \mathbf{r}} \\ &= \int d^3 r \xi(r) \delta^D(\mathbf{k} - \mathbf{k}') e^{-i\mathbf{k}' \cdot \mathbf{r}} \\ &= \int d^3 r \xi(r) e^{-i\mathbf{k} \cdot \mathbf{r}}. \end{aligned} \quad (\text{A.21})$$

Thus the power spectrum and the two-point correlation function are related by a Fourier Transform operation. If we write Eq.(A.21) in spherical coordinates, we get

$$\begin{aligned} P(k) &= \int_0^\infty r^2 dr \xi(r) \left(\int_0^{2\pi} d\varphi \int_0^\pi d\vartheta e^{-ikr \cos \vartheta} \sin \vartheta \right) \\ &= \int_0^\infty r^2 dr \xi(r) \left(\int_0^{2\pi} d\varphi \int_{-1}^1 d\mu e^{-ikr\mu} \right) \\ &= 4\pi \int_0^\infty r^2 dr \xi(r) \frac{\sin(kr)}{kr}. \end{aligned} \quad (\text{A.22})$$

Its inverse relation will be

$$\xi(r) = \frac{1}{2\pi^2} \int_0^\infty k^2 dk P(k) \frac{\sin(kr)}{kr}, \quad (\text{A.23})$$

which for $r \rightarrow 0$ shows the variance $\xi(0)$ of the fluctuation field as the zero-th order moment. This means that the power spectrum provides a complete statistical description of the density field as long as it is Gaussian. Hence, higher-order spectral moments are defined as

$$\sigma_i^2 = \frac{1}{2\pi^2} \int_0^\infty dk P(k) k^{2(i+1)}. \quad (\text{A.24})$$

So, the zero-th order moment is the variance; third-order moment is called skewness, while fourth-order one is called kurtosis. These moments contain information about the shape of $P(k)$, just as moments of a probability distribution contain information about its shape.

A.4 Higher-Order Statistics

If we consider pair of points the correlation function Eq.(A.19) is the two-point correlation function for a continuous GRF defined in Eq.(A.20). It can be negative, having anti-correlation between two points in this case. It is also possible to define spatial covariance functions for $N > 2$, where N is the number of correlated points. For example the three-point correlation function is

$$\zeta(r, s, t) = \langle \delta(\mathbf{x})\delta(\mathbf{x} + \mathbf{r})\delta(\mathbf{x} + \mathbf{s}) \rangle \quad (\text{A.25})$$

with $|\mathbf{r} - \mathbf{s}| = t$. The mean is taken over all points \mathbf{x} , defining all possible triangles with sides r, s and t . Generalising to $N > 3$ is trivial. The full covariance function will contain terms depending on expectations values of lower order, called disconnected part, while the unpartitioned part of the N -point covariance function is called connected part. So, the correlation function of three points is

$$\langle \delta_1 \delta_2 \delta_3 \rangle = \langle \delta_1 \rangle \langle \delta_2 \delta_3 \rangle + \langle \delta_2 \rangle \langle \delta_1 \delta_3 \rangle + \langle \delta_3 \rangle \langle \delta_1 \delta_2 \rangle + \langle \delta_1 \delta_2 \delta_3 \rangle_c. \quad (\text{A.26})$$

For a GRF all the odd N expectations values are zero, so the quantity $\langle \delta_1 \delta_2 \delta_3 \rangle$ is null and $\langle \delta \rangle = 0$. The even-order connected parts can be expressed as combinations of GRF pairs $\delta_i \delta_j$ for Wick's Theorem. So, using the Wiener-Khintchine Theorem, the Fourier Transform of an N -point correlation function is proportional to terms involving only power spectrum products. In general the Fourier transform of the three-point correlation function is called bispectrum

$$\langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3) \rangle = (2\pi)^3 \delta^D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B(k_1, k_2, k_3), \quad (\text{A.27})$$

while the Fourier transform of a four-point correlation function, defined analogously to bispectrum, is the trispectrum. In particular, searching for non-Gaussianities means searching for non-null odd-order correlation functions. For more details consult Bartolo et al. (2007).

We can also define our correlation function in terms of a discrete distribution of masses, and link it to the probability of finding an interacting pair of point-masses in a given volume. The density field is $\rho(\mathbf{x}) = \sum_i m_i \delta^D(\mathbf{x} - \mathbf{x}_i)$.

Hence, the mean density of the points in a homogeneous and isotropic space is $n = \bar{\rho}/m$, the probability of finding a particle in an infinitesimal volume dV is $dP = n dV$. The joint probability of finding two particles, one in volume dV_1 and another in volume dV_2 is

$$dP = n^2 dV_1 dV_2 + n^2 \xi(r_{12}) dV_1 dV_2 = n^2 \frac{\langle \rho(\mathbf{x}) \rho(\mathbf{x} + \mathbf{r}) \rangle}{\bar{\rho}} dV_1 dV_2, \quad (\text{A.28})$$

where $\xi(r_{12})$ is the excess probability of finding the second particle at a distance r_{12} . So, observed a particle in dV_1 , the probability of finding a second particle placed at a distance r_{12} in dV_2 is $dP(1|2) = n [1 + \xi(r_{12})] dV_2$. This is called the Poisson clustering model. In general it is possible to define the total N -point correlation function $\xi^{(N)}$, which contains contributions from correlations of order $\leq N$. This is

$$d^N P = n^N [1 + \xi^{(N)}(\mathbf{r})] dV_1 \dots dV_N, \quad (\text{A.29})$$

where \mathbf{r} stands for all the r_{ij} separating N points. Again, let us consider a three-point correlation function

$$d^3 P_3 = n^3 [1 + \xi(r_{12}) + \xi(r_{13}) + \xi(r_{23}) + \zeta_{123}] dV_1 dV_2 dV_3, \quad (\text{A.30})$$

where ζ_{123} is the part which does not depend on two-point correlation function. The difference with the continuous δ case is that $\langle \rho \rangle = 1$. So a four-point correlation function will contain terms in ζ_{ijkl} , $\xi_i \xi_{jkl}$, and ξ_{ij} , which have to be subtracted from the connected part η_{1234} .

A.5 Limber's Equation in Fourier Space

Throughout this entire work we will have to deal with angular power spectra, *i.e.* the two-dimensional power spectra projected on the sky, whose coordinates are denoted by the vector $\boldsymbol{\theta}$, from a three-dimensional field power spectrum. This field is generally assumed to be homogeneous and isotropic, like the density contrast field $\delta [f_K(\chi_c \boldsymbol{\theta}, \chi_c)]$ whose coordinates form a local comoving isotropic Cartesian reference frame. Let us consider i projections $g_i(\boldsymbol{\theta})$ of this three-dimensional field δ along the light-cone corresponding to an observer at $\chi_c = 0$ and $t = t_0$. These two quantities are related by a projection kernel $q_i(\chi_c)$ as

$$g_i(\boldsymbol{\theta}) = \int_0^{\chi_c(z_H)} d\chi_c q_i(\chi_c) \delta [f_K(\chi_c \boldsymbol{\theta}, \chi_c)], \quad (\text{A.31})$$

and the projections $g_i(\boldsymbol{\theta})$ are homogeneous and isotropic as well as the full three-dimensional field. The correlation function between two projections is

$$C_{12}(\boldsymbol{\theta}) = \langle g_1(\boldsymbol{\theta}) g_2(\boldsymbol{\theta}') \rangle = \int_0^{\chi_c(z_H)} d\chi_c q_1(\chi_c) \int_0^{\chi_c(z_H)} d\chi'_c q_2(\chi'_c) \langle \delta [f_K(\chi_c \boldsymbol{\theta}, \chi_c)] \delta [f_K(\chi'_c \boldsymbol{\theta}', \chi'_c)] \rangle. \quad (\text{A.32})$$

At large scales $k \rightarrow 0$ and the power spectrum decreases like $P_\delta \propto k$. Hence we can assume that after a certain scale r_c there are no fluctuations, and the

correlation will not be null only for scales $|\chi_c - \chi'_c| \leq r_c$ within the horizon distance $\chi_c(z_H)$. In other words, albeit δ evolves cosmologically, we consider it to be constant over a time scale on which light travels across a comoving distance r_c . This means that the kernel functions $q_i(\chi_c)$ do not vary appreciably over a scale $\Delta\chi_c \leq r_c$, and we can write $f_K(\chi'_c) \approx f_K(\chi_c)$ and $q_i(\chi'_c) = q_i(\chi_c)$. The result is Limber's Equation (Limber, 1954)

$$C_{12}(\theta) = \int_0^{\chi_c(z_H)} d\chi_c q_1(\chi_c) q_2(\chi_c) \int_0^{\chi_c(z_H)} d\chi'_c C^{\delta\delta} \left[\sqrt{f_K^2(\chi_c)\theta^2 + (\Delta\chi_c)^2}, \chi_c \right], \quad (\text{A.33})$$

where the second argument in $C^{\delta\delta}$ denotes the dependence of the correlation function on cosmic time.

For our purposes, the Fourier space form found by (Kaiser, 1992) of this equation will be more interesting. Applying Fourier Transform in Eq.(A.32), we have

$$C_{12}(\theta) = \int d\chi_c q_1(\chi_c) \int d\chi'_c q_2(\chi'_c) \int \frac{d^3k}{(2\pi^3)} \int \frac{d^3k'}{(2\pi^3)} \times \langle \delta(k, \chi_c) \delta^*(k', \chi'_c) \rangle e^{-if_K(\chi_c)\mathbf{k}_\perp \cdot \boldsymbol{\theta}} e^{-if_K(\chi'_c)\mathbf{k}'_\perp \cdot \boldsymbol{\theta}'} e^{-ik_3\chi_c} e^{-ik'_3\chi'_c}. \quad (\text{A.34})$$

The Fourier vector \mathbf{k} has been decomposed in its perpendicular and parallel to LoS components (\mathbf{k}_\perp, k_3) , as well as the comoving distance vector $[f_K(\chi_c)\boldsymbol{\theta}, \chi_c]$. Now the correlation can be replaced by the power spectrum, introducing a $\delta^D(\mathbf{k} - \mathbf{k}')$ which carries out the integration over k' . Applying Limber's approximation for $q_i(\chi_c)$ and $f_K(\chi_c)$, we get

$$C_{12}(\theta) = \int d\chi_c q_1(\chi_c) q_2(\chi_c) \int \frac{d^3k}{(2\pi^3)} P_\delta(k, \chi_c) e^{-if_K(\chi_c)\mathbf{k}_\perp \cdot (\boldsymbol{\theta} - \boldsymbol{\theta}')} e^{-ik_3\chi_c} \int d\chi'_c e^{ik'_3\chi'_c}. \quad (\text{A.35})$$

The last integral can be written as $2\pi \delta^D(k_3)$: only such modes will contribute to the projected correlation function. Splitting the k integral in k_\perp and k_3 integrals (and solving for this latter), we obtain

$$\begin{aligned} C_{12}(\theta) &= \int d\chi_c q_1(\chi_c) q_2(\chi_c) \int \frac{d^2k_\perp}{(2\pi)^2} P_\delta(k_\perp, \chi_c) e^{-if_K(\chi_c)\mathbf{k}_\perp \cdot \boldsymbol{\theta}} \\ &= \int d\chi_c q_1(\chi_c) q_2(\chi_c) \int \frac{k dk}{2\pi} P_\delta(k_\perp, \chi_c) J_0[f_K(\chi_c)\theta k]. \end{aligned} \quad (\text{A.36})$$

If we define the multipole vector $\mathbf{l} = \mathbf{k}_\perp \chi_c$ as the dual Fourier vector mode of $\boldsymbol{\theta}$, we can write the projected two-dimensional power spectrum as the Fourier transform of Eq.(A.36), namely

$$\begin{aligned} C_{12}(\mathbf{l}) &= \int d^2\theta C_{12}(\theta) e^{i\mathbf{l} \cdot \boldsymbol{\theta}} \\ &= \int d\chi_c q_1(\chi_c) q_2(\chi_c) \int \frac{d^2k_\perp}{(2\pi)^2} P_\delta(k_\perp, \chi_c) (2\pi)^2 \delta^D[\mathbf{l} - f_K(\chi_c)\mathbf{k}_\perp] \\ &= \int d\chi_c \frac{q_1(\chi_c) q_2(\chi_c)}{f_K^2(\chi_c)} P_\delta\left(\frac{\mathbf{l}}{f_K(\chi_c)}, \chi_c\right), \end{aligned} \quad (\text{A.37})$$

which is our final result.

A.5.1 About Limber's Approximation Accuracy

Limber approximation is accurate only for small angular scales, and only for quantities which are integrated over a broad redshift range. For example, large deviation may appear when the convergence field is correlated with galaxies within a very thin redshift slice, or when we study signatures in the power spectrum which appear mainly or uniquely at very large scales, like non-Gaussianities (Jeong, 2010). For $l \approx 10$ we could have errors of about 10% up to 1% at $l \approx 100$, when the exact and the Limber expressions are compared.

We used Limber's approximation to compute the convergence power spectrum Eq. (2.118). Jeong (2010) found the exact result for this

$$C_l^{\kappa\kappa} = \frac{2}{\pi} \int_0^{z_s} dz \int_0^{z_s} dz' \frac{\rho_0^2}{\Sigma_c(z, z_s)H(z)\Sigma_c(z', z_s)H(z')} \times \int dk k^2 P_\delta(k, z, z') J_l[k f_K(\chi_c)\theta] J_l[k f_K(\chi'_c)\theta], \quad (\text{A.38})$$

where $\Sigma_c(z, z_s)$ is the critical surface density. Results show that Limber's approximation works very well for a wide range of source redshifts. For $l > 10$ the error caused by Limber's approximation is always much smaller than $\sim 1\%$.

Appendix B

Toy Models for Gravitational Lensing

In general there is a variety of models describing compact objects like planets, stars, black holes, or Massive Astrophysical Compact Halo Objects (MACHOs). The most simple models are axially symmetric, but they are not sufficient for an accurate description of the real lens system. For example they would need an elliptical component or perturbation terms in the lensing potential. These models have to be intended mostly as toy models with which we tested various parts of our code.

In this appendix we will briefly describe the models adopted to better explain weak lensing observables in Section 2.2.2 and test our code part for lensing simulation through bicubic interpolation in Section 5.5.3. Moreover we will describe the NFW model used to model a cluster signal we attempted to recover through our discrete estimator in Section 5.5.7. Through this appendix we will follow mainly Meneghetti (2011) and Narayan & Bartelmann (1996).

B.1 Axially Symmetric Lenses

All the models that will be introduced in next sections are axially symmetric lenses. Their surface density is independent on the position angle with respect to lens center, so $\Sigma(\xi) = \Sigma(|\xi|) = \Sigma(\xi)$, with $\xi = \mathcal{D}_L \theta$. The lens equations are one-dimensional, since all the light rays from a source lie on the same plane passing through the center of the lens, the source and the observer. Under this assumption, the deflection angle Eq. 2.38, is

$$\alpha(\xi) = \frac{\xi}{\xi^2} \frac{4G}{c^2} 2\pi \int_0^\xi d\xi' \xi' \Sigma(\xi') = \frac{4GM(\xi)}{c^2 \xi^2} \xi, \quad (\text{B.1})$$

where $M(\xi)$ is the projected mass enclosed in a radius ξ . So the mass outside this radius has no effect on the deflection. Identifying ξ with the axis passing through the lens center, we get

$$\alpha(\xi) = \frac{4GM(\xi)}{c^2 \xi}, \quad (\text{B.2})$$

or in dimensionless quantities with $\xi = \xi_0 x$

$$\alpha(x) = \frac{\mathcal{D}_L \mathcal{D}_{L_s}}{\xi_0 \mathcal{D}_s} \frac{4GM(\xi_0 x)}{c^2 \xi} \frac{\pi \xi_0}{\pi \xi_0} = \frac{M(\xi_0 x)}{\pi \xi_0^2 \Sigma_{\text{crit}}} \frac{1}{x} = \frac{m(x)}{x}, \quad (\text{B.3})$$

in which we defined the dimensionless mass $m(x)$. Its definition follows from deflection's one, namely

$$\alpha(x) = \frac{2}{x} \int_0^x dx' x' \kappa(x') \Rightarrow m(x) = 2 \int_0^x dx' x' \kappa(x'). \quad (\text{B.4})$$

The lens equation can be written as

$$y = x - \frac{m(x)}{x}. \quad (\text{B.5})$$

To find the Jacobian we need to write the deflection angle as a vector. Considering $\boldsymbol{x} = (x_1, x_2)$, the deflection angle points towards the lens center, so

$$\boldsymbol{\alpha}(\boldsymbol{x}) = \frac{m(x)}{x^2} \boldsymbol{x}. \quad (\text{B.6})$$

Its derivatives are

$$\begin{aligned} \frac{\partial \alpha_1}{\partial x_1} &= \frac{dm(x)}{dx} \frac{x_1^2}{x^3} + m(x) \frac{x_2^2 - x_1^2}{x^4}, \\ \frac{\partial \alpha_2}{\partial x_2} &= \frac{dm(x)}{dx} \frac{x_2^2}{x^3} + m(x) \frac{x_1^2 - x_2^2}{x^4}, \\ \frac{\partial \alpha_1}{\partial x_2} &= \frac{\partial \alpha_2}{\partial x_1} = \frac{dm(x)}{dx} \frac{x_1 x_2}{x^3} + 2m(x) \frac{x_1 x_2}{x^4}, \end{aligned} \quad (\text{B.7})$$

and from these it follows that

$$\mathcal{J} = I - \frac{m(x)}{x^4} \begin{pmatrix} x_2^2 - x_1^2 & -2x_1 x_2 \\ -2x_1 x_2 & x_1^2 - x_2^2 \end{pmatrix} - \frac{dm(x)}{dx} \frac{1}{x^3} \begin{pmatrix} x_1^2 & x_1 x_2 \\ x_1 x_2 & x_2^2 \end{pmatrix}. \quad (\text{B.8})$$

This allows us to compute shear and convergence components:

$$\begin{aligned} \kappa(x) &= \frac{dm(x)}{dx} \frac{1}{2x}, \\ \gamma_1(x) &= \frac{1}{2} (x_2^2 - x_1^2) \left(\frac{2m(x)}{x^4} - \frac{dm(x)}{dx} \frac{1}{x^3} \right), \\ \gamma_2(x) &= x_1 x_2 \left(\frac{dm(x)}{dx} \frac{1}{x^3} - \frac{2m(x)}{x^4} \right). \end{aligned} \quad (\text{B.9})$$

So, we can see that

$$\gamma(x) = \frac{m(x)}{x^2} - \kappa(x) = \bar{\kappa}(x) - \kappa(x), \quad (\text{B.10})$$

with

$$\bar{\kappa}(x) = \frac{m(x)}{x^2} = 2\pi \frac{\int_0^x dx' x' \kappa(x')}{\pi x^2} \quad (\text{B.11})$$

the mean surface mass density within x . The Jacobian determinant of the lens mapping is

$$\begin{aligned} \det A &= \frac{y}{x} \frac{dy}{dx} = \left(1 - \frac{m(x)}{x^2} \right) \left[1 - \frac{d}{dx} \left(\frac{m(x)}{x} \right) \right] \\ &= \left(1 - \frac{m(x)}{x^2} \right) \left(1 + \frac{m(x)}{x^2} - 2\kappa(x) \right) = \left(1 - \frac{\alpha(x)}{x} \right) \left(1 - \frac{d\alpha(x)}{dx} \right). \end{aligned} \quad (\text{B.12})$$

This tells us that axially symmetric lenses with monotonically decreasing mass have at most two critical lines, which are the circles $m(x) = x^2$ (tangential) and $d(m(x)/x)/dx = 1$ (radial).

B.2 Point Mass

The Point Mass (PM) is the simplest axially symmetric lens model, in which all the mass is concentrated in one point. In this situation only the mass inside a radius ξ affects the deflection angle, and the equations for such a lens are valid also for any region outside a spherical mass distribution.

For a PM lens with mass M , the surface density is $\Sigma(\xi) = M \delta^D(\xi)$ with $\xi = \mathcal{D}_L \theta$. So, from Eq. (2.38), we get

$$\alpha(\xi) = \frac{4GM}{c^2} \frac{\xi}{\xi^2}, \quad (\text{B.13})$$

with $\xi = |\xi|$. The lens equation is

$$\beta = \theta - \frac{4GM}{c^2} \frac{\mathcal{D}_{L_s}}{\mathcal{D}_L \mathcal{D}_s} \frac{\theta}{\theta^2} = \theta - \theta_E^2 \frac{\theta}{\theta^2}, \quad (\text{B.14})$$

where we have defined the Einstein radius

$$\theta_E = \sqrt{\frac{4GM}{c^2} \frac{\mathcal{D}_{L_s}}{\mathcal{D}_L \mathcal{D}_s}}, \quad (\text{B.15})$$

which represents the distance from the PM to the critical line where the magnification is infinite. A source behind the center of the lens will be stretched into a ring of radius θ_E . There is no preferred direction, since for a positive β -axis source also θ will lie on the positive θ -axis. So the problem is reduced to one dimension and the lens equation can be written in its adimensional form, namely

$$y = x - \frac{1}{x}, \quad (\text{B.16})$$

with $x = \theta/\theta_E$ and $y = \beta/\theta_E$. Its solutions are

$$x_{\pm} = \frac{1}{2} \left[y \pm \sqrt{y^2 - 4} \right], \quad (\text{B.17})$$

so two images for any source, irrespective of its distance y from the lens. A third image is absent because the mass is singular and time-delay surface is not continuously deformed. To recover the Einstein radius definition one has to set $y = 0$, *i.e.* a source directly behind the lens.

For a PM, we have $m(x) = \theta_E^2$ and $\bar{\kappa}(x) = (\theta_E/\theta)^2 = 1/x^2$. So the magnification can be computed as for any axially-symmetric lens from the Jacobian

$$\begin{aligned} \mu^{-1} &= \det \mathcal{J} = \frac{y}{x} \frac{\partial y}{\partial x} = \left(1 - \frac{\alpha}{x}\right) \left(1 - \frac{\partial \alpha}{\partial x}\right) \\ &= \left(1 - \frac{1}{x^2}\right) \left(1 + \frac{1}{x^2}\right) = 1 - \left(\frac{1}{x}\right)^4. \\ \mu &= \left[1 - \left(\frac{1}{x}\right)^4\right]^{-1}. \end{aligned} \quad (\text{B.18})$$

So when $x \rightarrow 1$, and so $\theta = \theta_E$, $\mu \rightarrow \infty$, and this defines the critical line. For a general source position x_{\pm} ,

$$\mu_{\pm} = \frac{1}{2} \pm \frac{y^2 + 2}{2y\sqrt{y^2 + 4}}. \quad (\text{B.19})$$

For $y \rightarrow \infty$, $\mu_{\pm} \rightarrow (1, 0)$, so for large angular separations between source and lens, one image always disappears because is demagnified, while the other is undistinguishable from the source because it is placed at the same position and has its same flux.

The total magnification is

$$\mu = |\mu_+| + |\mu_-| = \frac{y^2 + 2}{y\sqrt{y^2 + 4}}, \quad (\text{B.20})$$

and the magnification ratio is $|\mu_-/\mu_+| = x_-/x_+$.

Since the deflection for a PM is

$$\alpha(\theta) = \theta_E^2 \frac{\theta}{\theta^2}, \quad (\text{B.21})$$

with magnitude

$$\alpha(\theta) = \theta_E^2 \frac{1}{\theta}, \quad (\text{B.22})$$

the lensing potential of a PM is

$$\Phi(\theta) = \theta_E^2 \ln|\theta|, \quad (\text{B.23})$$

since $\nabla_{\theta} \ln \theta = \theta/\theta^2$. So the convergence is, applying the Poisson's equation Eq. (2.45),

$$\kappa(\theta) = \frac{1}{2} \nabla_{\theta}^2 \Phi(\theta) = \frac{1}{2r} \frac{\partial}{\partial \theta} \left(\theta \frac{\partial \Phi(\theta)}{\partial \theta} \right) = 0. \quad (\text{B.24})$$

A PM model is more appropriate for a compact object than for a galaxy or a cluster where the angular size of the mass distribution is comparable to changes in the deflection potential.

B.3 Singular Isothermal Sphere

The Singular Isothermal Sphere (SIS) model approximate better than PM the effect of a galaxy or a cluster. The density profile for this model is

$$\rho(r) = \frac{\sigma_v^2}{2\pi G r^2}, \quad (\text{B.25})$$

where σ_v is the one-dimensional velocity dispersion of the constituent self-gravitating particles of which the lens is made of. This distribution is Maxwellian at every r , and so this system is in thermal and hydrostatic equilibrium. This distribution leads to flat rotation curves, as observed for spiral galaxies. This

expression has a non-physical singularity at $r = 0$ which is usually avoided introducing a core radius over which the distribution is truncated.

If we project the 3D density along the line of sight, we get the surface density

$$\Sigma(\xi) = 2 \frac{\sigma_v^2}{2\pi G} \int_0^\infty \frac{dz}{\xi^2 + z^2} = \frac{\sigma_v^2}{\pi G} \frac{[\arctan(z/\xi)]_0^\infty}{\xi} = \frac{\sigma_v^2}{2G\xi}. \quad (\text{B.26})$$

If our length scale on the lens plane is ξ_0 , namely

$$\xi_0 = 4\pi \frac{\sigma_v^2}{c^2} \frac{\mathcal{D}_{L_s} \mathcal{D}_L}{\mathcal{D}_s}, \quad (\text{B.27})$$

we get

$$\Sigma(x) = \frac{\sigma_v^2}{2G\xi} \frac{\xi_0}{\xi_0} = \frac{1}{2x} \frac{c^2}{4\pi G} \frac{\mathcal{D}_s}{\mathcal{D}_L \mathcal{D}_{L_s}} = \frac{\Sigma_{\text{cr}}}{2x}. \quad (\text{B.28})$$

This defines the convergence for the SIS

$$\kappa(x) = \frac{1}{2x} = \frac{\theta_E}{2\theta}. \quad (\text{B.29})$$

So the lensing potential is

$$\Phi(\theta) = \theta_E \theta, \quad (\text{B.30})$$

and consequently the deflection angle is

$$\alpha(\theta) = \theta_E \frac{\theta}{\theta}, \quad (\text{B.31})$$

with magnitude

$$\alpha = \theta_E. \quad (\text{B.32})$$

In this case the Einstein radius is defined as

$$\theta_E = \frac{4\pi\sigma_v^2}{c^2} \frac{\mathcal{D}_{L_s}}{\mathcal{D}_s}, \quad (\text{B.33})$$

with

$$\sigma_v^2 = \frac{GM(\theta_E)}{2\xi}, \quad (\text{B.34})$$

where $M(\theta_E)$ is the mass enclosed in the Einstein radius. The lens equation is again one-dimensional and reads

$$y = x - \frac{x}{|x|} \rightarrow \beta = \theta - \theta_E \frac{\theta}{\theta}. \quad (\text{B.35})$$

When $\beta < \theta_E$ or $y < 1$, two solutions of the lens equation exist. These are

$$\theta_{\pm} = \beta \pm \theta_E. \quad (\text{B.36})$$

Thus the lens, the source, and the images at θ_{\pm} lie on a straight line. Technically a third image with zero flux is located at $\theta = 0$. The flux becomes finite if we introduce the core radius to make the singularity disappear.

To compute shear and magnification we need to compute the derivatives of the potential. Expressed in terms of dimensionless x , we have

$$\frac{\partial\Phi}{\partial x_i} = \frac{x_i}{x} \quad \frac{\partial\Phi}{\partial x_i \partial x_j} = \frac{\delta_{ij}x - x_i x_j / x}{x^2} = \frac{\delta_{ij}x^2 - x_i x_j}{x^3}. \quad (\text{B.37})$$

So,

$$\Phi_{,11} = \frac{x_1^2}{x^3} \quad \Phi_{,12} = -\frac{x_1 x_2}{x^3} \quad \Phi_{,22} = \frac{x_2^2}{x^3}. \quad (\text{B.38})$$

This leads to shear components

$$\begin{aligned} \gamma_1 &= \frac{1}{2}(\Phi_{,11} - \Phi_{,22}) = \frac{1}{2} \frac{x^2(\sin^2 \phi - \cos^2 \phi)}{x^3} = -\frac{1}{2} \frac{\cos 2\phi}{x}, \\ \gamma_2 &= \Phi_{,12} = -\frac{\sin \phi \cos \phi}{x} = -\frac{1}{2} \frac{\sin 2\phi}{x}. \end{aligned} \quad (\text{B.39})$$

Thus,

$$\gamma(x) = \sqrt{\gamma_1^2 + \gamma_2^2} = \frac{1}{2x} = \kappa(x). \quad (\text{B.40})$$

The magnification is

$$\mu_{\pm} = \frac{\theta_{\pm}}{\beta} = 1 \pm \frac{\theta_E}{\beta} = \left(1 \mp \frac{\theta_E}{\theta_{\pm}}\right)^{-1}. \quad (\text{B.41})$$

Hence, if the source lies outside the Einstein ring, there is only one image at $\theta = \theta_+ = \beta + \theta_E$. Sources at large distance from the lens can only be weakly magnified by it.

B.3.1 Softened Isothermal Sphere

The singularity of a SIS is avoided if we introduce a core region with angular radius θ_c in which the density is defined. This means that the potential is

$$\Phi = \theta_E \sqrt{\theta^2 + \theta_c^2}. \quad (\text{B.42})$$

So, the deflection angle is

$$\alpha(\theta) = \theta_E \frac{\theta}{\sqrt{\theta^2 + \theta_c^2}}, \quad (\text{B.43})$$

while shear and convergence are

$$\kappa = \theta_E \frac{\theta^2 + 2\theta_c^2}{2(\theta^2 + \theta_c^2)^{3/2}}, \quad \gamma_{1,2} = -\theta_E \frac{\theta^2}{2(\theta^2 + \theta_c^2)^{3/2}} (\cos 2\phi, \sin 2\phi). \quad (\text{B.44})$$

B.4 The Navarro-Frenk-White Lens Model

N-body simulations of hierarchical clustering of dark matter haloes predicts a certain mass profile, the Navarro-Frenk-White (NFW) density profile. This is

$$\rho(r) = \frac{\rho_s}{r/r_s (1 + r/r_s)^2}, \quad (\text{B.45})$$

where the scale density ρ_s is the normalisation of this profile and r_s is a scale radius. This density is defined within the wide mass range $3 \times 10^{11} \lesssim M_{\text{vir}}/(h^{-1} M_{\odot}) \lesssim 3 \times 10^{15}$. The logarithmic slope of this profile changes from -1 at the center to -3 for large radii. So, this is flatter than that of a SIS in the inner part of the halo, and steeper in the outer part.

These quantities are often described in terms of the concentration parameter $c = r_{200}/r_s$, with r_{200} being the radius of the sphere in which the average density is 200 times the critical density and the enclosed mass is M_{200} . Its value is

$$r_{200} = 1.63 \times 10^{-2} \left(\frac{M_{200}}{h^{-1} M_{\odot}} \right)^{1/3} \left[\frac{\Omega_0}{\Omega(z)} \right]^{-1/3} (1+z)^{-1} h^{-1} \text{ Kpc}. \quad (\text{B.46})$$

The mass of the cluster is linked to the concentration parameter via the relation $M = 4\pi r_s^3 \rho_s [\ln(1+c) - c/(1+c)]$. This is because

$$\rho_s = \frac{200}{3} \rho_{\text{crit}} \frac{c^3}{[\ln(1+c) - c/(1+c)]}. \quad (\text{B.47})$$

Simulations show that $r_s(z)$ change with mass systematically, in such a way that concentration is a characteristic function of M_{200} .

The surface mass density follows from the NFW density profile, namely

$$\Sigma(\theta/\theta_s) = \frac{2\rho_s r_s}{(\theta/\theta_s)^2 - 1} f(\theta/\theta_s), \quad (\text{B.48})$$

where $\theta = r/\mathcal{D}(z)$ and $\theta_s = r_s/\mathcal{D}(z)$. The function $f(x = \theta/\theta_s)$ is defined as

$$f(x) = \begin{cases} 1 - \frac{2}{\sqrt{x^2-1}} \arctan \sqrt{\frac{x-1}{x+1}}, & (x > 1) \\ 1 - \frac{2}{\sqrt{1-x^2}} \operatorname{arctanh} \sqrt{\frac{1-x}{1+x}}, & (x < 1) \\ 0, & (x = 1) \end{cases}. \quad (\text{B.49})$$

The lensing potential produced by the NFW profile is

$$\Phi_{\text{NFW}}(x) = 4\rho_s r_s \Sigma_{\text{cr}}^{-1} g(x), \quad (\text{B.50})$$

where Σ_{cr} is the critical surface density. The function $g(x)$ is defined as

$$g(x) = \frac{1}{2} \ln^2 \frac{x}{2} + \begin{cases} 2 \arctan^2 \sqrt{\frac{x-1}{x+1}}, & (x > 1) \\ -2 \operatorname{arctanh}^2 \sqrt{\frac{1-x}{1+x}}, & (x < 1) \\ 0, & (x = 1) \end{cases}. \quad (\text{B.51})$$

It follows that the deflection angle is

$$\alpha(x) = \frac{4\rho_s r_s \Sigma_{cr}^{-1}}{x} h(x), \quad (\text{B.52})$$

with

$$h(x) = \ln \frac{x}{2} + \begin{cases} \frac{2}{\sqrt{x^2-1}} \arctan \sqrt{\frac{x-1}{x+1}}, & (x > 1) \\ \frac{2}{\sqrt{1-x^2}} \operatorname{arctanh} \sqrt{\frac{1-x}{1+x}}, & (x < 1) \\ 1, & (x = 1) \end{cases}. \quad (\text{B.53})$$

The convergence can be written as

$$\kappa(x) = \frac{2\rho_s r_s \Sigma_{cr}^{-1}}{x^2 - 1} f(x), \quad (\text{B.54})$$

from which one can obtain the dimensionless mass

$$m(x) = 2 \int_0^x \kappa(x') x' dx' = 4\rho_s r_s \Sigma_{cr}^{-1} h(x). \quad (\text{B.55})$$

The lens equation of this model is usually solved through numerical simulations. At fixed halo mass the critical curves of an NFW lens are closer to its center than for SIS because of its flatter density profile. So, there the magnification is larger and decreases more slowly away from the critical curves. NFW are in this sense less efficient in image splitting, but comparably efficient in their magnification.

Appendix C

Discrete Fourier Transforms and FFTW Storing

In this Appendix we will describe the Discrete Fourier Transforms (DFTs) formalism, which is largely used within this thesis. Then we will show how we have applied FFTW, a widely used Fast Fourier transform library, to our simulation code. The main reference for this appendix is (Jeong, 2010), whose prescriptions to generate 2D and 3D fields have been followed and modified for a C++ code.

C.1 Fourier Series

Consider a real periodic function $f(\mathbf{x})$ whose period is L . Its Fourier transform $f(\mathbf{k})$ is sampled with fundamental frequency $\Delta k = 2\pi/L$, so

$$f(\mathbf{k}) = f(\Delta k \mathbf{n}_k), \quad \mathbf{n}_k = (l, m, n). \quad (\text{C.1})$$

Hence the real function can be written as

$$f(\mathbf{x}) = \frac{1}{V} \sum_{l,m,n} f(\Delta k \mathbf{n}_k) e^{i\Delta k \mathbf{n}_k \cdot \mathbf{x}}, \quad (\text{C.2})$$

where $V = L^3$. This is because $f(\mathbf{x}) = f(\mathbf{x} + \mathbf{L})$, so its Fourier transform will contain a term $e^{i\mathbf{k} \cdot (\mathbf{x} + \mathbf{L})}$. The periodicity will ensure us that $\mathbf{k} \cdot \mathbf{L} = 2\pi j$, for some integer j . Note that if we have different periods for all three directions, we will have three different fundamental frequencies with which the Fourier space function is sampled.

Now if we consider the real function to be discrete, namely $f(\mathbf{x}_r)$ with $\mathbf{x}_r = \Delta x \mathbf{n}_r$, where $H = L/N$ and $\mathbf{n}_r = (x, y, z)$, and defining $N_v = N^3$, its Fourier transform is

$$f(\mathbf{k}) = \frac{V}{N_v} \sum_{x,y,z} f(\Delta x \mathbf{n}_r) e^{-i\Delta x \mathbf{n}_r \cdot \mathbf{k}}. \quad (\text{C.3})$$

This means that $f(\mathbf{k})$ is periodic with period $2\pi/\Delta x$, namely $f(\mathbf{k}) = f(\mathbf{k} + 2\pi m/\Delta x)$. Summarising, the periodicity of a function implies the discreteness of its Fourier dual, and the discreteness of a function implies the periodicity of its Fourier dual.

C.2 Sampling and Aliasing

If we consider a sampled real function $\hat{\delta}(\mathbf{x}_r)$, we can define the Fourier pair using Eq. (C.3) and

$$\hat{\delta}(\mathbf{x}_r) = \int_{V_k} \frac{d^3k}{(2\pi)^3} \hat{\delta}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}_r}. \quad (\text{C.4})$$

We can divide the Fourier space into an infinite number of Fourier space volumes $(2\pi)^3/V$, so

$$\hat{\delta}(\mathbf{x}_r) = \delta(\mathbf{x}_r) \int \frac{d^3k}{(2\pi)^3} \delta(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}_r} = \int_{V_k} \frac{d^3k'}{(2\pi)^3} \sum_{\mathbf{n}_k} \delta\left(\mathbf{k}' - \frac{2\pi}{L}\mathbf{n}_k\right) e^{i\mathbf{k}'\cdot\mathbf{x}_r}. \quad (\text{C.5})$$

Comparing the last two equations, we see that

$$\hat{\delta}(\mathbf{k}) = \sum_{\mathbf{n}_k} \delta\left(\mathbf{k} - \frac{2\pi}{L}\mathbf{n}_k\right), \quad (\text{C.6})$$

and so the Fourier transform of discrete-sampled functions is aliased, since it will contain a sum of infinite copies of the transformed function. This fake signal will contaminate spuriously the Fourier space function. In particular a one-dimensional sampling function like

$$\Delta(x) = \sum_{n=-\infty}^{\infty} \delta_{x,n}^K \quad (\text{C.7})$$

will have Fourier transform

$$\Delta(k) = 2\pi \sum_{m=-\infty}^{\infty} \delta^D(k - 2\pi m). \quad (\text{C.8})$$

Using the sampling function in three dimension $\Delta(\mathbf{x}) = \Delta(x)\Delta(y)\Delta(k)$, we can define the sampled real function as

$$\hat{\delta}(\mathbf{x}) = \Delta\left(\frac{\mathbf{x}}{\Delta x}\right) \delta(\mathbf{x}). \quad (\text{C.9})$$

Using the similarity theorem¹, we get

$$\text{FT}\left[\Delta\left(\frac{\mathbf{x}}{\Delta x}\right)\right] = 2\pi\Delta x \sum_{m=-\infty}^{\infty} \delta^D(k\Delta x - 2\pi m) = 2\pi \sum_{m=-\infty}^{\infty} \delta^D\left(k - \frac{2\pi m}{\Delta x}\right). \quad (\text{C.10})$$

The Fourier transform of the sampled function $\hat{\delta}(\mathbf{x})$ from the convolution theorem stated in Section A.1.3,

$$\hat{\delta}(\mathbf{k}) = (2\pi)^3 \int \frac{d^3q}{(2\pi)^3} \sum_{\mathbf{n}_k} \Delta\left(q - \frac{2\pi\mathbf{n}_k}{\Delta x}\right) \delta(\mathbf{k} - \mathbf{q}) = \sum_{\mathbf{n}_k} \delta\left(\mathbf{k} - \frac{(2\pi)\mathbf{n}_k}{L}\right). \quad (\text{C.11})$$

In practice, the Fourier series is the Fourier transform of the sampled function with step Δx .

¹The similarity theorem states that the Fourier transform of $f\left(\frac{x}{a}\right)$ is $|a|f(ka)$. The three-dimensional version has a prefactor of $|a|^3$.

C.2.1 The Nyquist Sampling Theorem

Because of Eq. (C.11), we can state that aliasing can be avoided if the Nyquist frequency $k_{\text{Nyq}} = 2\pi/\Delta x/2$ is greater than the maximum component frequency. This is the Nyquist sampling theorem, and ensures us that aliasing is null if the power spectrum is truncated² to be zero for $|k| > k_{\text{Nyq}}$, because aliasing mixes the modes coming from $\pm x_r 2\pi/\Delta x$.

C.3 Discrete Fourier Transforms

If we have a real periodic function which is sampled with constant interval Δx , the sampled series of $\delta(x_r)$ forms a Fourier series of periodic function in Fourier space. So we have the Fourier pair

$$\delta(x_r) = \int_{V_k} \frac{d^3 k}{(2\pi)^3} \delta(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}_r}, \quad \delta(\mathbf{k}) = \frac{V}{N_v} \sum_{\mathbf{n}_r} \delta(x_r) e^{-i\mathbf{x}_r\cdot\mathbf{k}}, \quad (\text{C.12})$$

where $V/N_v = (\Delta x)^3$. The periodicity of $\delta(x)$ implies that $\delta(\mathbf{k})$ is non-null only if $\mathbf{k} = 2\pi\mathbf{n}_k/L = \Delta k\mathbf{n}_k$. So the Fourier space integration becomes a finite sum, namely

$$\delta(x_r) = \frac{1}{V} \sum_{\mathbf{k}_k} \delta(\mathbf{k}_k) e^{i\mathbf{k}_k\cdot\mathbf{x}_r} = \frac{1}{V} \sum_{\mathbf{n}_k} \delta(\Delta k\mathbf{n}_k) e^{i\frac{2\pi}{N}\mathbf{n}_k\cdot\mathbf{n}_r}. \quad (\text{C.13})$$

The orthogonality condition will involve Kroenecher deltas instead of continuous Dirac ones,

$$\sum_{\mathbf{n}_k} e^{-2\pi i\mathbf{n}_k\cdot\mathbf{n}_r/N} e^{2\pi i\mathbf{n}_k\cdot\mathbf{n}_r/N} = N_v \delta_{\mathbf{n}_k, \mathbf{n}_q \pm N\mathbf{n}_r}^{\text{K}}, \quad (\text{C.14})$$

so we can derive the inverse transform as

$$\delta_{\mathbf{k}_k} = \frac{V}{N_v} \sum_{\mathbf{n}_r} \delta(x_r) e^{-i\frac{2\pi}{N}\mathbf{n}_k\cdot\mathbf{n}_r}. \quad (\text{C.15})$$

So, the DFT pair is defined by Eqs. (C.13) and (C.15).

C.4 Using FFTW

DFTs are computationally slow to compute. Fast Fourier Transforms (FFTs) can optimize the sums by factorising the DFT matrix into a product of sparse factors, through the Cooley-Turkey (CT) algorithm. We will not treat this algorithm in detail, but interested readers can consult a plenty of books describing it. One of the widely used FFT libraries is the publically available FFTW³, the

²For cosmological density power spectrum this will not be exactly true, since the galaxy power spectrum contains a constant term which dominates on small scales, and aliasing can be computed analitically (Jeong, 2010).

³<http://www.fftw.org>

Fastest Fourier Transform in the West, and computes transforms of real and complex-valued arrays of arbitrary dimensions in $n \log n$ time. It supports several variants of the CT algorithm, and it is able to make a plan on which the code automatically choose which algorithm is better to use for that particular machine. In our work we used FFTW++⁴, a C++ extension which simplifies technical aspects such as planning, allocation, wisdom, and alignment. This is used jointly with the high-performance Array class⁵, designed for scientific computing.

The aim of this section is to provide a useful guide to correctly arrange a sampled Fourier function in order to transform it with FFTW.

For the aliasing theorem, the Fourier transform of this latter is equal to the aliased sum of infinitely many copies of underlying function DFTs. For The Nyquist sampling theorem the aliasing can be avoided if the Nyquist frequency $k_{Nyq}/2$ is greater than the maximum component frequency. One solution could be to truncate to zero the input power spectrum, but this is not the case for cosmological density power spectrum. Another solution is to use a negative frequency space and take advantage of the periodicity in Fourier space: in fact, the FFTW routines use to compute complex DFTs taking only the $k \leq k_{Nyq}$. Moreover, if Δk is the fundamental frequency, the modes $\mathbf{k}' = \mathbf{k} + \Delta k \mathbf{n}_k$ are inseparable from the \mathbf{k} modes and this means that negative Fourier modes are indistinguishable from the positive ones for a real transformed field. In other words only the positive frequency modes can be used to perform a DFT. From a computational point of view, the n_k indices run from 0 to $N - 1$, with N the grid dimension. We will put the modes greater than the Nyquist frequency into the negative Fourier space and we take advantage of the reality or Hermitian condition Eq. (A.3), since our field must be a real field.

Because of this, half of the modes inside the periodic box are not real degrees of freedom. To perform real-to-complex (r2c) or complex-to-real (c2r) Fourier transforms, FFTW requires that the third dimension, *i.e.* the fastest varying index, in Fourier space needs to be cut roughly in half. Having half cube in Fourier space eliminates redundant results and speeds up the code. The arrays are automatically stored with a row-major order as already arranged by the Array class we are using. This latter requirement reduces the direction of the fastest varying index by about a half.

In order to generate a three-dimensional Gaussian random field in Fourier space, we need only roughly half cube with modes obeying to:

$$\Delta k(l, m, n) = \begin{cases} \Delta k(l, m, n) & \text{for } 0 \leq (l, m, n) \leq N/2 \\ \Delta k[(l, m) - N] & \text{for } N/2 + 1 \leq (l, m) \leq N - 1. \end{cases} \quad (\text{C.16})$$

The extension to a rectangular volume is trivial. We need to be careful to generate a field that satisfies the Hermitian condition. In general we should hold it for the entire box, but since we need to generate only a half-box, we have to treat properly the $k_n = 0$ and $k_n = N/2$ surfaces of the box, in which the Fourier modes are not independent from each other, since the field is both real and complex in it.

⁴<http://fftwpp.sourceforge.net/>

⁵<http://www.math.ualberta.ca/~bowman/Array>

Moreover, consider that the output of FFTW is not normalized. So, if one performs a Fourier transform and then performs the inverse transform, we need to divide the result by a factor N^n , where n is the dimensionality number of the problem.

To better understand how to store these complex values in complex arrays, let us start from the two-dimensional case (which is useful when different maps are considered and stacked together, as made in Section 5.1).

C.4.1 Two-dimensional arrays

Let us consider a two-dimensional complex array A . We will need to store only half of the total complex values, just as pictured on Figure C.1, so we will need to run the last index only up to $m/2$, the Nyquist index.

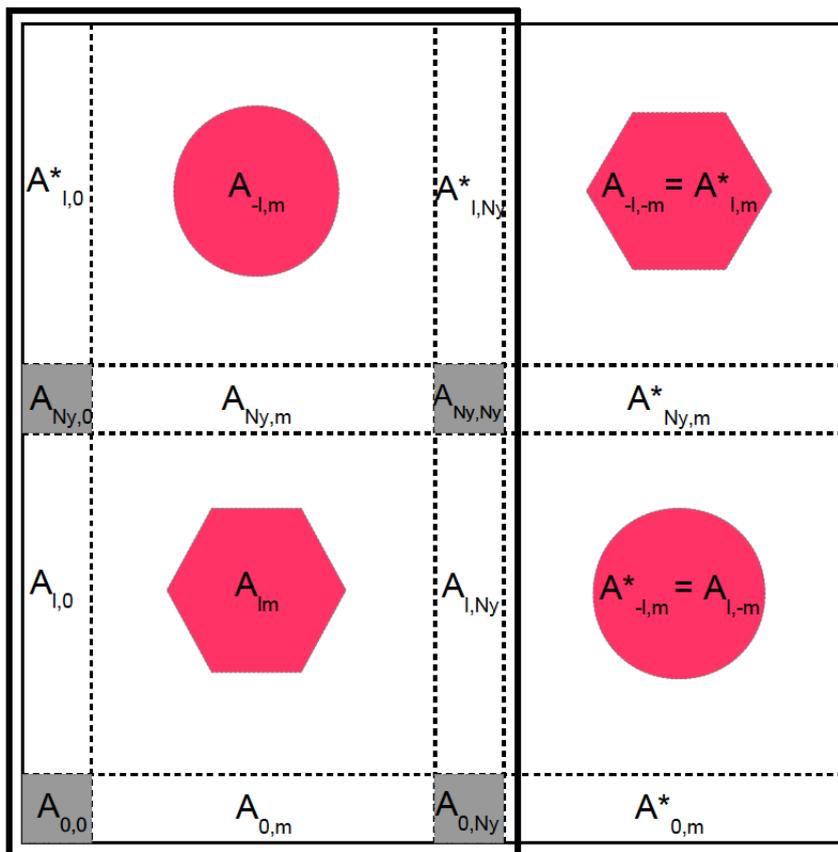


Figure C.1: Illustration showing how to store Fourier modes values into a complex array to create an Hermitian map for N even. Grey regions are pure real numbers in the complex array. The highlighted rectangle indicates the non-redundant half array to consider for FFTW storing.

In fact, because of reality condition we have

$$A[N - l, N - m] = A[-l, -m] = A^*[l, m], \tag{C.17}$$

so every grid point will have a complex conjugate in the same map, and half of the grid points are redundant. For the same reason one can show that the

$[0, 0]$ component (which is usually set to be null), and the ones involving the Nyquist frequency $[0, N/2]$, $[N/2, 0]$, $[N/2, N/2]$ have to be pure real numbers. Namely,

$$A[l, 0] = A^*[N - l, N] = A^*[N - l, 0] \quad A[l, N/2] = A^*[-l, -N/2] = A^*[N - l, N/2]. \quad (\text{C.18})$$

Of course this is analogous for the column index,

$$A[0, m] = A^*[N, N - m] = A^*[0, N - m] \quad A[N/2, m] = A^*[-N/2, -m] = A^*[N/2, N - m]. \quad (\text{C.19})$$

So, in order to create an Hermitian map, we need to:

1. prepare an array $A[0 : N - 1, 0 : N/2]$, rounded down if N is odd;
2. If N is even, assign a complex number to $A[0 : N - 1, 1 : N/2 - 1]$ and a real number to zero and Nyquist modes. Then assign complex numbers to a half row of $A[1 : N/2 - 1, 0]$ and $A[1 : N/2 - 1, N/2]$, and its complex conjugate to another half row $A[N/2 + 1 : N - 1, 0]$ and $A[N/2 + 1 : N - 1, N/2]$;
3. if N is odd, assign a complex number to $A[0 : N - 1, 1 : N/2]$, and a real number to $A[0, 0]$. Then assign complex number to a half row of $A[0, 1 : N/2]$, and its complex conjugate to another half row $A[N/2 + 1 : N - 1 : 0]$.

C.4.2 Three-dimensional arrays

The extension of what we have stated for the two-dimensional case is straight-forward. In general we always have the reality condition

$$A[N - l, N - m, N - n] = A[-l, -m, -n] = A^*[l, m, n] \quad (\text{C.20})$$

and only half cube needs to be stored, $A[0 : N - 1, 0 : N - 1, 0 : N/2]$. The zero and Nyquist mode are pure real numbers needs, $A[0, 0, 0] = 0$, $A[N/2, 0, 0]$, $A[0, N/2, 0]$, $A[0, 0, N/2]$, $A[N/2, N/2, 0]$, $A[N/2, 0, N/2]$, $A[0, N/2, N/2]$, and the full Nyquist mode $A[N/2, N/2, N/2]$.

For the same reason we have to be careful for the surfaces at $k = 0, N/2$, since they have to be defined like two independent two-dimensional Hermitian maps. So,

$$\begin{aligned} A[l, m, 0] &= A^*[N - l, N - m, N] = A^*[N - l, N - m, 0] \\ A[l, m, N/2] &= A^*[N - l, N - m, -N/2] = A^*[N - l, N - m, N/2]. \end{aligned} \quad (\text{C.21})$$

So now we can write the prescription

1. Prepare an array $A[0 : N - 1, 0 : N - 1, 0 : N/2]$, rounded down if N is odd;
2. If N is even, assign a complex number to $A[0 : N - 1, 0 : N - 1, 1 : N/2 - 1]$ and a real number to zero and Nyquist modes. Then assign complex numbers to a half row of $A[1 : N/2, 0 : N - 1, 0]$ and $A[1 : N/2, 0 : N - 1, N/2]$, and its complex conjugate to another half row $A[N/2 + 1 : N - 1, 0 : N - 1, 0]$ and $A[N/2 + 1 : N - 1, 0 : N - 1, N/2]$ respectively. Moreover assign the complex conjugate pair to $A[(0, N/2), 1 : N - 1, 0]$ and $A[(0, N/2), 0 : N - 1, N/2]$;

3. if N is odd, assign a complex number to $A[0 : N - 1, 0 : N - 1, 1 : N/2]$, and a real number to $A[0, 0, 0]$. Then assign complex number to a half row of $A[1 : N/2, 0 : N - 1, 0]$, and its complex conjugate to another half row $A[N/2 + 1 : N - 1 : 0 : N - 1, 0]$.

If for some reason the entire 3D Hermitian cube is required, consider that the Hermitian condition needs to be applied on the entire cube, and so also the faces at $(i, j) = [(0, N/2), (0, N/2)]$ need to be stored as two-dimensional Hermitian maps, and there will be 4 pairs of sub-cubes which are one the conjugate of the other.

Appendix D

21 cm Weak Lensing Reconstruction

In this appendix we will develop in detail the formalism useful to derive the formulae used in Chapter 4. This is mainly based on the Fourier space quadratic estimator for weak lensing potential developed first by Hu & Okamoto (2002) for the CMB and then extended by Zahn & Zaldarriaga (2006) to the 3D 21 cm case. The discrete estimator used in this work modifies the discrete estimator developed by Pourtsidou & Metcalf (2015) assuming that both source and lensing fields are Gaussian at EoR redshifts. With this formalism it will also be possible to extend this study for discrete sources at lower redshifts.

D.1 Line-of-Sight Discretisation in Fourier Space

The detected 21 cm intensity map can be modeled as an angular map detected at a given frequency channel. So we need to discretise the component along the line-of-sight in order to detect a small angular patch of the sky at the observation frequency ν .

Suppose we want to measure an intensity field $I(r)$. Its power spectrum can be written as $P(\mathbf{k}_\perp, k_\parallel)$, where we have split the components of the Fourier mode \mathbf{k} in a perpendicular and a parallel component respect to the line of sight. The radial coordinate is hence $r = \mathcal{D}(z)\theta + \mathcal{L}(z)r_\nu$, where $\mathcal{D}(z)$ is the comoving transverse (or angular diameter) distance and $\mathcal{L}(z)$ is the radial comoving distance corresponding to a given frequency bandwidth $\Delta\nu$ centered around the observation redshift z . This bandwidth is broken up in several channels, each with frequency resolution $\delta\nu$, and consequently we can discretise the parallel component of the Fourier mode in k_p modes, namely $k_\parallel = 2\pi k_p / \mathcal{L}(z)$. For a flat Universe the angular diameter distance is equivalent to Eq. (2.20). For small patches of the sky the flat-sky approximation holds, so we can define the Fourier space dual of the perpendicular angular component as the multipole vector defined by $\mathbf{l} = \mathbf{k}_\perp \mathcal{D}(z)$, which can be generalised into the multipole of the spherical harmonic decomposition for a full sky representation (like in the CMB case).

The Fourier transform of this field is

$$I(r) = \int \frac{d^3k}{(2\pi)^3} I(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} = \int \frac{d^2l}{(2\pi)^2} e^{i\mathbf{l}\cdot\boldsymbol{\theta}} \int \frac{dk_\parallel}{2\pi} \frac{I(\mathbf{k}_\perp, k_\parallel)}{\mathcal{D}^2(z)} e^{ik_\parallel r_\nu}. \quad (\text{D.1})$$

With $\tilde{I}(\mathbf{l}, k_{\parallel}) = I(\mathbf{k}_{\perp}, k_{\parallel})$, we can write

$$\begin{aligned} \langle \tilde{I}(\mathbf{l}, k_{\parallel}) \tilde{I}(\mathbf{l}', k'_{\parallel}) \rangle &= (2\pi)^3 \delta^{\text{D}}(\mathbf{k} - \mathbf{k}') \frac{P(\mathbf{k}_{\perp}, k_{\parallel})}{\mathcal{D}^4(z)} \\ &= (2\pi)^2 \delta^{\text{D}}(\mathbf{l} - \mathbf{l}') (2\pi) \delta^{\text{D}}(k_{\parallel} - k'_{\parallel}) \frac{P(\mathbf{k}_{\perp}, k_{\parallel})}{\mathcal{D}^2(z)} \\ &= (2\pi)^2 \delta^{\text{D}}(\mathbf{l} - \mathbf{l}') (2\pi) \delta_{k_p}^{\text{K}} \left(\frac{\mathcal{L}(z)}{2\pi} \right) \frac{P(\mathbf{k}_{\perp}, k_{\parallel})}{\mathcal{D}^2(z)}, \end{aligned} \quad (\text{D.2})$$

where we have used the discretisation of the parallel Fourier component. Hence if we define the field $\hat{I}(\mathbf{l}, k_p) = \tilde{I}(\mathbf{l}, k_{\parallel}) / \mathcal{L}(z)$, we have

$$\langle \hat{I}(\mathbf{l}, k_p) \hat{I}(\mathbf{l}', k'_p) \rangle = (2\pi)^2 \delta^{\text{D}}(\mathbf{l} - \mathbf{l}') C_{l, k_p} \delta_{k_p}^{\text{K}}, \quad (\text{D.3})$$

so that

$$I(\boldsymbol{\theta}, r_{\nu}) = \sum_{k_p} \int \frac{d^2 l}{(2\pi)^2} \hat{I}(\mathbf{l}, k_p) e^{i\mathbf{l} \cdot \boldsymbol{\theta}}, \quad (\text{D.4})$$

and the isotropic angular power spectrum for different k_p values has been defined as

$$C_{l, k_p} = (1 + \mu_k^2)^2 \frac{P \left[\sqrt{(l/\mathcal{D})^2 + (2\pi k_p / \mathcal{L})^2} \right]}{\mathcal{D}^2(z) \mathcal{L}(z)}, \quad (\text{D.5})$$

where we included the redshift space distortion contribution, as defined in Section 3.4.1 with $\mu_k = k_{\parallel}/k$ being the cosine of the angle between the wavevector and its component along the redshift direction. Note that P is the spherically averaged power spectrum.

D.2 Continuous 3D Quadratic Estimator

Considering that modes with different k_p are independent, our optimal estimator will sum these contributions without mixing them. We start from the assumption of weak lensing, performed by Taylor expanding the lensed field respect to the unlensed one:

$$\tilde{I}(\hat{\mathbf{n}}, r_{\nu}) = I(\hat{\mathbf{n}}, r_{\nu}) + \boldsymbol{\alpha}(\hat{\mathbf{n}}) \cdot \nabla_{\hat{\mathbf{n}}} I(\hat{\mathbf{n}}, r_{\nu}) + \dots, \quad (\text{D.6})$$

where $\boldsymbol{\alpha}(\hat{\mathbf{n}}) = -\nabla_{\hat{\mathbf{n}}} \Phi(\hat{\mathbf{n}})$ and the dots denote higher-order terms in the expansion. With our Fourier convention Eqs. (A.1) and (A.2), the gradients can be written as

$$\nabla_{\hat{\mathbf{n}}} \Phi(\hat{\mathbf{n}}) = i \int \frac{d^2 l}{(2\pi)^2} l \Phi(l) e^{i\mathbf{l} \cdot \hat{\mathbf{n}}} \quad \nabla_{\hat{\mathbf{n}}} I(\hat{\mathbf{n}}, r_{\nu}) = i \int \frac{d^2 l}{(2\pi)^2} l I(l, k_p) e^{i\mathbf{l} \cdot \hat{\mathbf{n}}}. \quad (\text{D.7})$$

This leads to

$$\begin{aligned} \tilde{I}(\mathbf{l}, k_p) &= I(\mathbf{l}, k_p) - \int d^2 \hat{\mathbf{n}} e^{-i\mathbf{l} \cdot \hat{\mathbf{n}}} \int \frac{d^2 l'}{(2\pi)^2} l' I(l', k_p) e^{i\mathbf{l}' \cdot \hat{\mathbf{n}}} \int \frac{d^2 l''}{(2\pi)^2} l'' \Phi(l'') e^{i\mathbf{l}'' \cdot \hat{\mathbf{n}}} + \dots \\ &= I(\mathbf{l}, k_p) - \int \frac{d^2 l'}{(2\pi)^2} l' \cdot (\mathbf{l} - \mathbf{l}') \Phi(\mathbf{l} - \mathbf{l}') I(l', k_p). \end{aligned} \quad (\text{D.8})$$

The unlensed field is isotropic so that $\langle I(\mathbf{l}, k_p) I^*(\mathbf{l}', k'_p) \rangle = (2\pi)^2 \delta^D(\mathbf{l} - \mathbf{l}') C_{l, k_p} \delta_{k_p}^K$. The averaging is made over different realisations of the source field $I(\mathbf{l}, k_p)$ and we neglected high-order terms. Lensing effect induces correlations among different l -modes that otherwise would not exist, making the observed field anisotropic for a fixed lensing potential. We can estimate the off-diagonal elements with $\mathbf{l} \neq \mathbf{l}' = \mathbf{l} - \mathbf{L}$ which will probe the lensing potential, namely

$$\begin{aligned} \langle \tilde{I}(\mathbf{l}, k_p) \tilde{I}^*(\mathbf{l}', k'_p) \rangle &= \langle I(\mathbf{l}, k_p) I^*(\mathbf{l}', k'_p) \rangle \\ &\quad - \int \frac{d^2 l'}{(2\pi)^2} \mathbf{l}' \cdot (\mathbf{l} - \mathbf{L} - \mathbf{l}') \Phi^*(\mathbf{l} - \mathbf{L} - \mathbf{l}') \langle I(\mathbf{l}, k_p) I^*(\mathbf{l}', k'_p) \rangle \\ &\quad - \int \frac{d^2 l'}{(2\pi)^2} \mathbf{l}' \cdot (\mathbf{l} - \mathbf{l}') \Phi(\mathbf{l} - \mathbf{l}') \langle I(\mathbf{l}', k_p) I^*(\mathbf{l} - \mathbf{L}, k_p) \rangle \\ &= (2\pi)^2 \delta^D(\mathbf{L}) C_{l, k_p} \delta_{k_p}^K + \left[\mathbf{l} \cdot \mathbf{L} C_{l, k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L, k_p} \right] \Phi(\mathbf{L}) \delta_{k_p}^K. \end{aligned} \quad (\text{D.9})$$

Here we have used the reality condition for the lensing potential $\Phi^*(\mathbf{L}) = \Phi(-\mathbf{L})$. This non-zero $\mathbf{L} \neq 0$ contribution justifies the construction of the following quadratic estimator

$$\hat{\Phi}(\mathbf{L}) = N_L^{\hat{\Phi}} \sum_{k_p} \int \frac{d^2 l}{(2\pi)^2} f(\mathbf{l}, \mathbf{L}, k_p) \tilde{I}(\mathbf{l}, k_p) \tilde{I}^*(\mathbf{l} - \mathbf{L}, k_p), \quad (\text{D.10})$$

in which the form of the filter $f(\mathbf{l}, \mathbf{L}, k_p)$ depends on the kind of field we are analysing and on its statistics and $N_L^{\hat{\Phi}}$ is a normalization function. In our case this has to be such that it minimizes the variance of the field under the condition that its ensemble average recovers the true lensing field, namely $\langle \hat{\Phi}(\mathbf{L}) \rangle = \Phi(\mathbf{L})$. This means

$$\sum_{k_p} \int \frac{d^2 l}{(2\pi)^2} f(\mathbf{l}, \mathbf{L}, k_p) \left[\mathbf{l} \cdot \mathbf{L} C_{l, k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L, k_p} \right] = \left[N_L^{\hat{\Phi}} \right]^{-1}. \quad (\text{D.11})$$

Now we need to compute the estimator's variance, so

$$\begin{aligned} \langle \hat{\Phi}(\mathbf{L}) \hat{\Phi}^*(\mathbf{L}) \rangle &= \sum_{k_p} \sum_{k'_p} \int \frac{d^2 l}{(2\pi)^2} \int \frac{d^2 l'}{(2\pi)^2} \left[N_L^{\hat{\Phi}} \right]^2 f^*(\mathbf{l}, \mathbf{L}, k_p) f(\mathbf{l}', \mathbf{L}, k'_p) \\ &\quad \times \langle \tilde{I}(\mathbf{l}, k_p) \tilde{I}^*(\mathbf{l} - \mathbf{L}, k_p) \tilde{I}^*(\mathbf{l}', k'_p) \tilde{I}(\mathbf{l}' - \mathbf{L}, k'_p) \rangle. \end{aligned} \quad (\text{D.12})$$

The four-point correlation function is then developed following Wick's theorem, so

$$\begin{aligned} \langle \tilde{I}(\mathbf{l}, k_p) \tilde{I}^*(\mathbf{l} - \mathbf{L}, k_p) \tilde{I}^*(\mathbf{l}', k'_p) \tilde{I}(\mathbf{l}' - \mathbf{L}, k'_p) \rangle &= (2\pi)^4 (2\pi)^2 \delta^D(k_p - k'_p) \\ &\quad \times \left[\tilde{C}_{l, k_p}^{\text{tot}} \delta^D(\mathbf{l} - \mathbf{l}') \tilde{C}_{l'-L, k_p}^{\text{tot}} \delta^D(\mathbf{l}' - \mathbf{l}) - \tilde{C}_{l, k_p}^{\text{tot}} \delta^D(\mathbf{l} + \mathbf{l}' - \mathbf{L}) \tilde{C}_{l', k_p}^{\text{tot}} \delta^D(\mathbf{l}' + \mathbf{l} - \mathbf{L}) \right], \end{aligned} \quad (\text{D.13})$$

where $\tilde{C}_{l, k_p}^{\text{tot}} = \tilde{C}_{l, k_p} + N_l$ is the sum of the measured lensed power spectrum and its noise. As pointed out by Mandel & Zaldarriaga (2006), when l will coincide

with the 21 cm brightness temperature field we can impose $\tilde{C}_{l,k_p} = C_{l,k_p}$, since the effect of lensing is small for an individual plane. With this we get

$$\begin{aligned} \langle \hat{\Phi}(\mathbf{L}) \hat{\Phi}^*(\mathbf{L}) \rangle &= (2\pi)^2 \sum_{k_p} \int \frac{d^2l}{(2\pi)^2} \delta^D(\mathbf{0}) f^2(\mathbf{l}, \mathbf{L}, k_p) [N_L^{\hat{\Phi}}]^2 \tilde{C}_{l,k_p}^{\text{tot}} \tilde{C}_{l-L,k_p}^{\text{tot}} \\ &\quad + (2\pi)^2 \sum_{k_p} \int \frac{d^2l}{(2\pi)^2} \delta^D(\mathbf{0}) f(\mathbf{l}, \mathbf{L}, k_p) f^*(\mathbf{L} - \mathbf{l}, \mathbf{L}, k_p) [N_L^{\hat{\Phi}}]^2 \tilde{C}_{l,k_p}^{\text{tot}} \tilde{C}_{l-L,k_p}^{\text{tot}}. \end{aligned} \quad (\text{D.14})$$

If $\mathbf{L} - \mathbf{l} \rightarrow \mathbf{l}$, so that $g^*(\mathbf{L} - \mathbf{l}, \mathbf{L}, k_p) = g^*(\mathbf{l}, \mathbf{L}, k_p)$ we have

$$\langle \hat{\Phi}(\mathbf{L}) \hat{\Phi}^*(\mathbf{L}) \rangle = 2(2\pi)^2 \delta^D(\mathbf{0}) [N_L^{\hat{\Phi}}]^2 \sum_{k_p} \int \frac{d^2l}{(2\pi)^2} f^2(\mathbf{l}, \mathbf{L}, k_p) \tilde{C}_{l,k_p}^{\text{tot}} \tilde{C}_{l-L,k_p}^{\text{tot}}. \quad (\text{D.15})$$

If we minimise this expression with respect to the filter $f^2(\mathbf{l}, \mathbf{L}, k_p)$, we get

$$f(\mathbf{l}, \mathbf{L}, k_p) = N_L^{\hat{\Phi}} \frac{[\mathbf{l} \cdot \mathbf{L} C_{l,k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L,k_p}]}{2\tilde{C}_{l,k_p}^{\text{tot}} \tilde{C}_{l-L,k_p}^{\text{tot}}} \delta_{k_p}^K. \quad (\text{D.16})$$

Substituting this into the normalisation condition, we have

$$[N_L^{\hat{\Phi}}]^{-1} = \sum_{k_p} \int \frac{d^2l}{(2\pi)^2} \frac{[\mathbf{l} \cdot \mathbf{L} C_{l,k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L,k_p}]^2}{2\tilde{C}_{l,k_p}^{\text{tot}} \tilde{C}_{l-L,k_p}^{\text{tot}}}. \quad (\text{D.17})$$

Using

$$2\pi \delta^D(\mathbf{0}) = \frac{2\pi}{dk_{\parallel}} \int \frac{dk_{\parallel}}{2\pi} \frac{2\pi}{dk_{\parallel}} \rightarrow \sum_{k_p} \quad (\text{D.18})$$

into the estimator variance, we get

$$\begin{aligned} \langle \hat{\Phi}(\mathbf{L}) \hat{\Phi}^*(\mathbf{L}) \rangle &= (2\pi)^2 \delta^D(\mathbf{0}) \left\{ \sum_{k_p} \int \frac{d^2l}{(2\pi)^2} \frac{[\mathbf{l} \cdot \mathbf{L} C_{l,k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L,k_p}]^2}{2\tilde{C}_{l,k_p}^{\text{tot}} \tilde{C}_{l-L,k_p}^{\text{tot}}} \right\}^{-1} \\ &= (2\pi)^2 \delta^D(\mathbf{0}) N_L^{\hat{\Phi}} \end{aligned} \quad (\text{D.19})$$

in which we have identified the normalization function with the estimator reconstruction noise which depends only on L . This results to be the inverse contribution of every independent k_p mode reconstruction noise, namely

$$N_L^{\hat{\Phi}} = \frac{1}{\sum_{k_p} [N_{L,k_p}^{\hat{\Phi}}]^{-1}}. \quad (\text{D.20})$$

So in general we can write

$$\langle \hat{\Phi}(\mathbf{L}) \hat{\Phi}^*(\mathbf{L}') \rangle = (2\pi)^2 \delta^D(\mathbf{L} - \mathbf{L}') (C_L^{\Phi\Phi} + N_L^{\hat{\Phi}}), \quad (\text{D.21})$$

where $C_L^{\Phi\Phi}$ is the potential lensing power spectrum.

D.2.1 Comparison with 2D CMB Analogous and Comments

If we identify the generic field $I(l, k_p)$ with the 21 cm brightness temperature fluctuation field (and hence identify the power spectrum C_{l, k_p} with its power spectrum Eq. (4.11), one can discuss and understand more 3D lensing reconstruction features if this estimator is compared with its 2D CMB analogous.

First of all, as already mentioned by (Metcalf & White, 2009) and (Zahn & Zaldarriaga, 2006), more small-scale modes are available for 21 cm lensing reconstruction, thanks to the absence of a CMB-like damping tail and the consequent flatness of the estimator noise on these scales. The only theoretical limit is set by the Jeans scale. On the other hand this means that the traces of baryonic oscillations are comparatively small.

Secondly, the CMB quadratic estimator is insensitive to an isotropic magnification if the power spectrum is scale-free as well as to shear contributions if the power spectrum is constant (Bucher et al., 2012). This is different for a 21 cm quadratic lensing estimator. $C_{l, k_p} \propto l^{-2}$ at small scales, the estimator will pick up shear signal from those scales coming from dark matter signal, which is traced by HI. The magnification comes from larger scales instead, where the matter power spectrum is constant and the reconstructed signal is shear-free.

Statistical detections of the lensing signal can be better than the ones provided by CMB experiments, thanks to the possibility to stack different k_p maps. And since the distance between observer and lens is a larger fraction of the whole distance to the source, one can cross-correlate 21 cm with CMB if they probe the same multipole range.

Moreover, as we have seen in Chapter 5 we succeed in reconstructing an image of the underlying lensing potential because the reconstruction noise is essentially flat at small scales. This is a consequence of the flat behavior of 21 cm power spectrum at those scales and the noise is constant up to scales where the S/N is above one. In general the level of the estimator noise will depend on the maximum multipole l_{\max} that is possible to observe.

With respect to the 2D case, we note that because of Eq. (D.17), the total noise level goes down as we use more k_p modes. But in general $P_{\Delta T}(k)$ is monotonically decreasing on all scales of interest and therefore using high k_p modes into the estimator reconstruction noise will be useless, since the noise tends to saturate as the signal quickly goes below the thermal noise level. Only the first 20 or so modes will effectively contribute to the total estimator noise.

As last remark, we point out that the effect of foregrounds can be better constrained as they are smooth in frequency, because of the possibility to detect more 21 cm maps at different frequencies. McQuinn et al. (2006) showed that generally the effect of foreground removal techniques is to make the first k_p modes unusable. The exact number of excluded modes depend on the foreground's nature and on the specific technique, other than the bandwidth. Excluding the first 3-4 modes will raise the noise level, but not in a drastic way, still preserving a large number of reconstructable modes with $S/N > 1$.

D.3 Discrete 21 cm Weak Lensing 3D Estimator

Our aim now is to introduce a discrete formalism which continuous limit gives the results of Zahn & Zaldarriaga (2006). The discrete Fourier Transform (DFT) of the intensity field $I(\mathbf{x})$ and its inverse transform are defined as

$$I_{\mathbf{k}} = \frac{\Omega_s}{N_s} \sum_{\mathbf{x}} I(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} \quad I(\mathbf{x}) = \frac{1}{\Omega_s} \sum_{\mathbf{k}} I_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}, \quad (\text{D.22})$$

where $\mathbf{x} = (\theta, r_\nu)$, $\mathbf{k} = (\mathbf{l}, k_p)$, $\Omega_s = \theta_\perp \times \theta_\perp$, and $N_s = N_\perp^2 \times N_\parallel$ for a square survey. Here the angular space resolution is defined as $\Delta\theta = \theta_\perp/N_\perp$, while the resolution element in Fourier space is $\Delta l = 2\pi/\theta_\perp$. The discrete angular coordinates are hence discretised with the integers (m, n) such that $\theta_{m,n} = \Delta\theta(m, n)$. We use the flat sky approximation.

Assuming that $I_{\mathbf{k}}$ is Gaussian we now compute the two-point correlation between discrete modes

$$\begin{aligned} \langle I_{\mathbf{k}} I_{\mathbf{k}'}^* \rangle &= \frac{\Omega_s^2}{N_s^2} \sum_{\mathbf{x}} \sum_{\mathbf{x}'} \langle I(\mathbf{x}) I(\mathbf{x}') \rangle e^{-i\mathbf{k}\cdot\mathbf{x}} e^{i\mathbf{k}'\cdot\mathbf{x}'} \\ &= \frac{\Omega_s^2}{N_s^2} \sum_{\mathbf{x}} \sum_{\mathbf{x}'} \xi_{\mathbf{x}\mathbf{x}'} e^{-i\mathbf{k}\cdot\mathbf{x}} e^{i\mathbf{k}'\cdot\mathbf{x}'} \\ &= \frac{\Omega_s^2}{N_s^2} \sum_{\mathbf{x}} \sum_{\mathbf{x}'} \sum_{\mathbf{k}''} P_{\mathbf{k}''} e^{i\mathbf{k}''\cdot(\mathbf{x}-\mathbf{x}')} e^{-i\mathbf{k}\cdot\mathbf{x}} e^{i\mathbf{k}'\cdot\mathbf{x}'} \\ &= \frac{\Omega_s^2}{N_s^2} \sum_{\mathbf{x}} \sum_{\mathbf{x}'} \sum_{\mathbf{k}''} P_{\mathbf{k}''} e^{i(\mathbf{k}''-\mathbf{k})\cdot\mathbf{x}} e^{i(\mathbf{k}'-\mathbf{k}'')\cdot\mathbf{x}'} \\ &= \Omega_s^2 \sum_{\mathbf{k}''} P_{\mathbf{k}''} \delta_{\mathbf{k}\mathbf{k}''}^{\mathbf{K}} \delta_{\mathbf{k}'\mathbf{k}''}^{\mathbf{K}} \\ &= \Omega_s^2 P_{\mathbf{k}} \delta_{\mathbf{k}\mathbf{k}'}^{\mathbf{K}}, \end{aligned} \quad (\text{D.23})$$

where the discrete power spectrum $P_{\mathbf{k}}$ is related to the continuous one by

$$P_{\mathbf{k}} = \frac{P(k)}{V_s} = \frac{P(k)}{\Omega_s \mathcal{D}^2 \mathcal{L}} = \frac{C_{l,k_p}}{\Omega_s}. \quad (\text{D.24})$$

This defines the angular power spectrum C_{l,k_p} , with $l = k_\perp \mathcal{D}$ and $k_\parallel = 2\pi k_p / \mathcal{L}$, with \mathcal{D} and \mathcal{L} the comoving distances along the perpendicular and the parallel direction respectively. So we get

$$\langle I_{\mathbf{k}} I_{\mathbf{k}'}^* \rangle = \Omega_s C_{l,k_p} \delta_{l,l'}^{\mathbf{K}} \delta_{k_p,k_p'}^{\mathbf{K}}. \quad (\text{D.25})$$

Now we will proceed analogously to what developed in the previous section. Defining the gradients Eqs. (D.7) in the discrete case and using these and Eq. (D.1) into Eq. (D.6), we get for the lensed field in Fourier space

$$\tilde{I}_{l,k_p} \simeq I_{l,k_p} + \frac{1}{\Omega_s} \sum_{l'} l' \cdot (l - l') I_{l',k_p} \Phi_{l-l'} + \dots \quad (\text{D.26})$$

So, the correlation between two modes is, up to first order,

$$\langle \tilde{I}_{l,k_p} \tilde{I}_{l-L,k'_p}^* \rangle = \Omega_s C_{l,k_p} \delta_{L,0}^K \delta_{k_p,k'_p}^K + \delta_{k_p,k'_p}^K \left[\mathbf{l} \cdot \mathbf{L} C_{l,k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L,k_p} \right] \Phi_L. \quad (\text{D.27})$$

Neglecting the $L = 0$ mode, we will use this relation to find a quadratic estimator of the form

$$\hat{\Phi}_L = \sum_{k_p} \sum_l f(\mathbf{l}, \mathbf{L}, k_p) \tilde{I}_{l,k_p} \tilde{I}_{l-L,k_p}^*. \quad (\text{D.28})$$

This has to be unbiased, *i.e.* for the condition $\langle \hat{\Phi}_L \rangle = \phi_L$ we have

$$\sum_{k_p} \sum_l f(\mathbf{l}, \mathbf{L}, k_p) \left[\mathbf{l} \cdot \mathbf{L} C_{l,k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L,k_p} \right] = 1. \quad (\text{D.29})$$

To find the form of the filter $f(\mathbf{l}, \mathbf{L}, k_p)$ we need to write down the variance of the estimator Eq. (D.28). This will lead to the four-point correlation function

$$\begin{aligned} \langle |\hat{\Phi}_L|^2 \rangle &= \sum_l \sum_{k_p} \sum_{l'} \sum_{k'_p} f(\mathbf{l}, \mathbf{L}, k_p) f^*(\mathbf{l}', \mathbf{L}, k'_p) \langle \tilde{I}_{l,k_p} \tilde{I}_{l-L,k_p}^* \tilde{I}_{l',k'_p}^* \tilde{I}_{l'-L,k'_p} \rangle \\ &= 2\Omega_s^2 \sum_{k_p} \sum_l f^2(\mathbf{l}, \mathbf{L}, k_p) C_{l,k_p}^{\text{tot}} C_{l-L,k_p}^{\text{tot}}, \end{aligned} \quad (\text{D.30})$$

in which $C_{l,k_p}^{\text{tot}} = C_{l,k_p} + N_l$ the sum of source power spectrum and thermal noise contribution. Now we can find the optimal filter using the Lagrangian multiplier technique. Given a Lagrange multiplier A_L we will minimise the quantity

$$\langle |\hat{\Phi}_L|^2 \rangle - A_L \times \left\{ \sum_{k_p} \sum_l f(\mathbf{l}, \mathbf{L}, k_p) \left[\mathbf{l} \cdot \mathbf{L} C_{l,k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L,k_p} \right] - 1 \right\} \quad (\text{D.31})$$

with respect to the filter $f(\mathbf{l}, \mathbf{L}, k_p)$. So,

$$\begin{aligned} \frac{\partial \langle |\hat{\Phi}_L|^2 \rangle - A_L \times \{ \dots \}}{\partial g(\mathbf{l}, \mathbf{L}, k_p)} &= 4\Omega_s^2 \sum_{k_p} \sum_l f(\mathbf{l}, \mathbf{L}, k_p) C_{l,k_p}^{\text{tot}} C_{l-L,k_p}^{\text{tot}} \\ &\quad - A_L \sum_{k_p} \sum_l \left[\mathbf{l} \cdot \mathbf{L} C_{l,k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L,k_p} \right] = 0 \end{aligned} \quad (\text{D.32})$$

So we get the filter

$$f(\mathbf{l}, \mathbf{L}, k_p) = \frac{A_L}{4\Omega_s^2} \frac{\left[\mathbf{l} \cdot \mathbf{L} C_{l,k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L,k_p} \right]}{C_{l,k_p}^{\text{tot}} C_{l-L,k_p}^{\text{tot}}}, \quad (\text{D.33})$$

which, put into normalization condition Eq. (D.29), gives

$$A_L = 4 \left\{ \sum_{k_p} \sum_l \frac{\left[\mathbf{l} \cdot \mathbf{L} C_{l,k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L,k_p} \right]^2}{C_{l,k_p}^{\text{tot}} C_{l-L,k_p}^{\text{tot}}} \right\}^{-1}. \quad (\text{D.34})$$

The estimator variance is hence again identified with the estimator noise, since

$$\langle |\hat{\Phi}_L|^2 \rangle = \Omega_s \left\{ \sum_{k_p} \sum_l \frac{[\mathbf{l} \cdot \mathbf{L} C_{l,k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L,k_p}]^2}{2\Omega_s C_{l,k_p}^{\text{tot}} C_{l-L,k_p}^{\text{tot}}} \right\}^{-1} = \Omega_s N_L^{\hat{\Phi}}. \quad (\text{D.35})$$

We conclude this section noting that if

$$\sum_l \rightarrow \Omega_s \int \frac{d^2 l}{(2\pi)^2} \quad \Omega_s \rightarrow (2\pi)^2 \delta^D(\mathbf{0}), \quad (\text{D.36})$$

we get the continuous results Eq. (D.19).

So for the general estimator two-point correlation we get the relation

$$\langle \hat{\Phi}(\mathbf{L}) \hat{\Phi}^*(\mathbf{L}') \rangle = \Omega_s (C_L^{\Phi\Phi} + N_L^{\hat{\Phi}}) \delta_{\mathbf{L},\mathbf{L}'}^K. \quad (\text{D.37})$$

D.4 Discrete Estimator with a Beam Function

Here we will develop in more detail the computations which lead to Eqs. (4.22) and (4.21) in Section 4.4.

The observed field point will be, including Sky and receiver noises,

$$\tilde{I}_{\mathbf{x},r} = \sum_{\mathbf{x}'} W_{\mathbf{x}\mathbf{x}'} (\tilde{I}_{\mathbf{x}',r} + n_{\mathbf{x}',r}^{\text{Sky}}) + n_{\mathbf{x},r}^{\text{Rcv}}, \quad (\text{D.38})$$

with Fourier transform

$$\tilde{I}_{l,k_p} = W_l (\tilde{I}_{l,k_p} + n_l^{\text{Sky}}) + n_l^{\text{Rcv}}. \quad (\text{D.39})$$

If we proceed just as in the previous section, we have, excluding the $\mathbf{L} = \mathbf{0}$ mode,

$$\langle \tilde{I}_{l,k_p} \tilde{I}_{l-L,k_p}^* \rangle = W_l W_{l-L}^* [\mathbf{L} \cdot \mathbf{l} C_{l,k_p} + \mathbf{L} \cdot (\mathbf{L} - \mathbf{l}) C_{|l-L,k_p}] \Phi_L. \quad (\text{D.40})$$

So we search for an estimator like

$$\hat{\phi}_L = \sum_{k_p} \sum_l f_{l,\mathbf{L},k_p} \tilde{I}_{l,k_p} \tilde{I}_{l-L,k_p}^*. \quad (\text{D.41})$$

The estimator needs to be unbiased, namely

$$\sum_{k_p} \sum_l f_{l,\mathbf{L},k_p} W_l W_{l-L}^* [\mathbf{l} \cdot \mathbf{L} C_{l,k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L,k_p}] = 1. \quad (\text{D.42})$$

The estimator variance is

$$\begin{aligned} \langle |\hat{\Phi}_L|^2 \rangle &= \sum_l \sum_{k_p} \sum_{\mathbf{l}'} \sum_{k_p'} f(\mathbf{l}, \mathbf{L}, k_p) f^*(\mathbf{l}', \mathbf{L}, k_p') \langle \tilde{I}_{l,k_p} \tilde{I}_{l-L,k_p}^* \tilde{I}_{\mathbf{l}',k_p'}^* \tilde{I}_{\mathbf{l}'-L,k_p'} \rangle \\ &= 2\Omega_s^2 \sum_{k_p} \sum_l f^2(\mathbf{l}, \mathbf{L}, k_p) C_{l,k_p}^T C_{l-L,k_p}^T, \end{aligned} \quad (\text{D.43})$$

with $C_{l,k_p}^T = |W_l|^2 (C_{l,k_p} + N_{l,k_p}^{\text{Sky}}) + N_{l,k_p}^{\text{Rcv}}$. In order to find the optimal filter we have to minimize the expression

$$\langle |\hat{\Phi}_L|^2 \rangle - \mathcal{A}_L \times \left\{ \sum_{k_p} \sum_l f_{l,L,k_p} W_l W_{l-L}^* \left[\mathbf{l} \cdot \mathbf{L} C_{l,k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L,k_p} \right] - 1 \right\}, \quad (\text{D.44})$$

using the Lagrangian multiplier \mathcal{A}_L . We find

$$f_{l,L,k_p} = \frac{\mathcal{A}_L}{4\Omega_s^2} \frac{W_l W_{l-L}^* \left[\mathbf{l} \cdot \mathbf{L} C_{l,k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L,k_p} \right]}{C_{l,k_p}^T C_{l-L,k_p}^T}. \quad (\text{D.45})$$

Putting this into Eq. (D.42), we have

$$\mathcal{A}_L = 4 \left\{ \sum_{k_p} \sum_l \frac{|W_l|^2 |W_{l-L}|^2 \left[\mathbf{l} \cdot \mathbf{L} C_{l,k_p} - \mathbf{L} \cdot (\mathbf{l} - \mathbf{L}) C_{l-L,k_p} \right]^2}{C_{l,k_p}^T C_{l-L,k_p}^T} \right\}^{-1}. \quad (\text{D.46})$$

So in the end we get the estimator Eq. (4.21)

$$\hat{\phi}_L = \frac{\mathcal{N}_L^{\hat{\phi}}}{2\Omega_s} \sum_{l,k_p} \left\{ \frac{W_l W_{l-L}^* \left[\mathbf{L} \cdot \mathbf{l} C_{l,k_p} + \mathbf{L} \cdot (\mathbf{L} - \mathbf{l}) C_{l-L,k_p} \right]}{C_{l,k_p}^T C_{l-L,k_p}^T} \right\} \tilde{\mathcal{I}}_{l,k_p} \tilde{\mathcal{I}}_{l-L,k_p}^* \quad (\text{D.47})$$

The estimator noise will be

$$\mathcal{N}_L^{\hat{\phi}} = \left\{ \frac{1}{2\Omega_s} \sum_{l,k_p} \frac{|W_l|^2 |W_{l-L}|^2 \left[\mathbf{L} \cdot \mathbf{l} C_{l,k_p} + \mathbf{L} \cdot (\mathbf{L} - \mathbf{l}) C_{l-L,k_p} \right]^2}{C_{l,k_p}^T C_{l-L,k_p}^T} \right\}^{-1}, \quad (\text{D.48})$$

in order to hold the relation

$$\langle \hat{\Phi}(\mathbf{L}) \hat{\Phi}^*(\mathbf{L}') \rangle = \Omega_s \left(C_L^{\Phi\Phi} + \mathcal{N}_L^{\hat{\phi}} \right) \delta_{L,L'}^K. \quad (\text{D.49})$$

D.5 Fast Quadratic Estimator Derivation

In this section we will give an explicit derivation¹ for Eq.(4.24), found by applying to beamed 3D 21 cm case what can be found on other works already cited on this paper. Our starting point is the quadratic estimator expression, namely

$$\hat{\phi}_L = \frac{\mathcal{N}_L^{\hat{\phi}}}{2\Omega_s} \sum_{l,k_p} \left\{ \frac{W_l W_{l-L}^* \left[\mathbf{L} \cdot \mathbf{l} C_{l,k_p} + \mathbf{L} \cdot (\mathbf{L} - \mathbf{l}) C_{l-L,k_p} \right]}{C_{l,k_p}^T C_{l-L,k_p}^T} \right\} \tilde{\mathcal{I}}_{l,k_p} \tilde{\mathcal{I}}_{l-L,k_p}^* \quad (\text{D.50})$$

¹Interested readers are advised to consult (Anderes, 2013) if they want to see how this procedure can be demonstrated for generic estimators in 2D case, like polarization-polarization or polarization-temperature estimators.

with $C_{l,k_p}^T = |W_l|^2 (C_{l,k_p} + N_{l,k_p}^{Sky}) + N_{l,k_p}^{Rcv}$ and the estimator reconstruction noise $N_L^{\hat{\phi}}$ given in Eq.(4.22). In order to recover a beamed expression for the faster estimator of Section 4.3.1, we can rewrite Eq.(D.50) in a smarter way:

$$\hat{\phi}_L = -\frac{N_L^{\hat{\phi}}}{2\Omega_s} (iL) \cdot \sum_{k_p} \sum_l \left[\frac{ilW_l C_{l,k_p} \tilde{\mathcal{F}}_{l,k_p}}{C_{l,k_p}^T} \frac{W_{l-L}^* \tilde{\mathcal{F}}_{l-L,k_p}^*}{C_{l-L,k_p}^T} - \frac{i(l-L) W_{l-L}^* C_{l-L,k_p} \tilde{\mathcal{F}}_{l-L,k_p}^*}{C_{l-L,k_p}^T} \frac{W_l \tilde{\mathcal{F}}_{l,k_p}}{C_{l,k_p}^T} \right]. \quad (D.51)$$

Now, if we define the following small-scale filtered fields

$$\mathcal{F}_{l,k_p} = \frac{W_l \tilde{\mathcal{F}}_{l,k_p}}{C_{l,k_p}^T}, \quad \mathcal{G}_{l,k_p} = \frac{W_l C_{l,k_p} \tilde{\mathcal{F}}_{l,k_p}}{C_{l,k_p}^T}, \quad (D.52)$$

our estimator will be

$$\hat{\phi}_L = -\frac{N_L^{\hat{\phi}}}{2\Omega_s} (iL) \cdot \sum_{k_p} \sum_l \left\{ il\mathcal{G}_{l,k_p} \mathcal{F}_{l-L,k_p}^* + [i(l-L)\mathcal{G}_{l-L,k_p}]^* \mathcal{F}_{l,k_p} \right\}. \quad (D.53)$$

If we consider the first l -sum for a given k_p mode, we can see that it is a convolution in Fourier space and, for the convolution theorem, this is equivalent to a product of two real space functions:

$$\begin{aligned} \sum_l il\mathcal{G}_l \mathcal{F}_{l-L}^* &= \sum_l il\mathcal{G}_l \left[\sum_{\theta} e^{-i(l-L)\theta} \mathcal{F}_{\theta} \right]^* \\ &= \sum_{\theta} e^{-iL\theta} \mathcal{F}_{\theta}^* \sum_l e^{il\theta} il\mathcal{G}_l = \sum_{\theta} e^{-iL\theta} \mathcal{F}_{\theta} (\nabla_{\theta} \mathcal{G}_{\theta}). \end{aligned} \quad (D.54)$$

Analogously proceeding, one can show that the second sum in D.53 leads to the same result. So in the end, we find for every independent k_p mode

$$\hat{\phi}_L = -\frac{N_L^{\hat{\phi}}}{\Omega_s} (iL) \cdot \sum_{k_p} \left[\sum_{\theta} e^{-iL\theta} \mathcal{F}_{\theta} (\nabla_{\theta} \mathcal{G}_{\theta}) \right]_{k_p}, \quad (D.55)$$

where the subscript k_p means that every Fast Fourier Transform involving these filtered fields has to be computed for a fixed k_p contribution. The result of these FFTs will be a vectorial field \mathcal{H}_{L,k_p} in Fourier space, so we recover our final form for the beamed quadratic estimator presented on Eq.(4.24):

$$\hat{\phi}_L = -\frac{N_L^{\hat{\phi}}}{\Omega_s} (iL) \cdot \sum_{k_p} \mathcal{H}_{L,k_p}. \quad (D.56)$$

Hence, this Fourier space estimator is basically the sum of the FFTs performed over k_p modes of the multiplication between the small-scale filtered real field \mathcal{F}_{θ} and the gradient of the other filtered field \mathcal{G}_{θ} . The result may depend on the

way the derivatives are implemented in the code. Generally spectral derivative are accurate enough if Fourier space fields are Hermitian and have periodic boundaries conditions. Fourth-order finite differences methods give slightly different results. So the real space optimal estimator ϕ_θ is the divergence of the vectorial field \mathcal{H}_θ normalized by its Fourier space variance, as one can notice from the presence of the operator iL in Eq.(4.24).

Appendix E

Non-Uniform Thermal Noise Derivation

Our purpose here is to derive Eq. (5.21) and to discuss the noise power spectrum properties comparing it with the uniform result and varying the observational frequency on which the bandwidth $\Delta\nu$ is applied.

E.1 Noise Angular Power Spectrum

Our starting point is the measured flux in visibility space for a given field $F(\mathbf{x})$ in angular coordinates:

$$V(\mathbf{U}) = \frac{\partial S_\nu}{\partial T} \int d^2x F(\mathbf{x}) A(\mathbf{x}) e^{-2\pi i \mathbf{U} \cdot \mathbf{x}} \quad (\text{E.1})$$

(Zaldarriaga et al., 2004), where $A(\mathbf{x})$ is the primary beam, and $\partial S_\nu / \partial T$ converts temperatures into fluxes and is equal to $2k_B / \lambda^2$ in the Rayleigh-Jeans limit, with k_B the Boltzmann constant. We know that $\Omega_s = (2\pi)^2 / d^2 l = 1 / d^2 U$. This is because $l = 2\pi U$. The power spectrum of our field is defined by

$$\langle F(\mathbf{U}) F^*(\mathbf{U}') \rangle = \delta^D(\mathbf{U} - \mathbf{U}') C_{l=2\pi U}. \quad (\text{E.2})$$

To find the expression for our thermal noise we write down the averaged value (over various skies) of the squared observed flux in visibility space, so

$$\begin{aligned} \langle V(\mathbf{U}) V^*(\mathbf{U}') \rangle &= \left(\frac{\partial S_\nu}{\partial T} \right)^2 \int d^2 U' |A(\mathbf{U} - \mathbf{U}')|^2 C_{l=2\pi U'} \\ &\approx \left(\frac{\partial S_\nu}{\partial T} \right)^2 C_{l=2\pi U} \int d^2 U' |A(\mathbf{U} - \mathbf{U}')|^2 \end{aligned} \quad (\text{E.3})$$

Now, let us consider the noise rms per observed visibility

$$\sigma_N = \left(\frac{2k_B T_{\text{sys}}}{A_{\text{eff}}} \right) \frac{1}{\sqrt{\Delta\nu t_U}}, \quad (\text{E.4})$$

where $A_{\text{eff}} = \varepsilon \pi D_{\text{el}}^2 / 4$ is the effective collecting area for one interferometer element (station or dish), with D_{el} the diameter of the element and ε its efficiency.

$T_{\text{sys}}(\nu)$ is the system thermal temperature, $\Delta\nu$ is the considered frequency bandwidth, and t_U is the integration time for one visibility pixel. If we consider that SKA-like experiments will have more than one polarization channel, the rms is reduced by a factor $\sqrt{N_{\text{pol}}}$ (Morales, 2005), so

$$\sigma_N = \left(\frac{2k_B T_{\text{sys}}}{A_{\text{eff}}} \right) \frac{1}{\sqrt{N_{\text{pol}} \Delta\nu t_U}}. \quad (\text{E.5})$$

Its square can be compared to Eq.(E.3), so that

$$\begin{aligned} C_{l,\Delta\nu}^N &= \left(\frac{2k_B T_{\text{sys}}}{A_{\text{eff}} \partial S_\nu / \partial T} \right)^2 \frac{1}{N_{\text{pol}} \Delta\nu t_U \int d^2 U' |A(\mathbf{U} - \mathbf{U}')|^2} \\ &= \left(\frac{\lambda^2 T_{\text{sys}}}{A_{\text{eff}}} \right)^2 \frac{d^2 U}{N_{\text{pol}} \Delta\nu t_U}, \end{aligned} \quad (\text{E.6})$$

where in the last passage we have considered that $A(\mathbf{U})$ is different from zero in an area $d^2 U$ and has to be normalized to one, so $\int d^2 U' |A(\mathbf{U} - \mathbf{U}')|^2 \sim 1/d^2 U$. For a non-regular beam shape this expression would be more complicated by the way. Let us explicit the observation time per each visibility element, so

$$t_U = d^2 U t_p n(\mathbf{U}, \nu) = d^2 U n(\mathbf{U}, \nu) \frac{t_o}{N_p} = d^2 U n(\mathbf{U}, \nu) \frac{t_o N_b \Omega_s}{S_{\text{area}}}, \quad (\text{E.7})$$

where our Field of View is set by the station dimension $\Omega_s = 1/d^2 U = (\lambda/D_{\text{min}})^2$ and we have written the observational time in function of the pointing time. An SKA-like experiment can point different sky areas to increase the number of independent measurements on scales smaller than the telescope field of view. The time per pointing is decreased as $t_p = t_o/N_p$, where $N_p = S_{\text{area}}/N_b \Omega_s$ is the number of pointings across a given survey area and N_b is the number of beams. Notice that $S_{\text{area}} > N_b \Omega_s$.

If we consider $N_p = 1$, the noise power spectrum is

$$C_{l,\Delta\nu}^N = \left(\frac{\lambda^2 T_{\text{sys}}}{A_{\text{eff}}} \right)^2 \frac{1}{N_{\text{pol}} \Delta\nu t_o n(\mathbf{U}, \nu)} \quad (\text{E.8})$$

Here $n(\mathbf{U}, \nu)$ is the average number density of baselines averaged over a 24 hrs period, usually only a function of U . This assumption simplifies the calculations considerably and holds true because of rotational invariance in visibility space. As already pointed out by Villaescusa-Navarro et al. (2014), its normalization is frequency dependent. The Field of View will change with the frequency, or in other words the resolution element in visibility space is frequency dependent ($U = D/\lambda$). Moreover, this function must hold the normalization condition $\int d^2 U n(\mathbf{U}, \nu) = N_{\text{stat}}^2/2$, the total station number squared. This means that the behavior of this function will change keeping the same functional form, since the array element is fixed. Suppose we have a $n(\mathbf{U}, \nu_f)$ at a given fiducial frequency ν_f . If we consider a lower frequency, $U_{\text{min}} = D_{\text{min}}/\lambda$ (set by the single interferometer element dimension) and $U_{\text{max}} = D_{\text{max}}/\lambda$ (set by

the maximum baseline dimension) will have a lower value. In order to keep the same normalization condition, this function has to shrink and increase its values. Viceversa, for a higher frequency U_{\min} and U_{\max} will be bigger than the ones at the fiducial frequency, so $n(U, \nu > \nu_f)$ has to stretch and decrease. At the first order we can write

$$n(U, \lambda) = \frac{\lambda}{\lambda_f} n\left(U \frac{\lambda}{\lambda_f}, \lambda_f\right). \quad (\text{E.9})$$

This behavior can be better understood by looking at the general definition for the baseline array distribution function in visibility space

$$\begin{aligned} n(U, \nu) &= \frac{N_{\text{stat}}(N_{\text{stat}} - 1)}{2} \rho_{2\text{D}}(U, \nu) \\ &= \frac{N_{\text{stat}}(N_{\text{stat}} - 1)}{2} B(\nu) \int_{D_{\min}}^{D_{\max}} 2\pi D \, dD \rho_{\text{ant}}(D) \int_0^{2\pi} d\phi \rho_{\text{ant}}(D'), \end{aligned} \quad (\text{E.10})$$

where $D' = |\mathbf{D}'| = |\mathbf{D} - \lambda \mathbf{U}| = \sqrt{D^2 + \lambda^2 U^2 - 2\lambda D U \cos \phi}$ and $\rho_{2\text{D}}(U, \nu)$ is the two-dimensional normalized baseline distribution. The normalization constant $B(\nu)$ is determined by the condition

$$\int_{D_{\min}}^{D_{\max}} U \, dU \int_0^\pi d\phi \rho_{2\text{D}}(U, \nu) = \pi \int_{D_{\min}}^{D_{\max}} U \, dU \rho_{2\text{D}}(U, \nu) = 1. \quad (\text{E.11})$$

The above normalization condition can be discretised, and, if $U_u = u\Delta U$, the normalization condition reads

$$\pi \Delta U \sum_{u=1}^{n_{\text{bin}}} U_u \rho_u = 1, \quad (\text{E.12})$$

where $\Delta U = (U_{\max} - U_{\min})/n_{\text{bin}}$ and n_{bin} is the number of visibility space bins. So the knowledge of $n(U, \nu)$ is based on the actual form that is used for the two-dimensional baseline density distribution. For example, Villaescusa-Navarro et al. (2014) used a function of the form $\rho_{\text{ant}}(D) = (A/D) e^{-(D/D_{\text{core}})^2/2}$, where the core radius $D_{\text{core}} = 1$ Km is the radius in which the array is considered dense. A is the physical area of the station, $A = A_{\text{eff}}/\varepsilon$. Generally, baselines larger than 4 – 5 Km do not contribute to the total sensitivity of the instrument. On the other hand, large baselines are useful for accurate measurements of foreground sources and for removal techniques.

The various $n(U, \lambda)$ at different observational redshifts from $z = 7$ to $z = 11.5$ are computed using Eq.(E.9), and using the fiducial $n_f(U, \lambda)$ provided by J. Pritchard (via private conversation). The resulting curves are shown in Figure E.1.

As explained in Section 5.6.1, SKA-Low uses aperture arrays and, above a given critical frequency, the area of a station will go as ν^{-2} , being constant below that, when the array becomes dense. Thus

$$A(\nu) = \begin{cases} (\nu_c/\nu)^2 & \text{for } \nu \geq \nu_c \\ 1 & \text{for } \nu < \nu_c. \end{cases} \quad (\text{E.13})$$

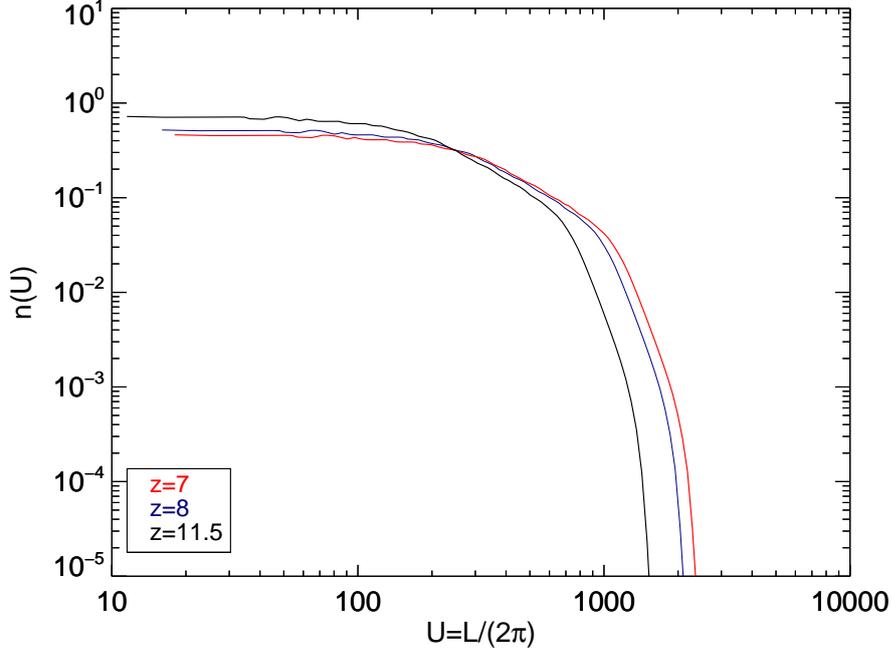


Figure E.1: $n(U, \lambda)$ for $z = 7$ (red), $z = 8$ (blue), and $z = 11.5$ (black).

Moreover, the Field of View is frequency dependent and SKA-Low can be equipped with PAFs (Phased Array Feeds), allowing the feeds to be packed to get a large number of beams, and causing an overlapping of beams over a certain critical frequency. This is also needed to keep noise uniformity across sky maps. So, at any frequency,

$$\Omega_s(\nu) = \Omega_s(\nu_c) \left(\frac{\nu_c}{\nu} \right)^2. \quad (\text{E.14})$$

If we take into account all of this, we can write our thermal noise power spectrum as

$$C_{l,\Delta\nu}^N = \left[\frac{\lambda^2}{A_{\text{eff}}(\nu_c)} \left(\frac{\nu}{\nu_c} \right)^2 \right]^2 \frac{T_{\text{sys}}^2}{N_{\text{pol}} B t_o n(U, \lambda = c/\nu)} \quad (\text{E.15})$$

In order to take into account the recent de-scoping of SKA (Bull, 2015), we need to half the number of stations previously considered. Since $n(U, \nu)$ has to be proportional to the number of stations squared Eq. (E.10), we will divide the actual $n(U, \nu)$ at any frequency by a factor 4, producing a factor 4 increase in the thermal noise power spectrum.

E.2 The Uniform Limit

The noise power spectrum in uniform approximation is, as seen in Section 5.4,

$$C_{l,\Delta\nu}^N = \frac{(2\pi)^3 T_{\text{sys}}^2}{B t_o f_{\text{cov}}^2 L_{\text{max}}^2} = \frac{T_{\text{sys}}^2}{B t_o} \frac{2\pi D_{\text{max}}^4}{N_{\text{stat}}^2 A_{\text{eff}}^2 D_{\text{max}}^2} \frac{\lambda^2}{D_{\text{max}}^2}, \quad (\text{E.16})$$

where in the last passage we have applied the definition of covering fraction as $f_{\text{cov}} = A_{\text{coll}}/D_{\text{max}}^2 = N_{\text{stat}} A_{\text{eff}}/D_{\text{max}}^2$ and maximum multipole $L_{\text{max}} = 2\pi D_{\text{max}}/\lambda$. We

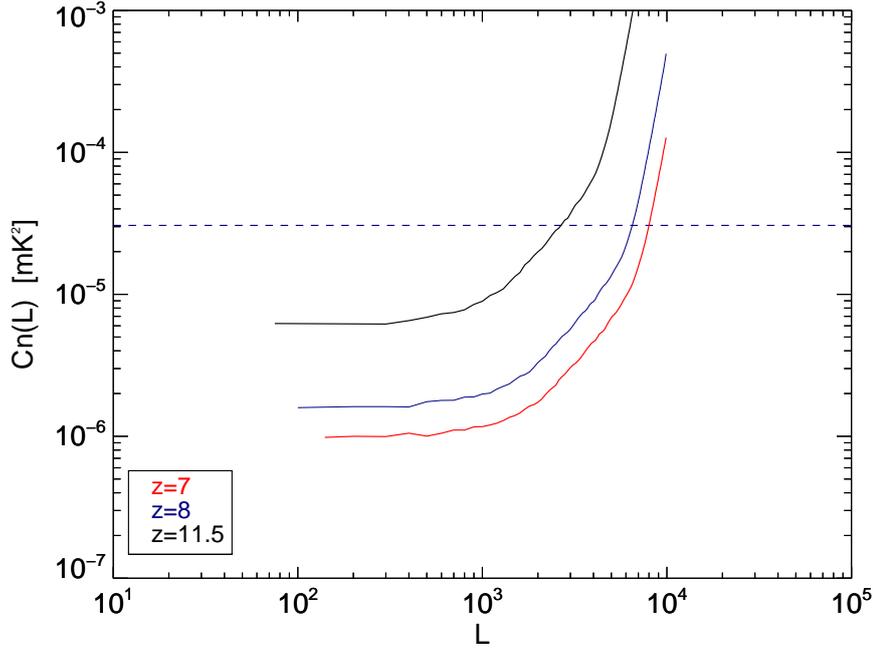


Figure E.2: The thermal noise C_L^N for $z = 7$ (red), $z = 8$ (blue), and $z = 11.5$ (black). The dashed line correspond to the old result at $z = 8$ with descoping and polarization channel contribution.

can re-arrange the previous expression as

$$C_{l,\Delta\nu}^N = \frac{T_{\text{sys}}^2}{Bt_o} \left(\frac{\lambda^2}{A_{\text{eff}}} \right)^2 \frac{1}{n(U, \lambda)}, \quad (\text{E.17})$$

which is equivalent to Eq. E.15, if we include the aperture array frequency dependence and the polarisation channel contributions. In this case we have used the uniform approximation to define the baseline density in visibility space, namely

$$n(U, \lambda) = \frac{N_{\text{stat}} (N_{\text{stat}} - 1) \lambda^2}{2\pi (D_{\text{max}}^2 - D_{\text{min}}^2)} = \frac{N_{\text{stat}} (N_{\text{stat}} - 1)}{2\pi (U_{\text{max}}^2 - U_{\text{min}}^2)}, \quad (\text{E.18})$$

for which $N_{\text{stat}} (N_{\text{stat}} - 1) \sim N_{\text{stat}}^2$ for large N_{stat} , and usually $D_{\text{min}} \ll D_{\text{max}}$.

E.2.1 Comparing Uniform and non-Uniform Cases

We can see that the main difference with the non-uniform case is due to the uniform approximation of the baseline array density in visibility space, which, depending on the actual station distribution on the ground as seen in Eq. (E.10), can cause the noise to be substantially different over a wide range of L s. If we consider our fiducial $n(U, \nu)$, we can compute the uniform and the non-uniform noises at $z = 8$ to compare them. The result is displayed in Figure E.2, and we have included for both cases frequency dependence of effective area, polarisation channel contribution, and de-scoping (so $N_{\text{stat}} = 433$). We have used the SKA1-Low R1 specifications described in Section 5.6.2.

Here we can see that, at least in the flat part of the uniform noise spectrum, the difference between the two noise power spectra is of about a factor $\sim 16 - 18$, which is equal roughly to the difference between the uniform $n(U, \nu)$ and the non-uniform one in the flat part. Here we show also the thermal noise for other two redshifts, namely $z = 7$ and $z = 11.5$. Note that the FoV is frequency dependent, so $U_{\min} = L_{\min}/(2\pi)$ changes with redshift.

Appendix F

The Schechter Luminosity Function

In this appendix we will present the main features of the Schechter luminosity function originally introduced in Press & Schechter (1974). These expression will be useful to perform the computations and implement the simulations presented in Chapter 6.

F.1 Schechter Function Definition

The Schechter function is defined as the comoving number of galaxies per unit volume per luminosity interval, namely

$$\frac{dn(\mathcal{L}, z)}{d\mathcal{L}} d\mathcal{L} = \Phi(\mathcal{L}) d\mathcal{L} = \Phi^*(z) \left(\frac{\mathcal{L}}{\mathcal{L}^*(z)} \right)^\alpha e^{-\mathcal{L}/\mathcal{L}^*(z)} d\left(\frac{\mathcal{L}}{\mathcal{L}^*(z)} \right). \quad (\text{F.1})$$

Φ^* is the normalization density of the Schechter function, α is the slope of the power-law, and \mathcal{L}^* is the characteristic luminosity of the function. Usually they can be redshift dependent, but throughout this work we will assume a conservative no-evolution scenario. The used values are taken from the HIPASS survey and specified in Section 6.1.

Since for the 21 cm emission the luminosity is proportional to the HI density, the luminosity moments are equivalent to HI mass moments whose distribution function is defined similarly to the luminosity one:

$$\frac{dn(M)}{dM} dM = \Phi(M) dM = \Phi^* \left(\frac{M}{M^*} \right)^\alpha e^{-M/M^*} d\left(\frac{M}{M^*} \right). \quad (\text{F.2})$$

Sometimes it is useful to write the Schechter function in absolute magnitude \mathcal{M} . Knowing that

$$\frac{\mathcal{L}}{\mathcal{L}_\odot} = 10^{0.4(\mathcal{M}_\odot - \mathcal{M})}, \quad (\text{F.3})$$

one obtains

$$\Phi(\mathcal{M}) d\mathcal{M} = (0.4 \ln 10) \Phi^* 10^{0.4(\alpha+1)(\mathcal{M}^* - \mathcal{M})} e^{-10^{0.4(\mathcal{M}^* - \mathcal{M})}} d\mathcal{M}. \quad (\text{F.4})$$

F.2 Schechter Function Moments

The integral of the Schechter luminosity function, *i.e.* the zero-th order moment, is the average number of galaxies over the explored volume by the survey

$$\bar{\eta} = \frac{N_{\text{gal}}}{V_{\text{tot}}} = \int \Phi(\mathcal{L}) d\mathcal{L} = \Phi^* \Gamma(\alpha + 1), \quad (\text{F.5})$$

where $\Gamma(\alpha+1)$ is the gamma function, and this integral diverges as $x = \mathcal{L}/\mathcal{L}^* \rightarrow 0$. This is not a serious problem since in reality we will never observe the entire range of available luminosities. the integral will be truncated at $x_{\text{min}} = 0.01$ and $x_{\text{max}} = 10$.

Using the HIPASS values specified in Section 6.1, we get $N_{\text{gal}} = 776890$ for a volume $V = \Omega_s \mathcal{D}^2(z) \mathcal{L}(z) = 1.40436 \times 10^7 \text{ Mpc}^3$ computed at $z = 2.5$. $\Omega_s = 25 \text{ deg}^2$ is the angular area of the survey, while $\mathcal{D}(z)$ is the comoving angular distance. The radial length $\mathcal{L}(z)$ has been computed with a bandwidth of $B = 5 \text{ MHz}$.

In general the n -th moments of the Schechter function are defined as

$$\langle \mathcal{L}^n \rangle = \frac{\Phi^*}{\bar{\eta}} (\mathcal{L}^*)^n \int \left(\frac{\mathcal{L}}{\mathcal{L}^*} \right)^{\alpha+n} e^{-\mathcal{L}/\mathcal{L}^*} d\left(\frac{\mathcal{L}}{\mathcal{L}^*} \right) = \frac{\Phi^*}{\bar{\eta}} (\mathcal{L}^*)^n \Gamma(\alpha + 1 + n). \quad (\text{F.6})$$

We can see that the moments ratio defining the shot noise power spectrum Eq. (6.9) is a ratio of Gamma functions which does not involve any knowledge on the characteristic luminosity. In fact, the continuous shot noise power spectrum is

$$P^{\text{shot}} = \frac{V_{\text{tot}} C^{\text{shot}}}{\Omega_s} = \frac{V_{\text{tot}} \langle \mathcal{L}^2 \rangle}{N_{\text{gal}} \langle \mathcal{L} \rangle^2} = \frac{1}{\Phi^*} \frac{\Gamma(\alpha + 3)}{[\Gamma(\alpha + 2)]^2}. \quad (\text{F.7})$$

Substituting our usual values, we get $P^{\text{shot}} = 95.88 \text{ Mpc}^3$.

Appendix G

High-Order Poissonian Statistics Computations

In this Appendix we will follow the computations performed by Pourtsidou & Metcalf (2015) and show in detail the computations of the high-order correlation functions which lead to the expressions introduced in Chapter 6.

G.1 Statistics for Unclustered Sources

We are interested in finding the moments of a discrete Poisson distribution. In general for a given discrete probability distribution function $p(X)$, the n th-moment is defined as the expectation value

$$E[(X - E[X])^n] = \sum_{-\infty}^{\infty} (X - E[X])^n p(x)\Delta x, \quad (\text{G.1})$$

where Δx is the sampling step for the variable X . In practice we will make use of the moment generating function method. If X is a discrete variable randomly distributed, the derivatives of the function computed at $t = 0$

$$M_X(t) = \sum_j p_X(X_j) e^{tX_j}, \quad (\text{G.2})$$

will correspond to the expectation values of a given power z of the discrete random variable, namely

$$E[X^z] = M_X^{(z)}(0). \quad (\text{G.3})$$

So, being the mean $E[X] = M_X'(0)$, the variance is

$$E[(X - E[X])^2] = E[X^2] - E^2[X] = M_X''(0) - M_X'(0)^2. \quad (\text{G.4})$$

We will make use of the third and fourth moments as well,

$$E[(X - E[X])^3] = E[X^3] - 3E[X]E[X^2] + 3E^2[X]E[X] - E^3[X] \quad (\text{G.5})$$

$$E[(X - E[X])^4] = E[X^4] - 4E[X]E[X^3] + 6E^2[X]E^2[X^2] - 4E^3[X]E[X] + E^4[X] \quad (\text{G.6})$$

Considering a Poisson distribution,

$$\begin{aligned} M_X(t) &= \sum_{X=0}^{\infty} e^{tX} \frac{e^{-\lambda} \lambda^X}{X!} = e^{-\lambda} \sum_{X=0}^{\infty} \frac{(e^t \lambda)^X}{X!} \\ &= e^{-\lambda} e^{\lambda e^t}, \end{aligned} \quad (\text{G.7})$$

the mean of a Poisson random variable X is hence

$$M'_X(0) = \lambda e^{t+\lambda e^t - \lambda} \Big|_{t=0} = \lambda. \quad (\text{G.8})$$

If we identify our variable with the fluctuation in the number of sources in a pixel i , $\delta n_i = n_i - \bar{n}_i$, trivially $\langle \delta n_i \rangle = 0$. The second moment is

$$\langle \delta n_i^2 \rangle = \langle (n_i - \bar{n}_i)^2 \rangle = \langle n_i^2 \rangle - \bar{n}_i = \bar{n}_i, \quad (\text{G.9})$$

since $\langle n_i^2 \rangle = M''_{n_i}(0) = \bar{n}_i(\bar{n}_i + 1)$. Analogously we have for the third moment,

$$M_{n_i}^{(3)}(0) = \langle n_i^3 \rangle = \bar{n}_i(\bar{n}_i^2 + 3\bar{n}_i + 1), \quad (\text{G.10})$$

so that

$$\langle \delta n_i^3 \rangle = \langle (n_i - \bar{n}_i)^3 \rangle = \langle n_i^3 \rangle - \bar{n}_i^3 - 3\bar{n}_i^2 = \bar{n}_i. \quad (\text{G.11})$$

For the fourth moment we have

$$\langle \delta n_i^4 \rangle = \langle (n_i - \bar{n}_i)^4 \rangle = \langle n_i^4 \rangle - 6\bar{n}_i^3 - 4\bar{n}_i^2 - \bar{n}_i^4 = \bar{n}_i(1 + 3\bar{n}_i), \quad (\text{G.12})$$

since we found

$$M_{n_i}^{(4)}(0) = \langle n_i^4 \rangle = \bar{n}_i(\bar{n}_i^3 + 6\bar{n}_i^2 + 7\bar{n}_i + 1). \quad (\text{G.13})$$

G.1.1 Estimator Variance for Unclustered Sources

In this section we will compute the four-point correlation function of the brightness fluctuation field $\delta S(\mathbf{j})$ useful to get Eq. (6.13). To lighten the notation we will ignore for the moment the factor Ω_s/N_{vol} in Eq. (6.4), which will be properly added at the end. Remembering that $\mathbf{j} = (\mathbf{j}_{\perp}, j_{\parallel})$ and considering the magnification to be unity, our fourth-moment is \mathcal{I}

$$\begin{aligned} \mathcal{I} &= \langle \delta \tilde{S}(\mathbf{j}) \delta \tilde{S}^*(\mathbf{j} - \mathbf{m}) \delta \tilde{S}^*(\mathbf{j}') \delta \tilde{S}(\mathbf{j}' - \mathbf{m}) \rangle \\ &= \left\langle \sum_{\mathbf{i}} \sum_{\mathcal{L}_1} \mathcal{L}_1 \delta n_{\mathbf{i}}^{\mathcal{L}_1} e^{i2\pi \mathbf{i} \cdot \mathbf{j} / N_{\text{vol}}} \sum_{\mathbf{p}} \sum_{\mathcal{L}_2} \mathcal{L}_2 \delta n_{\mathbf{p}}^{\mathcal{L}_2} e^{-i2\pi \mathbf{p} \cdot (\mathbf{j} - \mathbf{m}) / N_{\text{vol}}} \right. \\ &\quad \left. \times \sum_{\mathbf{q}} \sum_{\mathcal{L}_3} \mathcal{L}_3 \delta n_{\mathbf{q}}^{\mathcal{L}_3} e^{-i2\pi \mathbf{q} \cdot \mathbf{j}' / N_{\text{vol}}} \sum_{\mathbf{r}} \sum_{\mathcal{L}_4} \mathcal{L}_4 \delta n_{\mathbf{r}}^{\mathcal{L}_4} e^{i2\pi \mathbf{r} \cdot (\mathbf{j}' - \mathbf{m}) / N_{\text{vol}}} \right\rangle. \end{aligned} \quad (\text{G.14})$$

Not all of these averaged terms will contribute to the final variance expression. Let us consider the few non-null terms first:

- $i = p = q = r$ (and $\mathcal{L}_1 = \mathcal{L}_2 = \mathcal{L}_3 = \mathcal{L}_4$)

$$\begin{aligned} I_0 &= \left\langle \sum_i \sum_{\mathcal{L}} \mathcal{L}^4 (\delta n_i^{\mathcal{L}})^4 e^{i2\pi i \cdot (j-j+m-j'+j'-m)/N_{\text{vol}}} \right\rangle = \left\langle \sum_i \sum_{\mathcal{L}} \mathcal{L}^4 \bar{n}_i (1 + 3\bar{n}_i) e^0 \right\rangle \\ &= N_{\text{vol}} \langle \mathcal{L}^4 \rangle \bar{\eta} \delta V (1 + 3\bar{\eta} \delta V). \end{aligned} \quad (\text{G.15})$$

- $i = q \neq p = r$

$$\begin{aligned} I_1 &= \left\langle \sum_i \sum_{\mathcal{L}_1} \mathcal{L}_1^2 (\delta n_i^{\mathcal{L}_1})^2 e^{i2\pi i \cdot (j-j')/N_{\text{vol}}} \right\rangle \left\langle \sum_p \sum_{\mathcal{L}_2} \mathcal{L}_2^2 (\delta n_p^{\mathcal{L}_2})^2 e^{i2\pi i \cdot (j'-j)/N_{\text{vol}}} \right\rangle \\ &= N_{\text{vol}} (N_{\text{vol}} - 1) \langle \mathcal{L}^2 \rangle^2 \bar{\eta}^2 (\delta V)^2 \delta_{j-j'}^{\text{K}} \delta_{j'-j}^{\text{K}}. \end{aligned} \quad (\text{G.16})$$

- $i = r \neq p = q$

$$\begin{aligned} I_2 &= \left\langle \sum_i \sum_{\mathcal{L}_1} \mathcal{L}_1^2 (\delta n_i^{\mathcal{L}_1})^2 e^{i2\pi i \cdot (j+j'-m)/N_{\text{vol}}} \right\rangle \left\langle \sum_p \sum_{\mathcal{L}_2} \mathcal{L}_2^2 (\delta n_p^{\mathcal{L}_2})^2 e^{i2\pi p \cdot (m-j-j')/N_{\text{vol}}} \right\rangle \\ &= N_{\text{vol}} (N_{\text{vol}} - 1) \langle \mathcal{L}^2 \rangle^2 \bar{\eta}^2 (\delta V)^2 \delta_{j+j'-m}^{\text{K}} \delta_{m-j-j'}^{\text{K}}. \end{aligned} \quad (\text{G.17})$$

The following contributions are instead null because they are proportional to δ_j^{K} , $\delta_{j'}^{\text{K}}$, δ_m^{K} , δ_{j-m}^{K} or $\delta_{j'-m}^{\text{K}}$:

- $i = p = q \neq r$

$$I_3 = \left\langle \sum_i \sum_{\mathcal{L}} \mathcal{L}^3 (\delta n_i^{\mathcal{L}})^3 e^{i2\pi i \cdot (m-j')/N_{\text{vol}}} \right\rangle \left\langle \sum_r \sum_{\mathcal{L}_4} \mathcal{L}_4 \delta n_r^{\mathcal{L}_4} e^{i2\pi r \cdot (j'-m)/N_{\text{vol}}} \right\rangle = 0. \quad (\text{G.18})$$

- $i = p = r \neq q$

$$I_4 = \left\langle \sum_i \sum_{\mathcal{L}} \mathcal{L}^3 (\delta n_i^{\mathcal{L}})^3 e^{i2\pi i \cdot j'/N_{\text{vol}}} \right\rangle \left\langle \sum_q \sum_{\mathcal{L}_3} \mathcal{L}_3 \delta n_q^{\mathcal{L}_3} e^{i2\pi q \cdot j'/N_{\text{vol}}} \right\rangle = 0. \quad (\text{G.19})$$

- $i = r = q \neq p$

$$I_5 = \left\langle \sum_i \sum_{\mathcal{L}} \mathcal{L}^3 (\delta n_i^{\mathcal{L}})^3 e^{i2\pi i \cdot (j-m)/N_{\text{vol}}} \right\rangle \left\langle \sum_p \sum_{\mathcal{L}_2} \mathcal{L}_2 \delta n_p^{\mathcal{L}_2} e^{-i2\pi p \cdot (j-m)/N_{\text{vol}}} \right\rangle = 0. \quad (\text{G.20})$$

- $p = q = r \neq i$

$$I_6 = \left\langle \sum_p \sum_{\mathcal{L}} \mathcal{L}^3 (\delta n_p^{\mathcal{L}})^3 e^{-i2\pi p \cdot j/N_{\text{vol}}} \right\rangle \left\langle \sum_i \sum_{\mathcal{L}_1} \mathcal{L}_1 \delta n_i^{\mathcal{L}_1} e^{i2\pi i \cdot j/N_{\text{vol}}} \right\rangle = 0. \quad (\text{G.21})$$

- $i = p \neq q = r$

$$I_7 = \left\langle \sum_i \sum_{\mathcal{L}_1} \mathcal{L}_1^2 (\delta n_i^{\mathcal{L}_1})^2 e^{i2\pi i \cdot m/N_{\text{vol}}} \right\rangle \left\langle \sum_q \sum_{\mathcal{L}_3} \mathcal{L}_3^2 (\delta n_q^{\mathcal{L}_3})^2 e^{-i2\pi q \cdot m/N_{\text{vol}}} \right\rangle. \quad (\text{G.22})$$

- $i \neq p \neq q = r$

$$\begin{aligned} I_8 &= \left\langle \sum_i \sum_{\mathcal{L}_1} \mathcal{L}_1 \delta n_i^{\mathcal{L}_1} e^{i2\pi i \cdot j / N_{\text{vol}}} \right\rangle \left\langle \sum_p \sum_{\mathcal{L}_2} \mathcal{L}_2 \delta n_p^{\mathcal{L}_2} e^{-i2\pi p \cdot (j-m) / N_{\text{vol}}} \right\rangle \\ &\quad \times \left\langle \sum_q \sum_{\mathcal{L}_3} \mathcal{L}_3^2 (\delta n_q^{\mathcal{L}_3})^2 e^{i2\pi q \cdot (-m) / N_{\text{vol}}} \right\rangle = 0. \end{aligned} \quad (\text{G.23})$$

- $i \neq p = q \neq r$

$$\begin{aligned} I_9 &= \left\langle \sum_i \sum_{\mathcal{L}_1} \mathcal{L}_1 \delta n_i^{\mathcal{L}_1} e^{i2\pi i \cdot j / N_{\text{vol}}} \right\rangle \left\langle \sum_p \sum_{\mathcal{L}_2} \mathcal{L}_2^2 (\delta n_p^{\mathcal{L}_2})^2 e^{-i2\pi p \cdot (j-m+q) / N_{\text{vol}}} \right\rangle \\ &\quad \times \left\langle \sum_r \sum_{\mathcal{L}_4} \mathcal{L}_4 \delta n_r^{\mathcal{L}_4} e^{i2\pi r \cdot (j'-m) / N_{\text{vol}}} \right\rangle = 0. \end{aligned} \quad (\text{G.24})$$

- $i = p \neq q \neq r$

$$\begin{aligned} I_{10} &= \left\langle \sum_i \sum_{\mathcal{L}_1} \mathcal{L}_1^2 (\delta n_i^{\mathcal{L}_1})^2 e^{i2\pi i \cdot m / N_{\text{vol}}} \right\rangle \left\langle \sum_q \sum_{\mathcal{L}_3} \mathcal{L}_3 \delta n_q^{\mathcal{L}_3} e^{-i2\pi q \cdot j' / N_{\text{vol}}} \right\rangle \\ &\quad \times \left\langle \sum_r \sum_{\mathcal{L}_4} \mathcal{L}_4 \delta n_r^{\mathcal{L}_4} e^{i2\pi r \cdot (j'-m) / N_{\text{vol}}} \right\rangle = 0. \end{aligned} \quad (\text{G.25})$$

- $r = p \neq i \neq q$

$$\begin{aligned} I_{11} &= \left\langle \sum_i \sum_{\mathcal{L}_1} \mathcal{L}_1 \delta n_i^{\mathcal{L}_1} e^{i2\pi i \cdot j / N_{\text{vol}}} \right\rangle \left\langle \sum_q \sum_{\mathcal{L}_3} \mathcal{L}_3 \delta n_q^{\mathcal{L}_3} e^{-i2\pi q \cdot j' / N_{\text{vol}}} \right\rangle \\ &\quad \times \left\langle \sum_p \sum_{\mathcal{L}_2} \mathcal{L}_2^2 (\delta n_p^{\mathcal{L}_2})^2 e^{i2\pi p \cdot (j'-j) / N_{\text{vol}}} \right\rangle = 0. \end{aligned} \quad (\text{G.26})$$

- $i = q \neq p \neq r$

$$\begin{aligned} I_{12} &= \left\langle \sum_i \sum_{\mathcal{L}_1} \mathcal{L}_1^2 (\delta n_i^{\mathcal{L}_1})^2 e^{i2\pi i \cdot (j-j') / N_{\text{vol}}} \right\rangle \left\langle \sum_p \sum_{\mathcal{L}_2} \mathcal{L}_2 \delta n_p^{\mathcal{L}_2} e^{-i2\pi p \cdot (j-m) / N_{\text{vol}}} \right\rangle \\ &\quad \times \left\langle \sum_r \sum_{\mathcal{L}_4} \mathcal{L}_4 \delta n_r^{\mathcal{L}_4} e^{i2\pi r \cdot (j'-m) / N_{\text{vol}}} \right\rangle = 0. \end{aligned} \quad (\text{G.27})$$

- $i = r \neq p \neq q$

$$\begin{aligned} I_{13} &= \left\langle \sum_i \sum_{\mathcal{L}_1} \mathcal{L}_1^2 (\delta n_i^{\mathcal{L}_1})^2 e^{i2\pi i \cdot (j+j'-m) / N_{\text{vol}}} \right\rangle \left\langle \sum_p \sum_{\mathcal{L}_2} \mathcal{L}_2 \delta n_p^{\mathcal{L}_2} e^{-i2\pi p \cdot (j-m) / N_{\text{vol}}} \right\rangle \\ &\quad \times \left\langle \sum_q \sum_{\mathcal{L}_3} \mathcal{L}_3 \delta n_q^{\mathcal{L}_3} e^{-i2\pi q \cdot j' / N_{\text{vol}}} \right\rangle = 0. \end{aligned} \quad (\text{G.28})$$

- $i \neq r \neq p \neq q$

$$\begin{aligned} I_{14} &= \left\langle \sum_i \sum_{\mathcal{L}_1} \mathcal{L}_1 \delta n_i^{\mathcal{L}_1} e^{i2\pi i \cdot j / N_{\text{vol}}} \right\rangle \left\langle \sum_p \sum_{\mathcal{L}_2} \mathcal{L}_2 \delta n_p^{\mathcal{L}_2} e^{-i2\pi p \cdot (j-m) / N_{\text{vol}}} \right\rangle \\ &\quad \times \left\langle \sum_q \sum_{\mathcal{L}_3} \mathcal{L}_3 \delta n_q^{\mathcal{L}_3} e^{-i2\pi q \cdot j' / N_{\text{vol}}} \right\rangle \left\langle \sum_r \sum_{\mathcal{L}_4} \mathcal{L}_4 \delta n_r^{\mathcal{L}_4} e^{i2\pi r \cdot (j'-m) / N_{\text{vol}}} \right\rangle = 0. \end{aligned} \quad (\text{G.29})$$

So the variance is

$$\mathcal{I} = \mathcal{I}_0 + \mathcal{I}_1 + \mathcal{I}_2. \quad (\text{G.30})$$

Let us restore the Ω/N_{vol} factors,

$$\begin{aligned} \mathcal{I} &= \frac{\Omega_s^4}{N_{\text{vol}}^3} \langle \mathcal{L}^4 \rangle \bar{\eta} \delta V (1 + 3\bar{\eta} \delta V) + \frac{\Omega_s^4 (N_{\text{vol}} - 1)}{N_{\text{vol}}^3} \langle \mathcal{L}^2 \rangle^2 \bar{\eta}^2 (\delta V)^2 \delta_{j-j'}^{\text{K}} \delta_{j'-j}^{\text{K}} \\ &\quad + \frac{\Omega_s^4 (N_{\text{vol}} - 1)}{N_{\text{vol}}^3} \langle \mathcal{L}^2 \rangle^2 \bar{\eta}^2 (\delta V)^2 \delta_{j+j'-m}^{\text{K}} \delta_{m-j-j'}^{\text{K}}. \end{aligned} \quad (\text{G.31})$$

Now we want to write the variance for the dimensionless brightness fluctuation field $\Delta S(j)$. Hence, we need to divide the above expression by a factor $\bar{S}^4 = (\bar{\eta} \delta V \langle \mathcal{L} \rangle)^4$,

$$\begin{aligned} \frac{\mathcal{I}}{\bar{S}^4} &= \langle \Delta \tilde{S}(j) \Delta \tilde{S}^*(j-m) \Delta \tilde{S}^*(j') \Delta \tilde{S}(j'-m) \rangle \\ &= \frac{\Omega_s^4}{N_{\text{vol}}^3} \frac{\langle \mathcal{L}^4 \rangle}{\langle \mathcal{L} \rangle^4} \frac{1}{(\bar{\eta} \delta V)^3} (1 + 3\bar{\eta} \delta V) + \frac{\Omega_s^4 (N_{\text{vol}} - 1)}{N_{\text{vol}}^3} \frac{\langle \mathcal{L}^2 \rangle^2}{\langle \mathcal{L} \rangle^4} \frac{1}{(\bar{\eta} \delta V)^2} \delta_{j-j'}^{\text{K}} \delta_{j'-j}^{\text{K}} \\ &\quad + \frac{\Omega_s^4 (N_{\text{vol}} - 1)}{N_{\text{vol}}^3} \frac{\langle \mathcal{L}^2 \rangle^2}{\langle \mathcal{L} \rangle^4} \frac{1}{(\bar{\eta} \delta V)^2} \delta_{j+j'-m}^{\text{K}} \delta_{m-j-j'}^{\text{K}}. \end{aligned} \quad (\text{G.32})$$

Since $N_{\text{gal}} = \bar{\eta} \delta V N_{\text{vol}}$, we have

$$\begin{aligned} \frac{\mathcal{I}}{\bar{S}^4} &= \frac{\Omega_s^4}{N_{\text{gal}}^3} \frac{\langle \mathcal{L}^4 \rangle}{\langle \mathcal{L} \rangle^4} \left(1 + 3 \frac{N_{\text{gal}}}{N_{\text{vol}}} \right) + \frac{\Omega_s^4 (N_{\text{vol}} - 1)}{N_{\text{vol}}} \frac{\langle \mathcal{L}^2 \rangle^2}{\langle \mathcal{L} \rangle^4} \frac{1}{N_{\text{gal}}^2} \delta_{j-j'}^{\text{K}} \delta_{j'-j}^{\text{K}} \\ &\quad + \frac{\Omega_s^4 (N_{\text{vol}} - 1)}{N_{\text{vol}}} \frac{\langle \mathcal{L}^2 \rangle^2}{\langle \mathcal{L} \rangle^4} \frac{1}{N_{\text{gal}}^2} \delta_{j+j'-m}^{\text{K}} \delta_{m-j-j'}^{\text{K}}. \end{aligned} \quad (\text{G.33})$$

And finally, with C^{shot} from the first expression in Eq. (6.13), we get

$$\begin{aligned} \langle |\hat{\mu}(m_{\perp})|^2 \rangle &= \frac{\mathcal{I}}{\bar{S}^4} \frac{1}{(\Omega_s C^{\text{shot}} N_{\text{vol}})^2} \\ &= \frac{1}{N_{\text{gal}}} \frac{\langle \mathcal{L}^4 \rangle}{\langle \mathcal{L} \rangle^4} \left(1 + 3 \frac{N_{\text{gal}}}{N_{\text{vol}}} \right) + 2 \frac{N_{\text{vol}} - 1}{N_{\text{vol}}^2}. \end{aligned} \quad (\text{G.34})$$

G.2 Statistics of Clustered Sources

Now we are going to find the fundamental statistics of our clustered brightness fluctuation field, in which there are correlation between a discrete Gaussian random field and a discrete Poisson random field.

G.2.1 Power Spectrum

The second-order statistic of a discrete clustered field will contain contributions from the auto-correlation term $i = i'$ and from the cross-correlation

term $i \neq i'$,

$$\begin{aligned}
\langle \delta S(\mathbf{x}) \delta S^*(\mathbf{x}') \rangle &= \left\langle \sum_{\mathcal{L}} \mathcal{L} \delta n_{\mathbf{x}}^{\mathcal{L}} \sum_{\mathcal{L}'} \mathcal{L}' \delta n_{\mathbf{x}'}^{\mathcal{L}'} \right\rangle \\
&= \left\langle \sum_{\mathcal{L}} \mathcal{L}^2 (\delta n_{\mathbf{x}}^{\mathcal{L}})^2 \right\rangle + \left\langle \sum_{\mathcal{L}} \sum_{\mathcal{L}'} \mathcal{L} \mathcal{L}' \delta n_{\mathbf{x}}^{\mathcal{L}} \delta n_{\mathbf{x}'}^{\mathcal{L}'} \right\rangle \\
&= \langle \mathcal{L}^2 \rangle \bar{\eta} \delta V \delta_{\mathbf{x}\mathbf{x}'}^{\mathbf{K}} + (\bar{\eta} \delta V \langle \mathcal{L} \rangle)^2 \xi_{\mathbf{x}\mathbf{x}'}. \tag{G.35}
\end{aligned}$$

So the dimensionless brightness fluctuation field correlation is

$$\langle \Delta S(\mathbf{x}) \Delta S^*(\mathbf{x}') \rangle = \frac{1}{\bar{\eta} \delta V} \frac{\langle \mathcal{L}^2 \rangle}{\langle \mathcal{L} \rangle^2} \delta_{\mathbf{x}\mathbf{x}'}^{\mathbf{K}} + \xi_{\mathbf{x}\mathbf{x}'}. \tag{G.36}$$

The Fourier Transform will give

$$\begin{aligned}
\langle \Delta S_{\mathbf{k}} \Delta S_{\mathbf{k}'}^* \rangle &= \frac{\Omega_s^2}{N_{\text{vol}}^2} \left[\sum_{\mathbf{x}} \sum_{\mathbf{x}'} \frac{1}{\bar{\eta} \delta V} \frac{\langle \mathcal{L}^2 \rangle}{\langle \mathcal{L} \rangle^2} e^{-i\mathbf{k}\cdot\mathbf{x}} e^{i\mathbf{k}'\cdot\mathbf{x}'} \delta_{\mathbf{x}\mathbf{x}'}^{\mathbf{K}} + \sum_{\mathbf{x}} \sum_{\mathbf{x}'} \xi_{\mathbf{x}\mathbf{x}'} e^{-i\mathbf{k}\cdot\mathbf{x}} e^{i\mathbf{k}'\cdot\mathbf{x}'} \right] \\
&= \frac{\Omega_s^2}{N_{\text{vol}}^2} \left[\frac{1}{\bar{\eta} \delta V} \frac{\langle \mathcal{L}^2 \rangle}{\langle \mathcal{L} \rangle^2} \sum_{\mathbf{x}} e^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} + \sum_{\mathbf{x}} \sum_{\mathbf{x}'} \xi_{\mathbf{x}\mathbf{x}'} e^{-i\mathbf{k}\cdot\mathbf{x}} e^{-i\mathbf{k}'\cdot\mathbf{x}'} \right] \\
&= \frac{\Omega_s^2}{N_{\text{vol}}^2} \left[\frac{N_{\text{vol}} \langle \mathcal{L}^2 \rangle}{\bar{\eta} \delta V \langle \mathcal{L} \rangle^2} \delta_{\mathbf{k}\mathbf{k}'}^{\mathbf{K}} + \sum_{\mathbf{x}} \sum_{\mathbf{x}'} \sum_{\mathbf{k}''} P_{\mathbf{k}''} e^{i(\mathbf{k}''-\mathbf{k})\cdot\mathbf{x}} e^{i(\mathbf{k}'-\mathbf{k}'')\cdot\mathbf{x}'} \right] \\
&= \Omega_s^2 \frac{1}{N_{\text{gal}} \langle \mathcal{L} \rangle^2} \delta_{\mathbf{k}\mathbf{k}'}^{\mathbf{K}} + \Omega_s^2 \sum_{\mathbf{k}''} P_{\mathbf{k}''} \delta_{\mathbf{k}'',\mathbf{k}}^{\mathbf{K}} \delta_{\mathbf{k}'',\mathbf{k}'}^{\mathbf{K}} \\
&= \Omega_s \left[C^{\text{shot}} + C_{l,k_p} \right] \delta_{l,l'}^{\mathbf{K}} \delta_{k_p,k'_p}^{\mathbf{K}}, \tag{G.37}
\end{aligned}$$

where in the last passage we have applied the definitions of the angular power spectra Eqs. (6.9) and (4.10) and we have translated the Fourier wavevectors in multipole coordinates.

G.2.2 Bispectrum

Now we are going to find the three-point statistics for our clustered brightness fluctuation field. Applying the definition Eq. (6.2), we can write the three-point correlation function as

$$\begin{aligned}
\langle \delta S(\mathbf{x}_1) \delta S(\mathbf{x}_2) \delta S(\mathbf{x}_3) \rangle &= \sum_{\mathbf{x}_1} \sum_{\mathbf{x}_2} \sum_{\mathbf{x}_3} \sum_{\mathcal{L}_1} \sum_{\mathcal{L}_2} \sum_{\mathcal{L}_3} \langle \mathcal{L}_1 \mathcal{L}_2 \mathcal{L}_3 \delta n_{\mathbf{x}_1}^{\mathcal{L}_1} \delta n_{\mathbf{x}_2}^{\mathcal{L}_2} \delta n_{\mathbf{x}_3}^{\mathcal{L}_3} \rangle \\
&= \sum_{\mathbf{x}_1} \sum_{\mathcal{L}_1} \mathcal{L}_1^3 \langle (\delta n_{\mathbf{x}_1}^{\mathcal{L}_1})^3 \rangle \\
&\quad + \sum_{\mathbf{x}_1} \sum_{\mathbf{x}_3} \sum_{\mathcal{L}_1} \sum_{\mathcal{L}_3} \mathcal{L}_1^2 \mathcal{L}_3 \langle \delta n_{\mathbf{x}_1}^{\mathcal{L}_1} \delta n_{\mathbf{x}_3}^{\mathcal{L}_3} \rangle \delta_{12}^{\mathbf{K}} + 2 \text{ perms} \\
&\quad + \sum_{\mathbf{x}_1} \sum_{\mathbf{x}_2} \sum_{\mathbf{x}_3} \sum_{\mathcal{L}_1} \sum_{\mathcal{L}_2} \sum_{\mathcal{L}_3} \mathcal{L}_1 \mathcal{L}_2 \mathcal{L}_3 \langle \delta n_{\mathbf{x}_1}^{\mathcal{L}_1} \delta n_{\mathbf{x}_2}^{\mathcal{L}_2} \delta n_{\mathbf{x}_3}^{\mathcal{L}_3} \rangle, \tag{G.38}
\end{aligned}$$

in which we can see the terms $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}_3$, $\mathbf{x}_1 \neq \mathbf{x}_3 = \mathbf{x}_2$ (plus the two permutations), and $\mathbf{x}_1 \neq \mathbf{x}_2 \neq \mathbf{x}_3$. Applying Eq. (G.11) for the first term, we have

$$\begin{aligned} \langle \delta S(\mathbf{x}_1) \delta S(\mathbf{x}_2) \delta S(\mathbf{x}_3) \rangle &= \bar{\eta} \delta V \langle \mathcal{L}^3 \rangle \delta_{1,2}^K \delta_{1,3}^K \\ &\quad + (\bar{\eta} \delta V)^2 \langle \mathcal{L}^2 \rangle \langle \mathcal{L} \rangle \left[\xi_{1,3} \delta_{1,2}^K + \xi_{2,3} \delta_{1,3}^K + \xi_{1,2} \delta_{2,3}^K \right] \\ &\quad + (\bar{\eta} \delta V)^3 \langle \mathcal{L} \rangle^3 \xi_{123}, \end{aligned} \quad (\text{G.39})$$

where for simplicity $\xi_{1,2} \delta_{2,3}^K = \xi(|\mathbf{x}_1 - \mathbf{x}_2|) \delta_{\mathbf{x}_2, \mathbf{x}_3}^K$. Dividing by \bar{S}^3 we obtain

$$\begin{aligned} \langle \Delta S(\mathbf{x}_1) \Delta S(\mathbf{x}_2) \Delta S(\mathbf{x}_3) \rangle &= \frac{1}{\eta^2 \delta V^2} \frac{\langle \mathcal{L}^3 \rangle}{\langle \mathcal{L} \rangle^3} \delta_{1,2}^K \delta_{1,3}^K \\ &\quad + \frac{1}{\bar{\eta} \delta V} \frac{\langle \mathcal{L}^2 \rangle}{\langle \mathcal{L} \rangle^2} \left[\xi_{1,3} \delta_{1,2}^K + \xi_{2,3} \delta_{1,3}^K + \xi_{1,2} \delta_{2,3}^K \right] \\ &\quad + \xi_{123}, \end{aligned} \quad (\text{G.40})$$

where we can notice a pure Poisson term, three terms representing the interplay between Poisson counts and clustering source, and the last term ξ_{123} , which is a pure clustering term. This latter, as seen from Appendix A.4, is the three-point correlation function, which is null for Gaussian density perturbations. So, the Fourier transform of Eq. (G.40) will not have contributions from the Bispectrum of Gaussian matter perturbations, and, splitting perpendicular and parallel modes to the l.o.s., we obtain

$$\begin{aligned} &\langle \Delta S(l_1, k_p) \Delta S(l_2, k'_p) \Delta S(l_3, k''_p) \rangle \\ &= \frac{\Omega_s^3}{N_{\text{vol}}^3} \frac{1}{\eta^2 \delta V^2} \frac{\langle \mathcal{L}^3 \rangle}{\langle \mathcal{L} \rangle^3} \sum_{\mathbf{x}_1, r_v} \sum_{\mathbf{x}_2, r'_v} \sum_{\mathbf{x}_3, r''_v} \delta_{1,2}^K \delta_{1,3}^K e^{-l_1 \cdot \mathbf{x}_1} e^{-l_2 \cdot \mathbf{x}_2} e^{-l_3 \cdot \mathbf{x}_3} e^{-ik_p r_v} e^{-ik'_p r'_v} e^{-ik''_p r''_v} \\ &\quad + \frac{\Omega_s^3}{N_{\text{vol}}^3} \frac{1}{\eta \delta V} \frac{\langle \mathcal{L}^2 \rangle}{\langle \mathcal{L} \rangle^2} \left[\sum_{\mathbf{x}_1, r_v} \sum_{\mathbf{x}_2, r'_v} \sum_{\mathbf{x}_3, r''_v} \xi_{1,3} \delta_{1,2}^K e^{-l_1 \cdot \mathbf{x}_1} e^{-l_2 \cdot \mathbf{x}_2} e^{-l_3 \cdot \mathbf{x}_3} e^{-ik_p r_v} e^{-ik'_p r'_v} e^{-ik''_p r''_v} \right. \\ &\quad \left. + (2 \text{ perms}) \right], \end{aligned} \quad (\text{G.41})$$

where $\delta_{1,2}^K$ encloses information on both perpendicular and parallel components (\mathbf{x}, r_v) , and both \mathbf{x} and \mathbf{l} are discretised coordinates. Thus,

$$\begin{aligned} &\langle \Delta S(l_1, k_p) \Delta S(l_2, k'_p) \Delta S(l_3, k''_p) \rangle \\ &= \frac{\Omega_s^3}{N_{\text{vol}}^3} \frac{1}{\eta^2 \delta V^2} \frac{\langle \mathcal{L}^3 \rangle}{\langle \mathcal{L} \rangle^3} \sum_{\mathbf{x}_1, r_v} e^{-(l_1 + l_2 + l_3) \cdot \mathbf{x}_1} e^{-i(k_p + k'_p + k''_p) r_v} \\ &\quad + \frac{\Omega_s^3}{N_{\text{vol}}^3} \frac{1}{\eta \delta V} \frac{\langle \mathcal{L}^2 \rangle}{\langle \mathcal{L} \rangle^2} \left[\sum_{\mathbf{x}_1, r_v} \sum_{\mathbf{x}_3, r''_v} \xi_{1,3} e^{-(l_1 + l_2) \cdot \mathbf{x}_1} e^{-l_3 \cdot \mathbf{x}_3} e^{-i(k_p + k'_p) r_v} e^{-ik''_p r''_v} + (2 \text{ perms}) \right], \\ &= \frac{\Omega_s^3}{N_{\text{vol}}^2} \frac{1}{(\bar{\eta} \delta V)^2} \frac{\langle \mathcal{L}^3 \rangle}{\langle \mathcal{L} \rangle^3} \delta_{l_1 + l_2, -l_3}^K \delta_{k_p + k'_p, -k''_p}^K + \frac{\Omega_s^2}{N_{\text{vol}}} \frac{1}{\bar{\eta} \delta V} \frac{\langle \mathcal{L}^2 \rangle}{\langle \mathcal{L} \rangle^2} \left[C_{l_3, k''_p} \delta_{l_1 + l_2, -l_3}^K \delta_{k_p + k'_p, -k''_p}^K \right. \\ &\quad \left. + C_{l_2, k'_p} \delta_{l_1 + l_3, -l_2}^K \delta_{k_p + k''_p, -k'_p}^K + C_{l_1, k_p} \delta_{l_2 + l_3, -l_1}^K \delta_{k_p + k''_p, -k_p}^K \right]. \end{aligned} \quad (\text{G.42})$$

If we define a pure Poisson shot bispectrum B^{shot} , we can write this expression as

$$\begin{aligned} & \langle \Delta S(l_1, k_p) \Delta S(l_2, k'_p) \Delta S(l_3, k''_p) \rangle \\ &= \Omega_s \left[B^{\text{shot}} + C^{\text{shot}} \left(C_{l_3, k''_p} \delta_{l_1+l_2, -l_3}^K \delta_{k_p+k'_p, -k''_p}^K + C_{l_2, k''_p} \delta_{l_1+l_3, -l_2}^K \delta_{k_p+k'_p, -k''_p}^K \right. \right. \\ & \left. \left. + C_{l_1, k_p} \delta_{l_2+l_3, -l_1}^K \delta_{k'_p+k''_p, -k_p}^K \right) \right], \end{aligned} \quad (\text{G.43})$$

with

$$B^{\text{shot}} = \frac{\Omega_s^2}{N_{\text{vol}}^2} \frac{1}{\eta^2 \delta V^2} \frac{\langle \mathcal{L}^3 \rangle}{\langle \mathcal{L} \rangle^3}. \quad (\text{G.44})$$

G.2.3 Trispectrum

Now we want to compute the four-point correlation function of our surface brightness fluctuation field. For simplicity we will consider a compact notation for indicating perpendicular and parallel components of the vectors involved. So we will have the contributions from $x_1 \neq x_2 \neq x_3 \neq x_4$, $x_1 = x_2 = x_3 = x_4$, $x_1 \neq x_2 \neq x_3 = x_4 + 5$ permutations, $x_4 \neq x_1 = x_2 = x_3 + 3$ permutations, and $x_1 = x_2 \neq x_3 = x_4 + 2$ permutations. Using also Eq. (G.12), we can explicitly write down these terms ordering them in luminosity moments $\langle \mathcal{L}^n \rangle$,

$$\begin{aligned} & \langle \delta S(x_1) \delta S(x_2) \delta S(x_3) \delta S(x_4) \rangle \\ &= \langle \mathcal{L}^4 \rangle \bar{\eta} \delta V (1 + 3\bar{\eta} \delta V) \delta_{x_1 x_2}^K \delta_{x_1 x_3}^K \delta_{x_1 x_4}^K \\ &+ \langle \mathcal{L}^3 \rangle \langle \mathcal{L} \rangle (\bar{\eta} \delta V)^2 \left(\xi_{x_1 x_4} \delta_{x_1 x_2}^K \delta_{x_1 x_3}^K + \xi_{x_1 x_3} \delta_{x_1 x_2}^K \delta_{x_1 x_4}^K + \xi_{x_1 x_2} \delta_{x_1 x_3}^K \delta_{x_1 x_4}^K + \xi_{x_1 x_4} \delta_{x_2 x_3}^K \delta_{x_2 x_4}^K \right) \\ &+ \langle \mathcal{L}^2 \rangle^2 (\bar{\eta} \delta V)^2 \left(\xi_{x_1 x_3} \delta_{x_1 x_2}^K \delta_{x_3 x_4}^K + \xi_{x_1 x_2} \delta_{x_1 x_3}^K \delta_{x_2 x_4}^K + \xi_{x_2 x_4} \delta_{x_1 x_4}^K \delta_{x_2 x_3}^K \right) \\ &+ \langle \mathcal{L}^2 \rangle \langle \mathcal{L} \rangle^2 (\bar{\eta} \delta V)^3 \left(\xi_{x_1 x_2} \delta_{x_3 x_4}^K + \xi_{x_1 x_3} \delta_{x_2 x_4}^K + \xi_{x_1 x_4} \delta_{x_2 x_3}^K + \xi_{x_2 x_3} \delta_{x_1 x_4}^K \right. \\ & \quad \left. + \xi_{x_2 x_4} \delta_{x_1 x_3}^K + \xi_{x_3 x_4} \delta_{x_1 x_2}^K \right) \\ &+ \langle \mathcal{L} \rangle^4 (\bar{\eta} \delta V)^4 (\xi_{x_1 x_2} \xi_{x_3 x_4} + \xi_{x_1 x_3} \xi_{x_2 x_4} + \xi_{x_2 x_3} \xi_{x_1 x_4}). \end{aligned} \quad (\text{G.45})$$

At this point we divide everything by $\bar{S} = \langle \mathcal{L} \rangle^4 (\bar{\eta} \delta V)^4$ to write the four-point correlation function of the dimensionless brightness fluctuation which is

$$\begin{aligned} & \langle \Delta S(x_1) \Delta S(x_2) \Delta S(x_3) \Delta S(x_4) \rangle \\ &= \frac{\langle \mathcal{L}^4 \rangle}{\langle \mathcal{L} \rangle^4} \frac{(1 + 3\bar{\eta} \delta V)}{(\bar{\eta} \delta V)^3} \delta_{x_1 x_2}^K \delta_{x_1 x_3}^K \delta_{x_1 x_4}^K \\ &+ \frac{\langle \mathcal{L}^3 \rangle}{\langle \mathcal{L} \rangle^3} \frac{1}{(\bar{\eta} \delta V)^2} \left(\xi_{x_1 x_4} \delta_{x_1 x_2}^K \delta_{x_1 x_3}^K + \xi_{x_1 x_3} \delta_{x_1 x_2}^K \delta_{x_1 x_4}^K + \xi_{x_1 x_2} \delta_{x_1 x_3}^K \delta_{x_1 x_4}^K + \xi_{x_1 x_4} \delta_{x_2 x_3}^K \delta_{x_2 x_4}^K \right) \\ &+ \frac{\langle \mathcal{L}^2 \rangle^2}{\langle \mathcal{L} \rangle^4} \frac{1}{(\bar{\eta} \delta V)^2} \left(\xi_{x_1 x_3} \delta_{x_1 x_2}^K \delta_{x_3 x_4}^K + \xi_{x_1 x_2} \delta_{x_1 x_3}^K \delta_{x_2 x_4}^K + \xi_{x_2 x_4} \delta_{x_1 x_4}^K \delta_{x_2 x_3}^K \right) \\ &+ \frac{\langle \mathcal{L}^2 \rangle}{\langle \mathcal{L} \rangle^2} \frac{1}{\bar{\eta} \delta V} \left(\xi_{x_1 x_2} \delta_{x_3 x_4}^K + \xi_{x_1 x_3} \delta_{x_2 x_4}^K + \xi_{x_1 x_4} \delta_{x_2 x_3}^K + \xi_{x_2 x_3} \delta_{x_1 x_4}^K \right. \\ & \quad \left. + \xi_{x_2 x_4} \delta_{x_1 x_3}^K + \xi_{x_3 x_4} \delta_{x_1 x_2}^K \right) \\ &+ (\xi_{x_1 x_2} \xi_{x_3 x_4} + \xi_{x_1 x_3} \xi_{x_2 x_4} + \xi_{x_2 x_3} \xi_{x_1 x_4}). \end{aligned} \quad (\text{G.46})$$

Note that the clustering is assumed to be Gaussian.

Now we can Fourier transform this expression, remembering that the products in the exponentials $x_i k_j$ are scalar products, namely

$$\begin{aligned}
& \langle \Delta S(k_1) \Delta S^*(k_2) \Delta S^*(k_3) \Delta S(k_4) \rangle \\
&= \frac{\Omega_s^4}{N_{\text{vol}}} \frac{\langle \mathcal{L}^4 \rangle}{\langle \mathcal{L} \rangle^4} \frac{1}{N_{\text{gal}}} \left(1 + 3 \frac{N_{\text{gal}}^3}{N_{\text{vol}}} \right) \sum_{x_1} e^{-i(k_1 - k_2 - k_3 + k_4)x_1} \\
&+ \frac{\Omega_s^4}{N_{\text{vol}}^2} \frac{\langle \mathcal{L}^3 \rangle}{\langle \mathcal{L} \rangle^3} \frac{1}{N_{\text{gal}}^2} \left\{ \sum_{x_1} \sum_{x_4} \xi_{x_1 x_4} e^{-i(k_2 + k_3 - k_1)x_1} e^{-ik_4 x_4} + \sum_{x_1} \sum_{x_3} \xi_{x_1 x_3} e^{-i(k_1 - k_2 + k_4)x_1} e^{ik_3 x_3} \right. \\
&\quad \left. \sum_{x_1} \sum_{x_2} \xi_{x_1 x_2} e^{-i(k_1 - k_3 + k_4)x_1} e^{ik_2 x_2} + \sum_{x_1} \sum_{x_4} \xi_{x_1 x_4} e^{i(k_2 + k_3 - k_4)x_4} e^{-ik_1 x_1} \right\} \\
&+ \frac{\Omega_s^4}{N_{\text{vol}}^2} \frac{\langle \mathcal{L}^2 \rangle^2}{\langle \mathcal{L} \rangle^4} \frac{1}{N_{\text{gal}}^2} \left\{ \sum_{x_1} \sum_{x_3} \xi_{x_1 x_3} e^{-i(k_1 - k_2)x_1} e^{i(k_3 - k_4)x_3} + \sum_{x_1} \sum_{x_2} \xi_{x_1 x_2} e^{-i(k_1 - k_3)x_1} e^{-i(k_4 - k_2)x_2} \right. \\
&\quad \left. + \sum_{x_2} \sum_{x_4} \xi_{x_2 x_4} e^{-i(k_1 + k_4)x_4} e^{i(k_3 + k_2)x_2} \right\} \\
&+ \frac{\Omega_s^4}{N_{\text{vol}}^3} \frac{\langle \mathcal{L}^2 \rangle}{\langle \mathcal{L} \rangle^2} \frac{1}{N_{\text{gal}}} \left\{ \sum_{x_1} \sum_{x_2} \sum_{x_3} \xi_{x_1 x_2} e^{-ik_1 x_1} e^{ik_2 x_2} e^{i(k_3 - k_4)x_3} + \sum_{x_1} \sum_{x_2} \sum_{x_3} \xi_{x_1 x_3} e^{-ik_1 x_1} e^{ik_3 x_3} e^{i(k_2 - k_4)x_2} \right. \\
&\quad + \sum_{x_1} \sum_{x_2} \sum_{x_4} \xi_{x_1 x_4} e^{-ik_1 x_1} e^{-ik_4 x_4} e^{i(k_2 + k_3)x_2} + \sum_{x_1} \sum_{x_2} \sum_{x_3} \xi_{x_2 x_3} e^{ik_2 x_2} e^{ik_3 x_3} e^{-i(k_1 + k_4)x_1} \\
&\quad + \sum_{x_1} \sum_{x_2} \sum_{x_4} \xi_{x_2 x_4} e^{ik_2 x_2} e^{-ik_4 x_4} e^{-i(k_1 - k_3)x_1} + \sum_{x_2} \sum_{x_3} \sum_{x_4} \xi_{x_3 x_4} e^{ik_3 x_3} e^{-ik_4 x_4} e^{-i(k_1 - k_2)x_2} \left. \right\} \\
&+ \frac{\Omega_s^4}{N_{\text{vol}}^4} \left\{ \sum_{x_1} \sum_{x_2} \sum_{x_3} \sum_{x_4} \xi_{x_1 x_2} \xi_{x_3 x_4} e^{-ik_1 x_1} e^{ik_2 x_2} e^{ik_3 x_3} e^{-ik_4 x_4} \right. \\
&\quad + \sum_{x_1} \sum_{x_2} \sum_{x_3} \sum_{x_4} \xi_{x_1 x_3} \xi_{x_2 x_4} e^{-ik_1 x_1} e^{ik_2 x_2} e^{ik_3 x_3} e^{-ik_4 x_4} \\
&\quad \left. + \sum_{x_1} \sum_{x_2} \sum_{x_3} \sum_{x_4} \xi_{x_1 x_4} \xi_{x_2 x_3} e^{-ik_1 x_1} e^{ik_2 x_2} e^{ik_3 x_3} e^{-ik_4 x_4} \right\}. \tag{G.47}
\end{aligned}$$

At this point we can consider that $k_1 = (l, k_p)$, $k_2 = (l - L, k_p)$, $k_3 = (l', k'_p)$, and $k_4 = (l' - L, k'_p)$. The computation is lengthy but relatively simple, since we have always to apply what we have done for the bispectrum in Section G.2.2 or for the power spectrum in Section G.2.1.

If we exclude all the unobservable $L = 0$ terms, we get

$$\begin{aligned}
& \langle \Delta S(l, k_p) \Delta S^*(l - L, k_p) \Delta S^*(l', k'_p) \Delta S(l' - L, k'_p) \rangle \\
&= \frac{\Omega_s^4 \langle \mathcal{L}^4 \rangle}{N_{\text{gal}}^3 \langle \mathcal{L} \rangle^4} \left(1 + 3 \frac{N_{\text{gal}}}{N_{\text{vol}}} \right) \\
&+ \frac{\Omega_s^3 \langle \mathcal{L}^3 \rangle (N_{\text{vol}} - 1)}{N_{\text{gal}}^2 \langle \mathcal{L} \rangle^3 N_{\text{vol}}} \left[C_{l'-L, k'_p} + C_{l', k'_p} + C_{l-L, k_p} + C_{l, k_p} \right] \\
&+ \frac{\Omega_s^3 \langle \mathcal{L}^2 \rangle^2 (N_{\text{vol}} - 1)}{N_{\text{gal}}^2 \langle \mathcal{L} \rangle^4 N_{\text{vol}}} \left[C_{L, 0} + C_{l-l', k_p-k'_p} + C_{l+l'-L, k_p+k'_p} \right] \\
&+ \frac{\Omega_s^3 \langle \mathcal{L}^2 \rangle (N_{\text{vol}} - 1)(N_{\text{vol}} - 2)}{N_{\text{gal}} \langle \mathcal{L} \rangle^2 N_{\text{vol}}^2} \left[C_{l', k'_p} \delta_{l, l'}^{\text{K}} \delta_{k_p, k'_p}^{\text{K}} \delta_{l, l'}^{\text{K}} \delta_{k_p, k'_p}^{\text{K}} + \right. \\
&\quad + C_{L-l', -k'_p} \delta_{l, L-l'}^{\text{K}} \delta_{k_p, -k'_p}^{\text{K}} \delta_{l, L-l'}^{\text{K}} \delta_{k_p, -k'_p}^{\text{K}} + C_{l', k'_p} \delta_{L-l, l'}^{\text{K}} \delta_{-k_p, k'_p}^{\text{K}} \delta_{L-l, l'}^{\text{K}} \delta_{-k_p, k'_p}^{\text{K}} \\
&\quad \left. + C_{l'-L, k'_p} \delta_{l, l'}^{\text{K}} \delta_{k_p, k'_p}^{\text{K}} \delta_{l, l'}^{\text{K}} \delta_{k_p, k'_p}^{\text{K}} \right] \\
&+ \Omega_s^2 \frac{(N_{\text{vol}} - 1)(N_{\text{vol}} - 2)(N_{\text{vol}} - 1)}{N_{\text{vol}}^3} \left[C_{l, k_p} C_{l'-L, k'_p} \delta_{l, l'}^{\text{K}} \delta_{k_p, k'_p}^{\text{K}} \delta_{l, l'}^{\text{K}} \delta_{k_p, k'_p}^{\text{K}} \right. \\
&\quad \left. + C_{l, k_p} C_{l', k'_p} \delta_{l, L-l'}^{\text{K}} \delta_{k_p, -k'_p}^{\text{K}} \delta_{l, L-l'}^{\text{K}} \delta_{k_p, -k'_p}^{\text{K}} \right]. \tag{G.48}
\end{aligned}$$

At this point the last thing to do is adding an uncorrelated Gaussian thermal noise to the clustered signal by sending $C_{l, k_p} \rightarrow C_{l, k_p} + N_l^{\text{Sky}} + N_l^{\text{Rcv}} = C_{l, k_p}^T$. Let us consider now the limit $N_{\parallel} \rightarrow \infty$, in order to simplify the constant factors, and so

$$\begin{aligned}
& \langle \Delta S(l, k_p) \Delta S^*(l - L, k_p) \Delta S^*(l', k'_p) \Delta S(l' - L, k'_p) \rangle \\
&= \frac{\Omega_s^4 \langle \mathcal{L}^4 \rangle}{N_{\text{gal}}^3 \langle \mathcal{L} \rangle^4} + \frac{\Omega_s^3 \langle \mathcal{L}^3 \rangle}{N_{\text{gal}}^2 \langle \mathcal{L} \rangle^3} \left[C_{l'-L, k'_p}^T + C_{l', k'_p}^T + C_{l-L, k_p}^T + C_{l, k_p}^T \right] \\
&+ \frac{\Omega_s^3 \langle \mathcal{L}^2 \rangle^2}{N_{\text{gal}}^2 \langle \mathcal{L} \rangle^4} \left[C_{L, 0}^T + C_{l-l', k_p-k'_p}^T + C_{l+l'-L, k_p+k'_p}^T \right] \\
&+ \frac{\Omega_s^3 \langle \mathcal{L}^2 \rangle}{N_{\text{gal}} \langle \mathcal{L} \rangle^2} \left[C_{l', k'_p}^T \delta_{l, l'}^{\text{K}} \delta_{k_p, k'_p}^{\text{K}} \delta_{l, l'}^{\text{K}} \delta_{k_p, k'_p}^{\text{K}} + C_{L-l', -k'_p}^T \delta_{l, L-l'}^{\text{K}} \delta_{k_p, -k'_p}^{\text{K}} \delta_{l, L-l'}^{\text{K}} \delta_{k_p, -k'_p}^{\text{K}} \right. \\
&\quad \left. + C_{l', k'_p}^T \delta_{L-l, l'}^{\text{K}} \delta_{-k_p, k'_p}^{\text{K}} \delta_{L-l, l'}^{\text{K}} \delta_{-k_p, k'_p}^{\text{K}} + C_{l'-L, k'_p}^T \delta_{l, l'}^{\text{K}} \delta_{k_p, k'_p}^{\text{K}} \delta_{l, l'}^{\text{K}} \delta_{k_p, k'_p}^{\text{K}} \right] \\
&+ \Omega_s^2 \left[C_{l, k_p}^T C_{l'-L, k'_p}^T \delta_{l, l'}^{\text{K}} \delta_{k_p, k'_p}^{\text{K}} \delta_{l, l'}^{\text{K}} \delta_{k_p, k'_p}^{\text{K}} + C_{l, k_p}^T C_{l', k'_p}^T \delta_{l, L-l'}^{\text{K}} \delta_{k_p, -k'_p}^{\text{K}} \delta_{l, L-l'}^{\text{K}} \delta_{k_p, -k'_p}^{\text{K}} \right]. \tag{G.49}
\end{aligned}$$

The first term is a pure Poisson trispectrum term, while the last one is a pure trispectrum Gaussian term. The others are interplays of Poisson and Gaussian terms. If we remember the definitions of B^{shot} Eq. (G.44) and C^{shot} Eq. (6.9), and if we define

$$T^{\text{shot}} = \frac{\Omega_s^3 \langle \mathcal{L}^4 \rangle}{N_{\text{gal}}^3 \langle \mathcal{L} \rangle^4}, \tag{G.50}$$

we can write down

$$\begin{aligned}
& \langle \Delta S(\mathbf{l}, k_p) \Delta S^*(\mathbf{l} - \mathbf{L}, k_p) \Delta S^*(\mathbf{l}', k'_p) \Delta S(\mathbf{l}' - \mathbf{L}, k'_p) \rangle \\
&= \Omega_s \left\{ T^{\text{shot}} + B^{\text{shot}} \left[C_{l'-L, k'_p}^T + C_{l', k'_p}^T + C_{l-L, k_p}^T + C_{l, k_p}^T \right] + \left(C^{\text{shot}} \right)^2 \left[C_{L, 0}^T + C_{l-l', k_p-k'_p}^T + C_{l+l'-L, k_p+k'_p}^T \right] \right\} \\
&+ \Omega_s^2 \left\{ C^{\text{shot}} \left[C_{l', k'_p}^T \delta_{l, l'}^K \delta_{k_p, k'_p}^K \delta_{l, l'}^K \delta_{k_p, k'_p}^K + C_{L-l', -k'_p}^T \delta_{l, L-l'}^K \delta_{k_p, -k'_p}^K \delta_{l, L-l'}^K \delta_{k_p, -k'_p}^K \right. \right. \\
&\quad \left. \left. + C_{l', k'_p}^T \delta_{L-l, l'}^K \delta_{-k_p, k'_p}^K \delta_{L-l, l'}^K \delta_{-k_p, k'_p}^K + C_{l'-L, k'_p}^T \delta_{l, l'}^K \delta_{k_p, k'_p}^K \delta_{l, l'}^K \delta_{k_p, k'_p}^K \right] \right. \\
&\quad \left. + \left[C_{l, k_p}^T C_{l'-L, k'_p}^T \delta_{l, l'}^K \delta_{k_p, k'_p}^K \delta_{l, l'}^K \delta_{k_p, k'_p}^K + C_{l, k_p}^T C_{l', k'_p}^T \delta_{l, L-l'}^K \delta_{k_p, -k'_p}^K \delta_{l, L-l'}^K \delta_{k_p, -k'_p}^K \right] \right\}.
\end{aligned} \tag{G.51}$$

This is the four-point correlation function that has been used to obtain Eq. (6.25).

Acknowledgements

I should cite too many deserving people in this section, but I am sure of forgetting someone. For this I apologize in advance. Four years of life represent a big part of my life, and I am glad to have shared with all of you, here in Bologna, and at home in Reggio Calabria, all my complaints, tears, happiness, laughs, notes, blood, sweat and dreams.

In primis I would like to thank my advisor and guide Ben Metcalf, whose scientific and human preparation is unquestionable, especially for teaching (or at least trying to teach) me how to be a researcher and for its infinite patience, considering my remarkable stubbornness and large number of complaints. He always tried to direct me on the right path (even if I almost always tended to derail away from it) by encouraging my curiosity and fighting my laziness. Thank you for always believing in me even in periods in which neither I was able to believe in myself. And there have been a lot of moments like these during this years.

I also need to thank Alkistis, who helped me an uncountable number of times, listening to my problems and trying to solve them with me, overlooking my meaningless conclusions or questions. I must admit that sometimes she has been quite tough, but, yes, maybe that was necessary. She has been one of my constant reference points in this work, given her big scientific preparation, and always providing useful comments.

Then I would like to thank all my PhD-mates who shared with me this experience every day in these four years and supported my madness: Andrea, Fabio, Emilio, Davide, Cristiana, Silvia, Iary, Ivan, Tania. Thanks for all the laughs, all the "cards", all the stabbings, and all the serious and joyful moments we have had. R11 will live forever!

Thanks to all the people I have bothered inside the department of Astronomy in Bologna, trying to get them involved in my (scientific and non scientific) problems: Lauro, Carlo, Federico, Dominik, Fabio, Mauro, Fabien, Nicolas, Margarita, and all the others' year PhD students and Post-Docs (you're too many!) in Bologna.

How not to mention the Threemons and all the people that have been connected to this incredible volley team with an enviable group of supporters, thanking especially the ones who shared with me not only the team experience but also the department's life: Fernanda, Cristina, Loredana, Margherita, Olga, Daniele, Gabriele, Cristiano, and all the others.

I would like to thank also the people involved in Joe il Fotone, for all the cheerful moments and notes they gave me: Paolo, Laura, Antonio, Francesco,

Sandro, Matteo.

Outside the department's life I would like to thank all the people who went along with me every day in these years here in Bologna, supporting me in every occasion and making my life happier. Thanks to Emanuele, my housemate, for sharing every single day in these four years, complaint, meal, nonsenses and cartoons; thanks to Diego, Paolo, Gabriele, Miki, and Alessandro who shared with me their passion for music, concerts, bass, spicy food, nonsenses (again), beers, Lagavulin, movies, series, and almost all the things I care the most. Like friendship.

Thanks to my Brothers and Sisters who live spread in all parts of Italy: Luca, Viviana, Massi, Chiara, Domenico, Gabriele. Without your constant presence in my life I would never be where I am and trying to follow my dreams. You are irreplaceable and no other words are needed. I love you. Thanks to all my other friends in Reggio and in other parts of Italy: Laura, Marco, Caterina, Paolo, Fabio, Roberta, Isa, Rita, Sergio, Nino, and all those who give me support by keeping constantly in touch and sharing their life with me, with important or silly conversations. I could not imagine my life without you all and I always have you in my heart.

I would like to immensely thank all my music bandmates: Roberto, Vincenzo, Marco, Filippo, Alessio. With the things we create together, you literally give me the reason to go on and the most incredible outburst I could never have. You are essential to keep my mental health up to an acceptable minimum level. And many thanks to Stingo Gallino and his magic board, who makes his miracles come true every day.

We are almost at the end and I still have to thank my family, my uncles, aunts and cousins, for encouraging me every day in my studies and my life. But the most special place is taken by my parents, Giusy and Paolo, the only ones who are capable of giving me unconditional support, love and presence. And it is because of them that I am here trying to follow my unstable dreams. There is really nobody like them. And this is not a trivial catchphrase. I love you.

And, also this time, last but not the least, I would like to thank Alessia, for choosing to share my life with me every day. Thanks for all the support you always give me, even if you don't realize it. If I have been able to taken on these years and all the challenges I have been through it's also, if not especially, because of your presence in my life. Thanks for taking my hand and keep walking with me. Barbabietola in Agrodolce Complessata. Tante Astrazioni.

Bibliography

- Ali S. S., Bharadwaj S., Pandey B., 2005, MNRAS, 363, 251
- Alonso D., Ferreira P. G., Santos M. G., 2014, MNRAS, 444, 3183
- Anderes E., 2013, Phys. Rev. D, 88, 083517
- Ansari R., Campagne J. E., Colom P., Le Goff J. M., Magneville C., Martin J. M., Moniez M., Rich J., Yèche C., 2012, A&A, 540, A129
- Baek S., Di Matteo P., Semelin B., Combes F., Revaz Y., 2009, A&A, 495, 389
- Baek S., Semelin B., Di Matteo P., Revaz Y., Combes F., 2010, A&A, 523, A4
- Barkana R., 2006, MNRAS, 372, 259
- Barkana R., Loeb A., 2005, ApJ, 624, L65
- Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
- Bartolo N., Matarrese S., Riotto A., 2007, ArXiv Astrophysics e-prints
- Battye R. A., Brown M. L., Browne I. W. A., Davis R. J., Dewdney P., Dickinson C., Heron G., Maffei B., Pourtsidou A., Wilkinson P. N., 2012, ArXiv e-prints
- Battye R. A., Browne I. W. A., Dickinson C., Heron G., Maffei B., Pourtsidou A., 2013a, MNRAS, 434, 1239
- Battye R. A., Browne I. W. A., Dickinson C., Heron G., Maffei B., Pourtsidou A., 2013b, MNRAS, 434, 1239
- Baumann D., 2009, ArXiv e-prints
- Bernardi G., Carretti E., Fabbri R., Sbarra C., Poppi S., Cortiglioni S., 2003, MNRAS, 344, 347
- Bernardi G., Greenhill L. J., Mitchell D. A., Ord S. M., Hazelton B. J., Gaensler B. M., de Oliveira-Costa A., Morales M. F., Udaya Shankar N., Subrahmanyan R., Wayth R. B., Lenc E., Williams C. L., et al. 2013, ApJ, 771, 105
- Bernstein G. M., 2009, ApJ, 695, 652
- Bharadwaj S., Ali S. S., 2005, MNRAS, 356, 1519

- Blandford R., Narayan R., 1985, MNRAS, 213, 591
- Boerner G., Ehlers J., 1988, A&A, 204, 1
- Bowden M., Taylor A. N., Ganga K. M., Ade P. A. R., Bock J. J., Cahill G., Carlstrom J. E., Church S. E., et al. 2004, MNRAS, 349, 321
- Bowman J. C., Roberts M., 2010, CoRR, abs/1008.1366
- Bowman J. D., Morales M. F., Hewitt J. N., 2009, ApJ, 695, 183
- Bucher M., Carvalho C. S., Moodley K., Remazeilles M., 2012, Phys. Rev. D, 85, 043016
- Bull P., 2015, ArXiv e-prints
- Bull P., Ferreira P. G., Patel P., Santos M. G., 2015, ApJ, 803, 21
- Burn B. J., 1966, MNRAS, 133, 67
- Carroll S. M., Press W. H., Turner E. L., 1992, ARA&A, 30, 499
- Carvalho C. S., Moodley K., 2010, Phys. Rev. D, 81, 123010
- Chang T.-C., Pen U.-L., Bandura K., Peterson J. B., 2010, ArXiv e-prints
- Chang T.-C., Pen U.-L., Peterson J. B., McDonald P., 2008, Physical Review Letters, 100, 091303
- Chapman E., Abdalla F. B., Harker G., Jelić V., Labropoulos P., Zaroubi S., Brentjens M. A., de Bruyn A. G., Koopmans L. V. E., 2012, MNRAS, 423, 2518
- Clifton T., Ferreira P. G., Padilla A., Skordis C., 2012, Phys. Rep., 513, 1
- Coles P., Lucchin F., 2002, *Cosmology: The Origin and Evolution of Cosmic Structure*, Second Edition
- Cooray A., 2006, Physical Review Letters, 97, 261301
- Copeland E. J., Sami M., Tsujikawa S., 2006, International Journal of Modern Physics D, 15, 1753
- Cropper M., Hoekstra H., Kitching T., Massey R., Amiaux J., Miller L., Mellier Y., Rhodes J., Rowe B., Pires S., Saxton C., Scaramella R., 2013, MNRAS, 431, 3103
- de Oliveira-Costa A., Tegmark M., Gaensler B. M., Jonas J., Landecker T. L., Reich P., 2008, MNRAS, 388, 247
- Dewdney P. E. F., 2013, SKA1 System Baseline Design, SKA Project Documentation
- Di Matteo T., Perna R., Abel T., Rees M. J., 2002, ApJ, 564, 576

- Dijkstra M., Gilfanov M., Loeb A., Sunyaev R., 2012, *MNRAS*, 421, 213
- Dijkstra M., Haiman Z., Loeb A., 2004, *ApJ*, 613, 646
- Dillon J. S., Liu A., Tegmark M., 2013, *Phys. Rev. D*, 87, 043005
- Dodelson S., 2003, *Modern cosmology*
- Eisenstein D. J., Hu W., 1998, *ApJ*, 496, 605
- Ewen H. I., Purcell E. M., 1951, *Nature*, 168, 356
- Field G. B., 1958, *Proceedings of the IRE*, 46, 240
- Font-Ribera A., Miralda-Escudé J., Arnau E., Carithers B., Lee K.-G., Noterdaeme P., Pâris I., Petitjean P., Rich J., Rollinde E., Ross N. P., Schneider D. P., White M., York D. G., 2012, *ArXiv e-prints*, 11, 59
- Friedrich M. M., Mellema G., Alvarez M. A., Shapiro P. R., Iliev I. T., 2011, in *Revista Mexicana de Astronomia y Astrofisica Conference Series Vol. 40 of Revista Mexicana de Astronomia y Astrofisica Conference Series, The Euler Characteristic as a Measure of the Topology of Cosmic Reionization*. pp 13–14
- Furlanetto S. R., Loeb A., 2004, *ApJ*, 611, 642
- Furlanetto S. R., Oh S. P., Briggs F. H., 2006, *Phys. Rep.*, 433, 181
- Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004, *ApJ*, 613, 1
- Gaensler B. M., Dickey J. M., McClure-Griffiths N. M., Green A. J., Wieringa M. H., Haynes R. F., 2001, *ApJ*, 549, 959
- Giocoli C., Meneghetti M., Metcalf R. B., Etori S., Moscardini L., 2014, *MNRAS*, 440, 1899
- Hamaker J. P., Bregman J. D., Sault R. J., 1996, *A&AS*, 117, 137
- Haverkorn M., Katgert P., de Bruyn A. G., 2004, *A&A*, 427, 169
- Hilbert S., Hartlap J., White S. D. M., Schneider P., 2009, *A&A*, 499, 31
- Hilbert S., Metcalf R. B., White S. D. M., 2007, *MNRAS*, 382, 1494
- Hirata C. M., Sigurdson K., 2007, *MNRAS*, 375, 1241
- Hoekstra H., 2013, *ArXiv e-prints*
- Hu W., Dodelson S., 2002, *ARA&A*, 40, 171
- Hu W., Okamoto T., 2002, *ApJ*, 574, 566
- Hu W., White M., 1997, *Phys. Rev. D*, 56, 596

- Jarvis M., Bacon D., Blake C., Brown M., Lindsay S., Raccanelli A., Santos M., Schwarz D. J., 2015, *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, p. 18
- Jelić V., Zaroubi S., Labropoulos P., Bernardi G., de Bruyn A. G., Koopmans L. V. E., 2010, *MNRAS*, 409, 1647
- Jelić V., Zaroubi S., Labropoulos P., Thomas R. M., Bernardi G., Brentjens M. A., de Bruyn A. G., Ciardi B., Harker G., Koopmans L. V. E., Pandey V. N., Schaye J., Yatawatta S., 2008, *MNRAS*, 389, 1319
- Jeong D., 2010, PhD thesis, University of Texas at Austin
- Jester S., Falcke H., 2009, 53, 1
- Joachimi B., Bridle S. L., 2010, *A&A*, 523, A1
- Joudaki S., Doré O., Ferramacho L., Kaplinghat M., Santos M. G., 2011, *Physical Review Letters*, 107, 131304
- Kaiser N., 1987, *MNRAS*, 227, 1
- Kaiser N., 1992, *ApJ*, 388, 272
- Kaiser N., Squires G., 1993, *ApJ*, 404, 441
- Kilbinger M., Fu L., Heymans C., Simpson E., Benjamin J., Erben T., Harnois-Déraps J., Hoekstra H., Hildebrandt H., Kitching T. D., Mellier Y., Miller L., Van Waerbeke L., Benabed K., et al. 2013, *MNRAS*, 430, 2200
- Kneib J.-P., Natarajan P., 2011, *A&A Rev.*, 19, 47
- Kochanek C. S., 2002, *ApJ*, 578, 25
- Kochanek C. S., Schechter P. L., 2004, *Measuring and Modeling the Universe*, p. 117
- Kovetz E. D., Kamionkowski M., 2013, *Phys. Rev. D*, 87, 063516
- Kuhlen M., Madau P., Montgomery R., 2006, *ApJ*, 637, L1
- Lee K.-G., Spergel D. N., 2011, *ApJ*, 734, 21
- Lewis A., 2005, *Phys. Rev. D*, 71, 083008
- Lewis A., Challinor A., 2006, *Phys. Rep.*, 429, 1
- Lidz A., Zahn O., McQuinn M., Zaldarriaga M., Dutta S., Hernquist L., 2007, *ApJ*, 659, 865
- Limber D. N., 1954, *ApJ*, 119, 655
- Liszt H., 2001, *A&A*, 371, 698

- Liu A., Tegmark M., 2011, *Phys. Rev. D*, 83, 103006
- Liu A., Tegmark M., 2012, *MNRAS*, 419, 3491
- Liu A., Tegmark M., Bowman J., Hewitt J., Zaldarriaga M., 2009, *MNRAS*, 398, 401
- Loeb A., Wyithe J. S. B., 2008, *Physical Review Letters*, 100, 161301
- Loeb A., Zaldarriaga M., 2004, *Physical Review Letters*, 92, 211301
- LSST Science Collaboration Abell P. A., Allison J., Anderson S. F., Andrew J. R., Angel J. R. P., Armus L., Arnett D., Asztalos S. J., Axelrod T. S., et al. 2009, *ArXiv e-prints*
- Lu T., Pen U.-L., 2008, *MNRAS*, 388, 1819
- McPherson A., 2015, *Report and Options for re-baselining of SKA-1*
- Mandel K. S., Zaldarriaga M., 2006, *ApJ*, 647, 719
- Mao Y., Shapiro P. R., Mellema G., Iliev I. T., Koda J., Ahn K., 2012, *MNRAS*, 422, 926
- Mao Y., Tegmark M., McQuinn M., Zaldarriaga M., Zahn O., 2008, *Phys. Rev. D*, 78, 023529
- Masui K. W., McDonald P., Pen U.-L., 2010, *Phys. Rev. D*, 81, 103527
- McQuinn M., Lidz A., Zahn O., Dutta S., Hernquist L., Zaldarriaga M., 2007, *MNRAS*, 377, 1043
- McQuinn M., Zahn O., Zaldarriaga M., Hernquist L., Furlanetto S. R., 2006, *ApJ*, 653, 815
- Meneghetti M., 2011
- Mesinger A., Furlanetto S., Cen R., 2011, *MNRAS*, 411, 955
- Metcalf R. B., Petkova M., 2014, *MNRAS*, 445, 1942
- Metcalf R. B., White S. D. M., 2007, *MNRAS*, 381, 447
- Metcalf R. B., White S. D. M., 2009, *MNRAS*, 394, 704
- Mirabel I. F., Dijkstra M., Laurent P., Loeb A., Pritchard J. R., 2011, *A&A*, 528, A149
- Moore D. F., Aguirre J. E., Parsons A. R., Jacobs D. C., Pober J. C., 2013, *ApJ*, 769, 154
- Morales M. F., 2005, *ApJ*, 619, 678
- Morales M. F., Bowman J. D., Hewitt J. N., 2006, *ApJ*, 648, 767

- Morales M. F., Wyithe J. S. B., 2010, *ARA&A*, 48, 127
- Naoz S., Barkana R., 2005, *MNRAS*, 362, 1047
- Narayan R., Bartelmann M., 1996, *ArXiv Astrophysics e-prints*
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Paciga G., Chang T.-C., Gupta Y., Nityanada R., Odegova J., Pen U.-L., Peterson J. B., Roy J., Sigurdson K., 2011, *MNRAS*, 413, 1174
- Padmanabhan H., Choudhury T. R., Refregier A., 2015, *MNRAS*, 447, 3745
- Parsons A., Pober J., McQuinn M., Jacobs D., Aguirre J., 2012, *ApJ*, 753, 81
- Partl A. M., Maselli A., Ciardi B., Ferrara A., Müller V., 2011, *MNRAS*, 414, 428
- Peacock J. A., 1999, *Cosmological Physics*
- Peacock J. A., Dodds S. J., 1996, *MNRAS*, 280, L19
- Pearson R., Sherwin B., Lewis A., 2014, *Phys. Rev. D*, 90, 023539
- Perlmutter S., Aldering G., Goldhaber G., Knop R. A., Nugent P., Castro P. G., Deustua S., Fabbro S., Goobar A., Groom D. E., Hook I. M., Kim A. G., Project T. S. C., et al. 1999, *ApJ*, 517, 565
- Péroux C., McMahon R. G., Storrie-Lombardi L. J., Irwin M. J., 2003, *MNRAS*, 346, 1103
- Petkova M., Metcalf R. B., Giocoli C., 2014, *MNRAS*, 445, 1954
- Petrovic N., Oh S. P., 2011, *MNRAS*, 413, 2103
- Pillepich A., Porciani C., Matarrese S., 2007, *ApJ*, 662, 1
- Planck Collaboration Ade P. A. R., Aghanim N., Arnaud M., Arroja F., Ashdown M., Aumont J., Baccigalupi C., Ballardini M., Banday A. J., et al. 2015, *ArXiv e-prints*
- Planck Collaboration Ade P. A. R., Aghanim N., Arnaud M., Ashdown M., Aumont J., Baccigalupi C., Banday A. J., Barreiro R. B., Bartlett J. G., et al. 2015, *ArXiv e-prints*
- Plaszczynski S., Lavabre A., Perotto L., Starck J.-L., 2012, *A&A*, 544, A27
- Pober J. C., Parsons A. R., DeBoer D. R., McDonald P., McQuinn M., Aguirre J. E., Ali Z., Bradley R. F., Chang T.-C., Morales M. F., 2013, *AJ*, 145, 65
- Pourtsidou A., Bacon D., Crittenden R., Metcalf R. B., 2015, *ArXiv e-prints*
- Pourtsidou A., Metcalf R. B., 2014, *MNRAS*, 439, L36
- Pourtsidou A., Metcalf R. B., 2015, *MNRAS*, 448, 2368

- Press W. H., Schechter P., 1974, *ApJ*, 187, 425
- Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 2002, *Numerical recipes in C++ : the art of scientific computing*
- Pritchard J., Ichiki K., Mesinger A., Metcalf R. B., Pourtsidou A., Santos M., Abdalla F. B., Chang T. C., Chen X., Weller J., Zaroubi S., 2015, *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, p. 12
- Pritchard J. R., Furlanetto S. R., 2007, *MNRAS*, 376, 1680
- Pritchard J. R., Loeb A., 2008, *Phys. Rev. D*, 78, 103511
- Pritchard J. R., Loeb A., 2012, *Reports on Progress in Physics*, 75, 086901
- Riess A. G., Filippenko A. V., Challis P., Clocchiatti A., Diercks A., Garnavich P. M., Gilliland R. L., Hogan C. J., Jha S., Kirshner R. P., Leibundgut B., Phillips M. M., Reiss D., Schmidt B. P., Schommer R. A., Smith R. C., Spyromilio J., et al. 1998, *AJ*, 116, 1009
- Rowan-Robinson M., 2004, *Cosmology*
- Rubin V. C., Coyne G. V., 1988, *Large-scale motions in the universe*
- Rybicki G. B., 2006, *ApJ*, 647, 709
- Rybicki G. B., Lightman A. P., 1986, *Radiative Processes in Astrophysics*
- Santos M., Alonso D., Bull P., Silva M. B., Yahya S., 2015, *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, p. 21
- Santos M., Ferramacho L., Silva M., Amblard A., Cooray A., , 2010, *SimFast21: Simulation of the Cosmological 21cm Signal*, *Astrophysics Source Code Library*
- Santos M. G., Bull P., Alonso D., Camera S., Ferreira P. G., Bernardi G., Maartens R., Viel M., Villaescusa-Navarro F., Abdalla F. B., Jarvis M., Metcalf R. B., Pourtsidou A., Wolz L., 2015, *ArXiv e-prints*
- Sault R. J., Hamaker J. P., Bregman J. D., 1996, *A&AS*, 117, 149
- Schäfer B. M., Heisenberg L., Kalovidouris A. F., Bacon D. J., 2012, *MNRAS*, 420, 455
- Schneider P., 2015, *Extragalactic Astronomy and Cosmology: An Introduction*
- Schneider P., Ehlers J., Falco E. E., 1992, *Gravitational Lenses*
- Schneider P., van Waerbeke L., Kilbinger M., Mellier Y., 2002, *A&A*, 396, 1
- Seitz S., Schneider P., Ehlers J., 1994, *Classical and Quantum Gravity*, 11, 2345
- Sereno M., Giocoli C., Etori S., Moscardini L., 2014, *ArXiv e-prints*

- Shaw J. R., Sigurdson K., Pen U.-L., Stebbins A., Sitwell M., 2014, *ApJ*, 781, 57
- Shaw J. R., Sigurdson K., Sitwell M., Stebbins A., Pen U.-L., 2015, *Phys. Rev. D*, 91, 083514
- Simard-Normandin M., Kronberg P. P., Button S., 1981, *ApJS*, 45, 97
- Smirnov O. M., 2011, *A&A*, 531, A159
- Smoot G. F., Debono I., 2014, *ArXiv e-prints*
- Squires G., Kaiser N., 1996, *ApJ*, 473, 65
- Sun X. H., Reich W., Han J. L., Reich P., Wielebinski R., Wang C., Müller P., 2011, *A&A*, 527, A74
- Switzer E. R., Masui K. W., Bandura K., Calin L.-M., Chang T.-C., Chen X.-L., Li Y.-C., Liao Y.-W., Natarajan A., Pen U.-L., Peterson J. B., Shaw J. R., Voytek T. C., 2013, *MNRAS*, 434, L46
- Takada M., Jain B., 2004, *MNRAS*, 348, 897
- Tassev S., Zaldarriaga M., Eisenstein D. J., 2013, *ArXiv e-prints*, 6, 36
- Taylor A. R., Stil J. M., Sunstrum C., 2009, *ApJ*, 702, 1230
- Tegmark M., Zaldarriaga M., 2002, *Phys. Rev. D*, 66, 103508
- Thomas R. M., Zaroubi S., Ciardi B., Pawlik A. H., Labropoulos P., Jelić V., Bernardi G., Brentjens M. A., de Bruyn A. G., Harker G. J. A., Koopmans L. V. E., Mellema G., Pandey V. N., Schaye J., Yatawatta S., 2009, *MNRAS*, 393, 32
- Trac H., Cen R., Loeb A., 2008, *ApJ*, 689, L81
- Trac H. Y., Gnedin N. Y., 2011, *Advanced Science Letters*, 4, 228
- Tucci M., Martínez-González E., Toffolatti L., González-Nuevo J., De Zotti G., 2004, *MNRAS*, 349, 1267
- Van Waerbeke L., Mellier Y., 2003, *ArXiv Astrophysics e-prints*
- Villaescusa-Navarro F., Viel M., Datta K. K., Choudhury T. R., 2014, *ArXiv e-prints*, 9, 50
- Visbal E., Loeb A., Wyithe S., 2009, 10, 30
- Waelkens A., Jaffe T., Reinecke M., Kitaura F. S., Enßlin T. A., 2009, *A&A*, 495, 697
- Wang X., Hu W., 2006, *ApJ*, 643, 585
- Wang X., Tegmark M., Santos M. G., Knox L., 2006, *ApJ*, 650, 529

- White M., Tinker J. L., McBride C. K., 2014, MNRAS, 437, 2594
- Wolz L., Abdalla F. B., Alonso D., Blake C., Bull P., Chang T. C., Ferreira P., Kuo C. Y., Santos M., Shaw J. R., 2015, *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, p. 35
- Wolz L., Abdalla F. B., Blake C., Shaw J. R., Chapman E., Rawlings S., 2014, MNRAS, 441, 3271
- Wyithe J. S. B., Loeb A., 2009, MNRAS, 397, 1926
- Wyithe S., Loeb A., 2007, ArXiv e-prints
- Yang X., Zhang P., 2011, MNRAS, 415, 3485
- Zahn O., Zaldarriaga M., 2006, ApJ, 653, 922
- Zaldarriaga M., Furlanetto S. R., Hernquist L., 2004, ApJ, 608, 622
- Zaroubi S., 2013, in Wiklind T., Mobasher B., Bromm V., eds, *Astrophysics and Space Science Library Vol. 396 of Astrophysics and Space Science Library, The Epoch of Reionization*. p. 45
- Zhang P., Pen U.-L., 2005, Physical Review Letters, 95, 241302
- Zhang P., Pen U.-L., 2006, MNRAS, 367, 169
- Zwaan M. A., Staveley-Smith L., Koribalski B. S., Henning P. A., Kilborn V. A., Ryder S. D., Barnes D. G., Bhathal R., Boyce P. J., de Blok W. J. G., Disney M. J., Drinkwater M. J., Ekers R. D., Freeman K. C., Gibson B. K., et al. 2003, AJ, 125, 2842
- Zwicky F., 1933, Helvetica Physica Acta, 6, 110